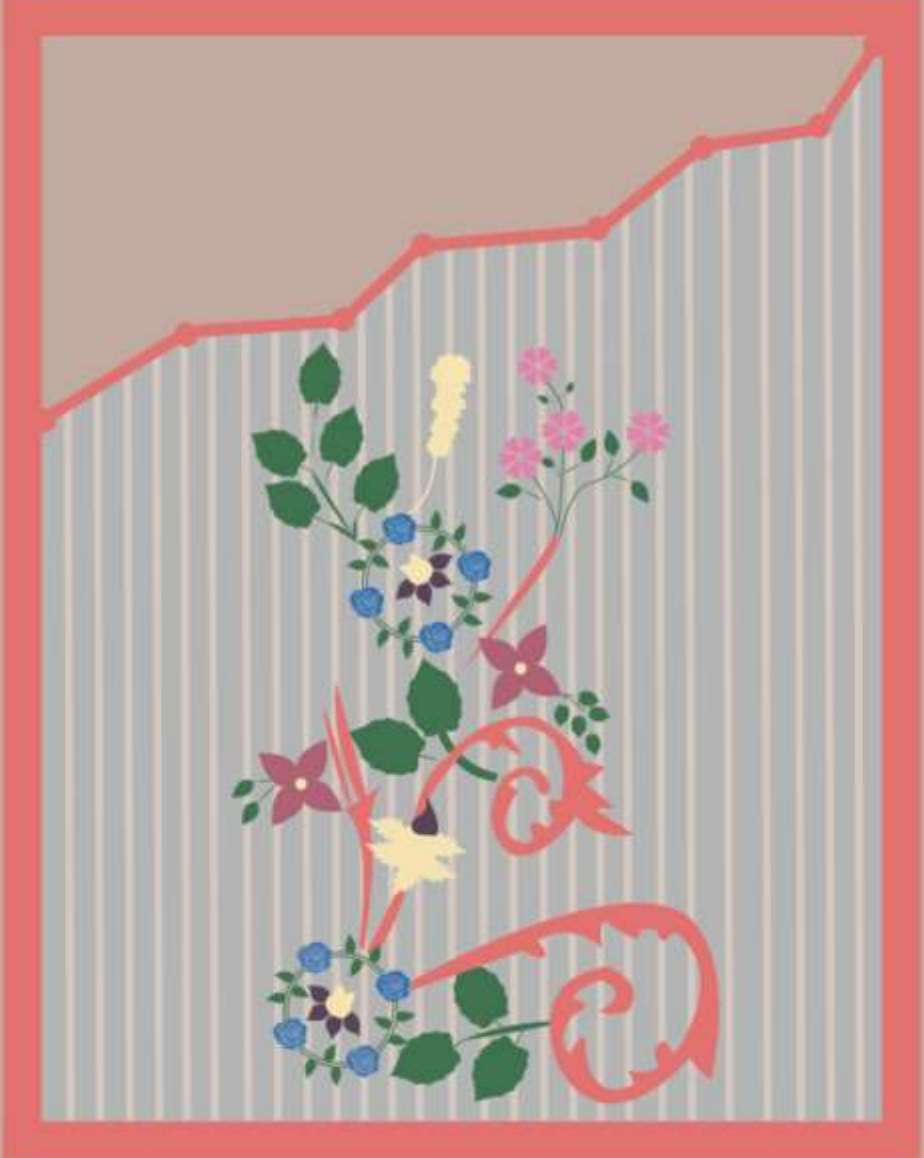


SAĞLIK ÇALIŞANLARI İÇİN ARAŞTIRMA VE PRATİK İSTATİSTİK ÖRNEK PROBLEMLER VE SPSS ÇÖZÜMLERİ

Editör : Prof. Dr. Zekeriya AKTÜRK



Bu kitap “Üret ve Paylaş!” prensibiyle, bilim insanlarına faydalı olmak amacıyla yazılmıştır.

Siz de biyoistatistik bilginizi geliştirerek insanlara daha faydalı olabilirsiniz!

“İnsanların en hayırlısı insanlara faydalı olanıdır.”

Hız. Muhammed

“İnsanların seni en çok sevdiği zaman, onların işine en çok yaradığın zamandır.”

Heinrich Karl Bukowski (Alman Şair ve Yazar, 1920-1994)

Kapak



Trabzon Dokuması

El yapımı tezgahta Şükriye AKTÜRK tarafından 1975 yılında dokunmuş, kumaşın üzerine Hatice DOLDURAN tarafından 2012 yılında el nakışı işlenmiştir.

Bu kumaş türü Karadeniz’de 1980’li yıllara kadar ev tipi tezgahlarda dokunmaktaydı. Kumaşın ipi elde işlenmiş ve örmeye hazırlanmıştır. Yün ipliğin daha dayanıklı olması için araya kendirden yapılan iplikler eklenmiştir.

Dokunan kumaşlar gömlek, atlet, havlu ve çarşaf gibi eşya üretiminde kullanılmaktaydı.

Arka Kapak



Karadeniz Peşgiri (Havlusı)

Bu kumaş da Trabzon’da el tezgahında Şükriye AKTÜRK tarafından dokunmuştur.

İplik koyun yününden tamamen elde hazırlanmış, boya da elde yapılmıştır.

Peşgir üzerine Hatice DOLDURAN tarafından Türk motifleri boyanmıştır.

Kapak çizimi: Abdullah Enis MANDAL

Sağlık Çalışanları İçin Araştırma ve Pratik İstatistik Örnek Problemler ve SPSS Çözümleri

Editör: Prof. Dr. Zekeriya AKTÜRK

Yazarlar: Prof. Dr. Zekeriya AKTÜRK¹, Prof. Dr. Hamit ACEMOĞLU²,
Doç. Dr. Memet IŞIK³, Prof. Dr. Fatih AKÇAY⁴

¹ Şifa Üniversitesi Tıp Fakültesi Aile Hekimliği Anabilim Dalı/Biyostatistik ve Tıp Bilişimi Anabilim Dalı, İzmir

² Atatürk Üniversitesi Tıp Fakültesi Tıp Eğitimi Anabilim Dalı/Biyostatistik ve Tıp Bilişimi Anabilim Dalı, Erzurum

³ Atatürk Üniversitesi Tıp Fakültesi Aile Hekimliği Anabilim Dalı, Erzurum

⁴ Atatürk Üniversitesi Tıp Fakültesi Tıbbi Biyokimya Anabilim Dalı, Erzurum

Aktürk Z, Acemoğlu H, Işık M, Akçay F. Sağlık Çalışanları İçin Araştırma ve Pratik İstatistik. 3. Baskı. Anadolu Matbaası. İstanbul 2015

ISBN: 978-975-00631-1-4

Basım Yeri: Anadolu Ofset, Cağaloğlu, İstanbul. Tel: 0532 5498961.

<http://www.aile.net>, zekeriya.akturk@gmail.com

Teşekkür

Kitabın hazırlanma aşamasında, dizgisinde ve tashihindeki desteklerinden dolayı Doç. Dr. Turan SET ve Hemş. Semra AKTÜRK'e, baskıdaki desteklerinden ötürü Aile Hekimliği Okulu Derneği'ne teşekkür ederiz.

İçindekiler

İçindekiler	9
Önsöz	19
Üçüncü Baskı İçin Önsöz	21
Kitap Hakkında – Genel Amaç ve Hedefler	23
Bölüm 1: Giriş	27
Bölüm 2: İstatistiksel Testler ve Kullanım Yerleri – Akış Şemaları.....	41
Örnek Problemler.....	44
Bölüm 3: Veri Çeşitleri.....	50
Türetilmiş Veriler.....	52
Durdurulmuş (Censored) Veri	53
Bölüm 4: Veri Girişi	54
Veri Girişi Yöntemleri	54
Veri Girişinin Planlanması	55
Kategorik Veriler	56
Nümerik Veriler.....	56
Anketlerin Numaralandırılması.....	57
Boş Verilerin Kodlanması.....	57
Alıştırma 1	58
Bölüm 5: Hata Ayıklama ve Uç Değerler	62
Hata arama	62
Eksik veriler (missing) varsa ne yapacağız?	64
Uç değerler.....	64
Alıştırma 2	66
Bölüm 6: Verilerin Grafiklerle Gösterilmesi	68
Frekans dağılımının şekli	71
Bölüm 7: Verilerin Tablolarla Gösterilmesi	76
Genel prensipler	76
SPSS’de tablo ayarları	79
SPSS’de özel tablolar.....	80
Bölüm 8: Merkezi Eğilim (Yığılma) Ölçütleri	82
Aritmetik ortalama	82
Ortanca.....	83
Mod.....	83

Geometrik ortalama	83
Ağırlıklı ortalama.....	86
Diğer Merkezi Eğilim Göstergeleri.....	87
Alıştırma 3	90
Bölüm 9: Yaygınlık Ölçütleri	91
Aralık (range).....	91
Persantil aralıkları	91
Varyans	92
Standart sapma	93
SPSS ile hesaplama.....	93
Birey içi ve bireyler arası farklılıklar	95
Bölüm 10: Teorik Dağılımlar: Normal Dağılım	97
Olasılık hesapları	97
Olasılık kuralları	98
Olasılık dağılımlarının teorisi	99
Normal dağılım (Gauss dağılımı)	101
Standart Normal Dağılım	102
Alıştırma 4	102
Bölüm 11: Teorik Dağılımlar: Diğer Dağılımlar	105
Diğer sürekli teorik dağılımlar	105
Tam sayılı teorik dağılımlar	108
Bölüm 12: Veri Dönüştürme.....	111
Veri dönüşümü nasıl yapılır?	112
Tipik dönüşümler	112
Bölüm 13: Örneklem ve Örneklem Dağılımları.....	121
Örneklemenin temsil edici olması	121
Örneklem varyasyonu	122
Ortalamanın örneklem dağılımı	122
Standart sapma mı standart hata mı?.....	123
Orantının örneklem dağılımı	123
Alıştırma 5	125
Bölüm 14: Güven Aralıkları	127
Ortalamanın güven aralığı.....	127
Orantılar için güven aralığı	129

Ortanca için güven aralığı	130
Güven aralıklarının yorumlanması.....	130
Serbestlik derecesi	131
Örnek	131
Farkların güven aralığı	133
Odds oranları için güven aralığı.....	134
Relative Deviate formülünün farkların güven aralığına uygulanması	135
Bölüm 15: Araştırma Tasarımı – I.....	137
Deneysel / gözlemsel araştırmalar	138
Araştırma çeşitleri.....	138
Nedensellik	143
Kesitsel / longitudinal araştırmalar	144
Kontrol grupları	144
Hatalar (bias).....	144
Alistırma 6	146
Bölüm 16: Araştırma Tasarımı – II.....	149
Varyasyon	149
Kopyalama (Replikasyon).....	149
Örneklem büyüklüğü	149
Özel araştırma tasarımları	150
Bölüm 17: Klinik Araştırmalar	153
Karşılaştırmalar.....	155
Araştırmanın çıktıları (sonlanım noktaları, endpoints, outcomes)	155
Gruplara ayırma	155
Körleme	156
Etik.....	156
Araştırma protokolü	156
Bölüm 18: Önceki Konuların Özeti	161
Alistırma 7	166
Bölüm 19: Kohort Araştırmaları	167
Kohortun seçilmesi	167
Kohort araştırmaların analizi	169
Kohort araştırmaların olumlu yönleri.....	171
Kohort araştırmaların olumsuz yönleri	171

Klinik kohortlar.....	171
Bölüm 20: Vaka Kontrol Çalışmaları	176
Vakaların seçilmesi.....	176
Kontrol grubunun seçilmesi	177
Eşleştirme.....	177
Eşleştirilmemiş vaka kontrol araştırmalarının analizi	178
Alıştırma 8	181
Bölüm 21: Hipotez Testleri.....	182
Hipotez testlerinde genel yaklaşım	182
H_0 ve H_1 hipotezi	182
İki yönlü ve tek yönlü hipotez.....	183
p değerinin yorumlanması.....	184
Parametrik ve parametrik olmayan testler	185
Hipotez testi mi güven aralığı mı?	186
Bölüm 22: Hipotez Testlerinde Hatalar	188
Karar verme	188
Yanlış karar verme	189
Araştırmanın gücü ve ilişkili faktörler	190
Birden çok hipotezi test etme.....	191
Alıştırma 9	195
Bölüm 23: Nümerik Veri Tek Grup (Tek Örneklemde t Testi)	197
Tek örneklemde t-testi varsayımları.....	197
Örnek uygulama.....	197
Örneğin SPSS ile yapılması	200
Varsayımlar karşılanamıyorsa.....	200
İşaret testi.....	200
Alıştırma 10	203
Bölüm 24: Nümerik Veri İki Bağımlı Grup (Bağımlı Örneklemelerde t Testi, Wilcoxon Testi).....	207
Bağımlı örneklemelerde t-testi.....	208
Varsayımları.....	208
Testin mantığı	208
Ek açıklama.....	208
Örnek uygulama.....	208
Örneğin SPSS ile yapılması	213

Varsayımlar karşılanamıyorsa.....	213
Wilcoxon işaretli sıra testi.....	213
Mantığı.....	213
Bir örnek üzerinden Wilcoxon işaretli sıra testi.....	214
Wilcoxon işaretli sıra testinin SPSS ile yapılması	218
Bölüm 25: Nümerik Veri İki Bağımsız Grup (Student t Testi, Mann-Whitney U Testi)	
.....	221
Bağımsız örneklerde t testi.....	222
Varsayımlar karşılanamıyorsa.....	228
Bölüm 26: Nümerik Veri İki'den Fazla Bağımlı Grup (Tekrarlayan Ölçümlerde ANOVA, Friedman Varyans Analizi)	
.....	231
Senaryo 1, Tek faktör (grup içi karşılaştırma)	231
Varsayımları.....	232
Testin mantığı	232
Örnek uygulama 1.....	232
General Linear Model	236
Varsayımlar karşılanamıyorsa.....	239
Senaryo 2, İki faktör (grup içi ve gruplar arası karşılaştırma)	240
Örnek uygulama 2.....	240
General Linear Model	244
Varsayımlar karşılanamıyorsa.....	247
Bölüm 27: Nümerik Veri İki'den Fazla Bağımsız Grup (Tek Yönlü ANOVA, Kruskal Wallis Testi)	
.....	250
ANOVA.....	250
Tek yönlü ANOVA ve lineerite testi	255
Varsayımlar karşılanamıyorsa.....	257
Bölüm 28: Nümerik Veri İki Kategorik Bağımsız Değişken (İki Yönlü ANOVA)....	262
İki yönlü ANOVA.....	262
Varsayımlar karşılanamıyorsa.....	269
Bölüm 29: Kategorik Veri Tek Grup (Tek Orantı Testi)	270
Tek orantı testi (test of a single proportion).....	270
Eğer bir tercih durumunu araştırıyorsak (Binomiyal test).....	272
Bölüm 30: Kategorik Veri İki Bağımlı Grup (McNemar Testi)	276
McNemar testi.....	276

Bölüm 31: İkidenden fazla nominal kategorisi olan bağımlı iki ölçüm (McNemar Bowker testi)	280
McNemar-Bowker testi	280
Bölüm 31: Kategorik Veri İki Bağımlı Grup (Kappa Uyum Analizi)	285
Basit uyum katsayısı eşitliği	286
Cohen'in kappa katsayısı (κ) eşitliği	286
Kappa katsayısına göre uyumun yorumlanması	287
Örneğin SPSS ile yapılması	287
Alıştırma 11	290
Çözüm için bakınız s. 470	290
Bölüm 32: Kategorik Veri İki Bağımsız Grup (Ki Kare Analizi)	293
Ki Kare (χ^2) testi	293
Bölüm 33: Kategorik Veri İkidenden Fazla Bağımsız Grup (Ki Kare Analizi)	299
Çoklu gruplarda Ki Kare (χ^2) testi	299
Ki Kare Trend testi	305
Ham veriler olmadan SPSS ile Ki Kare analizi	308
Bölüm 34: Kategorik Veri İkidenden Fazla Bağımlı Grup (Cochran Q Analizi)	311
Bölüm 35: Kategorik Veri İkidenden Fazla Bağımlı Kategori (Marginal Homogeneity Testi)	315
Bölüm 36: Kategorik Veri İkidenden Fazla Bağımsız Kategori (Ki Kare)	319
Bölüm 37: Kategorik Veri İkidenden Fazla Kategori – 1 Grup	321
Bölüm 38: İki Nümerik Bağımlı Değişken (Bland-Altman Analizi)	325
Bland-Altman analizinin kullanım yeri	325
SPSS ile Bland-Altman analizinin yapılması	326
Alıştırma 12	332
Bölüm 39: İki Nümerik Bağımsız Değişken (Korelasyon)	333
Pearson korelasyon katsayısı	333
Spearman korelasyon katsayısı	337
Bölüm 40: Lineer Regresyon	339
Varsayımlar	339
Modelin oluşturulması	340
SPSS ile uygulama	340
Alıştırma 13	347
Bölüm 41: Ortalamaya Doğru Gerileme	349
Bölüm 42: Lojistik Regresyon	351

Lojistik regresyon eşitliği.....	352
Modelin oluşturulması	352
Modelin uygunluğunun test edilmesi	353
Örnek uygulama.....	353
Alıştırma 14	359
Alıştırma 15	360
Bölüm 43: Kaplan Meier Sağkalım Analizi.....	361
Tanımlar.....	362
Varsayımlar.....	363
Sağkalım eğrilerinin karşılaştırılması	364
SPSS’te Kaplan –Meier Sağkalım Analizi.....	364
Alıştırma 16	371
Bölüm 44: Cox Regresyon Analizi	373
Zarar Oranı (Hazard Ratio, HR).....	374
SPSS’te Kaplan – Cox regresyon Analizi	374
Alıştırma 17	378
Bölüm 45: Nümerik Veri Uyum Analizi (Sınıf İçi Korelasyon).....	379
Bölüm 46: Örneklem Hesabı	382
Tanımlayıcı istatistikler için örneklem büyüklüğü formülleri.....	384
Hipotez testleri için örneklem büyüklüğünün hesaplanması.....	386
Alıştırma 18	390
Bölüm 47: Duyarlılık ve Özgüllük (ROC) Analizi	393
Duyarlılık ve Özgüllük Kavramları	393
PPV ve NPV Kavramları	394
ROC Analizi	395
Hangi Kesme Değeri?	395
Örnek uygulama.....	396
Bölüm 48: Geçerlilik ve Güvenilirlik Analizi.....	399
Geçerlilik ve Güvenilirlik Kavramları	400
Güvenilirlik Çeşitleri	401
Geçerlilik Çeşitleri	402
Bölüm 49: Tanı Testleri ve Sağlık Hizmetlerindeki Değeri	405
Tanı Testlerinin Özellikleri.....	406
Fazla Test İstenmesinin Sonuçları	407

Bayes Teoremi ve Tanı Testleri	408
Bölüm 50: Biyoistatistik Paket Programları	413
SPSS	414
Minitab.....	415
İstatistik Programlarının Karşılaştırılması	415
Bölüm 51: Rapor Yazma	419
Kullanılan yazı tipi.....	419
Sayfa Kenar Boşlukları	420
Satır Aralıkları	420
Başlık Düzeni.....	420
Dipnot ve Son Not	420
Sayfa Numaraları	421
Verilerin Tablolar ve Şekillerle Gösterilmesi	421
İstatistiksel Gösterimler	421
Araştırma Raporu Bölümleri.....	422
Bölüm 52: Araştırma Fonları	425
Üniversitelerin Bilimsel Araştırma Proje (BAP) destekleri	425
TÜBİTAK Araştırma Destek Programları	426
TÜBİTAK Bilim İnsanı Destekleme Programları	430
Devlet Planlama Teşkilatı (DPT).....	431
Bakanlık Araştırma Fonları.....	431
Kalkınma Ajansları Destek Programları	432
Uluslararası Fon Kuruluşları.....	432
Bölüm 53: Yazarlık Hakkı	435
Yazarlık Ölçütleri.....	435
Onursal (Hediye) Yazarlık	437
Yazar Paslaşması	437
Hayalet Yazarlık	437
Yazarlık Sırası nasıl olmalıdır?	438
Yazarlık Sorumlulukları.....	438
Teşekkür (Acknowledgment) kime yapılmalı.....	439
Bölüm 54: Dergi Seçimi	440
Bölüm 55: Kendini Değerlendirme Çalışması	441
Bölüm 56: Kendini Değerlendirme Çalışması Çözümleri	445

Ek 1: İstatistik tablolar.....	458
Z tablosu (Standart normal dağılım).....	458
t tablosu.....	459
Ki Kare tablosu	460
F tablosu.....	461
Pearson korelasyon katsayısı	462
Spearman korelasyon katsayısı	463
Ek 2: Alıştırmaların Çözümleri.....	464
Çözüm 1 (Sayfa 59)	464
Çözüm 2 (Sayfa 67)	466
Çözüm 3 (Sayfa 91)	467
Çözüm 4 (Sayfa 103)	470
Çözüm 5 (Sayfa 126)	470
Çözüm 6 (Sayfa 147)	472
Çözüm 7 (Sayfa 167)	473
Çözüm 8 (Sayfa 181)	474
Çözüm 9 (Sayfa 193)	475
Çözüm 10 (Sayfa 202)	475
Çözüm 11 (Sayfa 283)	477
Çözüm 12 (Sayfa 325)	477
Çözüm 13 (Sayfa 340)	481
Çözüm 14 (Sayfa 352)	484
Çözüm 15 (Sayfa 353)	485
Çözüm 16 (Sayfa 364)	486
Çözüm 17 (Sayfa 371)	487
Çözüm 18 (Sayfa 383)	489
Kaynaklar.....	493
İndeks	495

Önsöz

İnsan sağlığının önemli belirleyicilerinden birisi de tıptaki gelişmelerdir. Tıbbi gelişmelerden beklenen, başarılı koruyucu hizmetler, tanı ve tedavi yöntemlerini kullanarak geniş halk tabakalarına mümkün olan en uygun sağlık hizmetinin ulaştırılmasıdır. Bu hizmetin verilmesi için mezuniyet sonrası eğitimler ve sürekli tıp eğitimleri üzerinde sıklıkla durulmakla birlikte bu eğitimleri planlayacak, gerekli araştırmaları yapacak ve insan gücünü yetiştirecek akademik personelin yetiştirilmesi konusundaki çabalar da artırılmalıdır. Sağlık araştırmacıları genellikle üniversitelerde kıdemli eğiticilerin yanında usta-çırak ilişkisiyle veya kendi deneyim ve çabalarıyla öğrenerek yetişmektedir. Özellikle önemli bir iş gücü açığı ve eğitilmiş doktor, dolayısıyla akademisyen ihtiyacı olan birinci basamakta araştırma teknikleri ve istatistik bilgisine ihtiyaç vardır.

Geriye dönüp baktığımda, bu kitabın temellerinin akademisyenliğe ilk niyet ettiğim 1994 yıllarında Doç. Dr. Mehmet AKIN ile asistanlık eğitimim sırasında yaptığımız çalışmalarda atıldığını görüyorum. 1999 yılında Trakya Üniversitesi'nde akademisyenliğe başladıktan sonra Prof. Dr. Nezih DAĞDEVİREN hocamız ve iş arkadaşarımla yaptığımız çalışmaların da çok katkısı oldu. Araştırma ve istatistik konusunda yeterli kaynak bulamamamız ve/veya var olan kaynaklara ulaşamamamız kendimizi geliştirmemiz konusunda bizi zorladı. 2005 yılından sonra önceleri internet üzerinden verdiğimiz kurslar, daha sonra da Atatürk Üniversitesi'nde devam ettirdiğimiz sınıf dersleri bu çalışmayı geliştirmemizi sağladı ve Prof. Dr. Hamit ACEMOĞLU ile birlikte, ders notları olarak biriken materyalin kitap olarak basılmasını kararlaştırdık.

Dileğimiz, bilginin paylaşılması ve paylaşıldıkça da büyümesi, araştırma yapmak isteyen sağlık profesyonellerinin işlerinin kolaylaşmasıdır. Bu nedenle kitabımızın elektronik formunun kullanıma açık olmasını kararlaştırdık. Bununla birlikte, basılı nüshasını satın alarak destek olmak isterseniz bundan memnuniyet duyarız.

Kitabın erişime açık olmasının kendisini geliştirmek isteyen sağlık profesyonellerine katkıda bulunacağı, giderek iyileştirme ve yaygınlaştırma sayesinde daha geniş kitlelere ulaşılması, oluşturulacak olumlu rekabet ortamında daha iyi kitapların geliştirilip meslektaşlarımıza sunulması umulmaktadır. Bu açıdan kitabın daha sonraki baskıları için destek ve önerileriniz memnuniyetle kabul edilecektir.

İyi okumalar ve başarılı bilimsel çalışmalar diliyoruz.

(Eylül 2010)

Prof. Dr. Zekeriya AKTÜRK

Üçüncü Baskı İçin Önsöz

İnsan sürekli bir tekamül ve gelişim içerisinde. Bir taraftan kurslarımıza ve derslerimize devam ederken, bir taraftan da katılımcıların ihtiyaçları ve kendi eksiklerimizden hareketle gelişim çabası içerisindeyiz. Öncelikle sağlık çalışanlarına olmak üzere tüm akademisyenlere yönelik yaptığımız seminerler ve kurslar sonucunda ortaya çıkan kitabımız eğitim ve danışmanlık faaliyetlerimizle olgunlaştı. Bunun için tüm kursiyerlerimize ve danışmanlık alanlara teşekkür ederiz.

Birinci baskıda tespit ettiğimiz bazı hataları ikinci baskıda düzeltmiştik. Yine de göze çarpan hatalarımızın olduğunu fark ettik. İngilizce bir deyimde “*to err is human; to forgive, divine*” yani “*hata yapmak insana, affetmek Tanrı'ya aittir*” (Alexander Pope) denmektedir. İkinci baskıdaki hatalarımızı düzelttiğimizi düşünüyoruz ama biliyoruz ki, hata payımız her zaman olacaktır...

Bu baskıda ayrıca Lineer regresyon, Kaplan Meier sağ kalım analizi, Cox regresyon analizi, Kappa uyum analizi, Bland-Altman analizi, Tanı testlerinin sağlık hizmetlerindeki yeri, Biyoistatistik paket programları, Bilimsel rapor yazma, Araştırma fonları ve Yazarlık hakkı konularını kitaba ekledik. Bu konuların da okuyucunun ilgisini çekeceğini ve önemli bir ihtiyacı karşılayacağını düşünüyoruz.

Bilgiyi yayma gayretimiz hızını artırarak devam ediyor. 2014-2015 içerisinde Atatürk Üniversitesi'nin “İnternet Üzerinden Herkese Açık Kurs-İHAK” uygulaması (<http://atademix.atauni.edu.tr/>) çerçevesinde ücretsiz kurslar düzenledik. Kitabımızın üçüncü baskısının da öncekiler gibi internet üzerinden erişilebilir olmasını hedefliyoruz.

Çalışmamızın geliştirilmesinde katkıda bulunan tüm katılımcılara teşekkür ederiz. Burada kitabımızın geliştirilmesinde eleştiri ve görüşleriyle katkıda bulunan, uzaktan eğitim kursumuza katılmış kursiyerlerimizden bazı alıntılar da paylaşmak istiyoruz:

“Öncelikle hazırladığınız biyoistatistik kursunun şimdiye kadar katıldığım en donanımlı, en yenilikçi, en ilk kurs olduğunu ve diğer üniversiteler ve eğitim kurumlarına örnek olması gerektiğini vurgulamak isterim. Elinize sağlık. Bu ilkin sayenizde gerçekleşmesi de ayrı gurur vericidir (EAA-Trabzon).”

“Sizin 10. Biyoistatistik kursunuzu Youtube vasıtası ile izledim. Haddim olmayarak işinize yaraması ümidi ile kurs programınızı izlerken aklıma gelen önerilerimi madde madde sıraladım. Bu mesajı yazmamdaki amacım; bir ücret talep etmeden bilgisini cömertçe paylaşan sizin gibi değerli bir insanın amacına küçükte olsa katkı sağlayabilme ümididir (HN-Adana)”

“Bu mesajı yazmamdaki en önemli amaç bu denli kaliteli bir kursun Türkiye’de bir örneğinin daha olmamasıdır. Umarım ömrünüz uzun ve sağlıklı olursunuz. İstatistik öğrenmek isteyen bilim insanlarına Türkçe sunulmuş yüksek lisans eğitimi kalitesinde bu kursun her daim ulaşılabilir olması adına katkıda bulunmak istedim.”

“Sayın Hocam istatistiği çok seven biri olarak eğitiminizin altın fırsat olduğunu düşünüyorum. Mesaj listemdeki tüm arkadaşlarımı ve öğrencilerimi de yönlendirdim (ÇY-Gümüşhane).”

“Çok güzel, yararlı bir kurs. İş dışındaki zamanımızı olanaklar dahilinde bu tür kurslarla kendimizi geliştirmek için kullanmayı arzu ediyorum (Anonim).”

“Çok zor bir konunun basite indirgenerek anlaşılır bir şekilde sunulmuş olmasını beğendim (Anonim).”

“Ben genellikle zor beğenirim. Beğenmediğim derse veya konuya başımı çevirip bakmam. Ancak Hocamın derslerini iki yıldır severek takip ediyorum. Bir çok dostlarıma da bahsettim (Anonim).”

“İhtiyacımı karşılıyor, beni zaman aralıklarına sıkıştırmıyor, üstelik ücretsiz. Daha ne isteyebilirim ki... (SS-Erzurum)”

İyi okumalar ve başarılı bilimsel çalışmalar diliyoruz.

(Ağustos 2015)

Prof. Dr. Zekeriya AKTÜRK

Kitap Hakkında – Genel Amaç ve Hedefler

Zekeriya AKTÜRK

Ülkemizde tıp alanında araştırmaya olan ilgi, bilimsel çalışmalar ve yayınlar giderek artmaktadır. YÖK ve üniversitelerimizin akademik yükseltme ve atama kriterlerini belirlemiş olmaları ve bunların giderek zorlaşması da bilimsel araştırmalara olan ilgiyi artırmaktadır.

Diğer taraftan, aile hekimliği disiplininin hızla gelişmesi, birinci basamakta da şimdiye kadar hiç olmadığı kadar bilimsel çalışmalara ilgi duyulmasını sağlamıştır.

Şüphesiz bir bilimsel araştırmayı planlamak, verilerini toplamak, analizini yapmak ve yayına hazırlamak bir ekip işidir. Ancak, koşullarımız her zaman ekip desteği bulmamıza izin vermemektedir. Sonuçta, özellikle araştırmaların planlama ve istatistik analiz aşamalarında sorunlarla karşılaşmaktadır.

İstatistik ve araştırma konusunda çok değerli basılı eserler ve internet kaynakları vardır. Bununla birlikte, istatistik konusu özellikle klinisyenler için ürkütücü olmaya devam etmektedir. Hekimler karmaşık gözüken hesaplamalar yerine pratik olarak çalışmalarına yol gösterilmesini tercih etmektedir.

Araştırma görevlileri de asistanlık sürelerinin sonundaki tez çalışmalarında önemli zorluklarla karşılaşmakta, çalışmalarının istatistik analizleri için danışmanlığa ihtiyaç duymaktadır.

Klinisyen olarak özellikle aile hekimlerinin olmak üzere tüm sağlık çalışanlarının, araştırma görevlilerinin ve tıp öğrencilerinin kolaylıkla anlayıp uygulayabilecekleri bir esere ihtiyaç olduğunu düşünüyoruz.

Temel araştırma yöntemleri ve istatistik analizlerini içeren, örnek veri setleri ve SPSS’teki kullanımlarıyla desteklenmiş bu kitabın konusuyla ilgili ihtiyacı karşılayacağını düşünüyoruz.

Bu kitabın amacı, bilimsel araştırma açısından kendisini geliştirmek isteyen sağlık profesyonellerine destek olmaktır. Programın sonunda okuyucunun bir araştırmayı planlayabilmek, uygulayabilmek ve SPSS paket programını kullanarak temel istatistik analizleri yapabilmek için gerekli bilgi ve beceriyle donanmaları amaçlanmaktadır.

Kitaptaki konular örnek problemlerden hareket edilerek sistematik bir yaklaşımla işlenmiştir. Kitapta ayrıca birçok alıştırma uygulamasına yer verilmiştir. Alıştırmaların ilgili konudan sonra çözülmesi ve sayfa 464’teki çözümlere de bakılması yararlı olacaktır. Temel araştırma teknikleri ve istatistik analiz yöntemlerinin tamamının kapsanması hedeflenmiştir. Yaygın ihtiyaç

olması nedeniyle ayrıca lojistik regresyon ve örneklem hesabı gibi daha ileri konulara da değinilmiştir. Örneklerin çözümünde SPSS 15. Sürüm (bazı yerlerde 18. Sürüm veya 21. Sürüm) kullanılmıştır. Analizde kullanılacak veri setini internetten indirebilirsiniz: <https://drive.google.com/open?id=0B0ieO-Fy3z-5a3cydjlPQ0YtYk0>.

Haftada iki saat okuma ve iki saat de pratik yapmanız halinde içerisinde aşağıda belirtilen hedeflere ulaşabileceğinizi düşünüyoruz. Konular tamamlandıktan sonra Bölüm 55: Dergi Seçimi

Zekeriya AKTÜRK

Bu bölümde okuyucuların bilimsel bir eseri yayınlanmak üzere göndermek için dergi seçimi hakkında bilgi sahibi olmaları amaçlanmıştır.

Dergi önerisi	http://jane.biosemantics.org/suggestions.php
	https://journalsuggester.springer.com/
ISI	http://mjl.clarivate.com/
Index of Open Access Fees	http://eigenfactor.org/projects/openAccess/oa.php?catid=2&pt=0
Ücretsiz dergiler	http://www.journals4free.com/
PubMed	https://www.ncbi.nlm.nih.gov/nlmcatalog
Ulakbim	https://trdizin.gov.tr/statistics/listAcceptedJournals.xhtml
Sahte Yayıncılar	https://web.archive.org/web/20170111172306/https://scholarlyoa.com/publishers/
Sahte Dergiler	https://predatoryjournals.com/journals/

Bölüm 56: Kendini Değerlendirme Çalışması” Bölümündeki deneme sınavını yapmanızı ve “

Bölüm 57: Kendini Değerlendirme Çalışması Çözümleri” Bölümündeki puanlama kriterlerine göre puanınızı hesaplamanızı öneriyoruz. 70 puanın altında almanız halinde eksik olduğunuz konuları yeniden çalışmanızı öneririz.

Bu kitaptaki konuların okunması ve örnek ve alıştırmaların çözülmesi halinde aşağıdaki hedeflere ulaşmanız beklenir:

- Veri analizinde kullanılan temel teknikleri sayabilme
- Veri tiplerini sayabilme
- İstatistik yazılımlarına veri girme prensiplerini bilme
- Uç değerlerin önemini açıklayabilme ve yönetebilme
- Grafik çeşitlerini sayabilme ve açıklayabilme
- Bilimsel tablo hazırlama prensiplerini açıklayabilme
- Merkezi eğilim ölçütlerini açıklayabilme
- Yaygınlık ölçütlerini açıklayabilme
- Normal dağılımı açıklayabilme
- Diğer sürekli dağılımları açıklayabilme
- Poisson dağılımını açıklayabilme
- Ki kare dağılımını açıklayabilme
- Binomial dağılımı açıklayabilme
- Veri dönüşümünün gerektiği durumları açıklayabilme ve yapabilme
- Örnekleme yöntemlerini sayabilme
- Güven aralığını hesaplayabilme
- Araştırma çeşitlerini sayabilme ve sınıflandırabilme
- Randomize kontrollü çalışmaları açıklayabilme
- Kohort çalışmaları açıklayabilme
- Vaka kontrol çalışmalarını açıklayabilme
- Tip 1 hata, tip 2 hata araştırma gücü ifadelerini ve aralarındaki ilişkiyi açıklayabilme
- Tek örnekleme t testini açıklayabilme ve kullanım yerini bilme
- Bağımlı örneklemlerde t testini açıklayabilme ve kullanım yerini bilme
- Wilcoxon signed rank testini açıklayabilme ve kullanım yerini bilme
- Sign testini açıklayabilme ve kullanım yerini bilme

- Bağımsız örneklemelerde t testini açıklayabilme ve kullanım yerini bilme
- Wilcoxon rank sum testini açıklayabilme ve kullanım yerini bilme
- Mann whitney U testini açıklayabilme ve kullanım yerini bilme
- ANOVA testini açıklayabilme ve kullanım yerini bilme
- Kruskal Wallis testini açıklayabilme ve kullanım yerini bilme
- Nümerik veri, kategorik bağımsız değişkenden tanımlayabilme
- İki yönlü ANOVA açıklayabilme ve kullanım yerini bilme
- Tek örnekte z testini açıklayabilme ve kullanım yerini bilme
- McNemar testini açıklayabilme ve kullanım yerini bilme
- Ki kare testini açıklayabilme ve kullanım yerini bilme
- Ki kare trend testini açıklayabilme ve kullanım yerini bilme
- Fischer exact testini açıklayabilme ve kullanım yerini bilme
- Friedman testini açıklayabilme ve kullanım yerini bilme
- Cochran's Q testini açıklayabilme ve kullanım yerini bilme
- Marginal Homogeneity testini açıklayabilme ve kullanım yerini bilme
- Hipotez testlerini SPSS ile yapabilme ve çıktıları yorumlayabilme
- Çeşitli senaryolarda hangi istatistik analizler yapılacağını açıklayabilme
- Korelasyon analizini açıklayabilme, SPSS'te yapabilme ve sonuçları yorumlayabilme
- Lojistik regresyon analizini açıklayabilme, SPSS'te yapabilme ve sonuçları yorumlayabilme
- Örnekleme etki eden faktörleri açıklayabilme ve çeşitli yazılımları kullanarak örneklem hesabı yapabilme
- Duyarlılık ve özgüllük kavramlarını açıklayabilme
- ROC analizinin kullanım yerini açıklayabilme
- Güvenilirlik çeşitlerini sayabilme ve SPSS ile yapabilme
- Geçerlilik çeşitlerini açıklayabilme
- Kaplan Meier sağkalım analizinin kullanım yerini açıklayabilme
- Cox regresyon analizinin kullanım yerini açıklayabilme
- Kappa uyum analizinin kullanım yerini açıklayabilme
- Bland-Altman analizinin kullanım yerini açıklayabilme

- Yukarıda geçen testleri SPSS’te yapabilme ve çıktıları yorumlayabilme
- Bilimsel eserlerde yazar hakkı prensiplerini açıklayabilme
- Bilimsel çalışmalar için ulusal ve uluslararası fonlara örnek verebilme
- Bilimsel rapor yazma prensiplerini açıklayabilme
- Tanı testlerinin sağlık hizmetlerindeki yerini tartışabilme

Bölüm 1: Giriş

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucuların sağlıkta araştırma yapmanın önemini kavramaları ve araştırma yöntemleri konusunda temel bilgiler kazanmaları amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Bir araştırmanın aşamalarını sayabilmeleri
- Cevaplanabilir bir klinik soru oluşturabilmeleri
- Veri çeşitlerini sayabilmeleri
- Bir veride aranması gereken özellikleri sayabilmeleri
- Araştırma çeşitlerini sayabilmeleri
- Kohort, vaka kontrol ve randomize kontrollü çalışmayı açıklayabilmeleri
- Parametrik ve nonparametrik testlerin uygulanma yerlerini açıklayabilmeleri

Bir araştırma yapmanız için neler gerekiyor?

- İstek (Gerek var mı?)
- Zaman
- Fikir (Neyi araştırayım?)
- Beceri (Nasıl?)
- Destek (Proje, istatistik...)
- Para (?)

Genelde araştırma yapmama nedeni olarak maddi kaynak olmamasını öne süreriz. Oysa birçok araştırma kendi imkanlarımızla, ek bir para olmadan yapılabilir. Kaldı ki, üniversitelerin araştırma fonları, TÜBİTAK fonları, DPT ve Bakanlık fonları gibi birçok kaynağa da uygun projeler sunmamız halinde ulaşabiliriz (Bakınız sayfa 425). Öyleyse ne duruyoruz?

Araştırma yapmak için uygun bir sistematik izlemekte fayda vardır. Aksi halde istediğimiz başarıyı elde edemeyebilir ve kaynak ve zaman israfına yol açabiliriz. Araştırmanın belki en önemli aşaması konuyu belirleme ve nasıl araştıracağımızın cevabını vermektedir.

Bir araştırmada şu aşamalar olmalıdır:

1. Yanıtlanabilir bir soru sormak
2. Sorunuz daha önce cevaplanmış mı?

3. Hipotez kurma
4. Yöntem belirleme
5. Uygulama (araştırmayı yürütme)
6. Verileri bilgisayara girme ve istatistik analizleri yapma.
7. Bulguları yorumlama ve literatürle karşılaştırma
8. Rapor yazma ve yayınlama

Araştırma Fikri

Görüldüğü üzere belirtilen aşamalar içerisinde en değerlisi araştırma fikrini ortaya koyma ve yöntemi belirlemedir. Fikir yakalamak için uyanık olmalıyız; her an araştırabileceğimiz potansiyel konuları belirlemek üzere dikkatimiz açık olmalı. Fikirlerimizi aşağıdaki yollarla elde edebiliriz:

- Klinik uygulamalarımız ve karşılaştığımız durumlar
- Çalışma ortamımız
- Hastalardan gelen fikirler
- Arkadaşlarımızın önerileri ve görüşleri
- Beklenmeyen laboratuvar sonuçları
- ...

Belirlediğimiz fikirleri hemen araştırma imkanımız olmayacağından bunları not etmemiz ve gerektiğinde kullanmamız iyi olacaktır. Bu amaçla bir araştırma panosu (

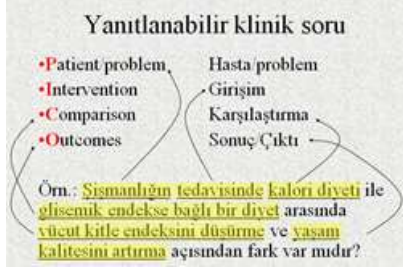
Tablo 1) oluşturabilirsiniz. Panonuz çalışma masanızın önünde olabileceği gibi bilgisayarınızda bir dosyadan da oluşabilir. Araştırma panosunda “araştırma fikirleri”, “yöntem yazılma aşamasında olan konular”, “uygulama aşamasında olan konular” ve “rapor yazım aşamasında olan konular” gibi başlıklarınız olmalıdır. Araştırma panosu sürekli dinamik olmalıdır ve konular sütunlar arasında yer değiştirerek yayına doğru gitmelidir.

Araştırma Panom			
Sıradaki çalışmalar	Yöntem yazım aşamasında	Uygulama aşamasında	Rapor yazım aşamasında
1. Halk arasında bitkisel ilaçların kullanma durumu 2. Kabızlığın zihinsel fonksiyonlara etkisi 3. Danışmanlığın ergenlerdeki riskli davranışlara etkisi. 4. .	1. Türkiye’de tıpta uzmanlık eğitimi 2. .	1. Yapılandırılmış egzersizin kilo vermeye etkisi 2. Birinci basamak için bir ulusal kayıt yazılımı geliştirilmesi 3. .	1. Edirneli kadınlarda osteoporoz prevalansı 2. Tıp öğrencilerinin bitkileri tanıma durumu 3. .

Tablo 1: Örnek bir araştırma panosu.

PICO

Araştırmaya başlarken cevaplanabilir bir klinik soru sorulmalıdır. Sorumuzda dört bileşen olmalıdır: **P**roblem/Hasta, **G**irişim (**I**ntervention), **K**arşılaştırma (**C**omparison) ve **Ç**ıktı (**O**utcome). Hatırlaması kolay olsun diye bu kelimelerin baş harfleri olan PICO akronimi akılda tutulabilir (**Şekil 1**).



Şekil 1: Yanıtlanabilir klinik soru

PICO akronimi S (setting, ortam) eklenerek PICOS olarak da kullanılmaktadır. Bu durumda araştırma sorusunda araştırmanın ortamı hakkında da bilgi verilmesi istenir.

Literatür tarama

Cevaplanabilir bir klinik soru sorduktan sonra sorumuzun daha önce cevaplanıp cevaplanmadığına bakmamız gerekir. Bunun için:

- Literatür taraması
- Tecrübeli meslektaşlarımızın fikirlerine başvurma
- Eğiticilere danışma
- Arkadaşlarımıza sorma
- Kitaplara bakma
- İnternette arama
- ...

gibi yöntemleri kullanabiliriz.

Hipotez kurma

Daha sonraki aşamada araştırma hipotezini kurmamız gerekir. Hipotez geleneksel olarak H_0 ve H_1 olarak ifade edilir. H_0 hep olumsuz olarak kurulur ve “incelenen durumlar arasında fark yoktur” gibi ifade edilir:

- H_0
 - Şişmanlığın tedavisinde kalori diyeti ile glisemik endekse bağlı bir diyet arasında vücut kitle endeksini düşürme ve yaşam kalitesini artırma açısından fark yoktur.
- H_1
 - Şişmanlığın tedavisinde kalori diyeti ile glisemik endekse bağlı bir diyet arasında vücut kitle endeksini düşürme ve yaşam kalitesini artırma açısından fark vardır.

Bazı prensipler

Bilimsel araştırma ve istatistik uygulamaları sırasında aşağıdaki noktalara dikkat edilmesini öneriyoruz:

- Birden çok hipotezi bir arada test edebiliriz, ama olabildiğince az hipotez kurmalıyız.
- Hastalarımızın tüm sorunlarına eğilip hiç birini çözmemek yerine bir soruna eğilip onu çözmek daha yararlı olur.
- Araştırmamızın yöntemi ve örnekleme her hipoteze göre ayrı ayrı değerlendirilmelidir.
- Gerekliyse mutlaka erken dönemde bir istatistik uzmanına danışılmalıdır.

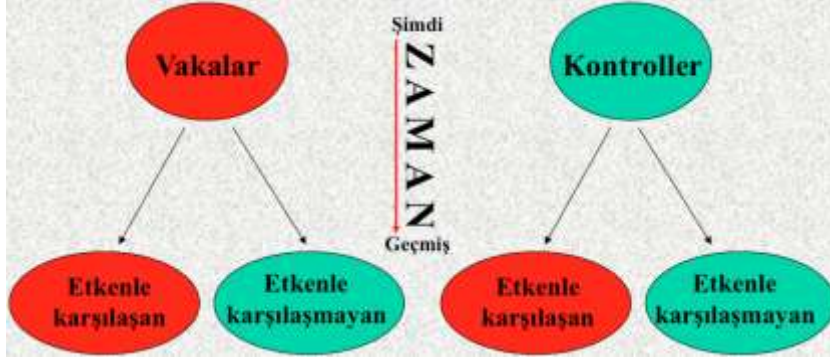
Araştırma yöntemleri

Araştırma sorumuzu ortaya koyup literatür taraması yaptıktan sonra sorumuzun cevabını hangi araştırma tasarımıyla ortaya koyabileceğimizi belirlemeliyiz.

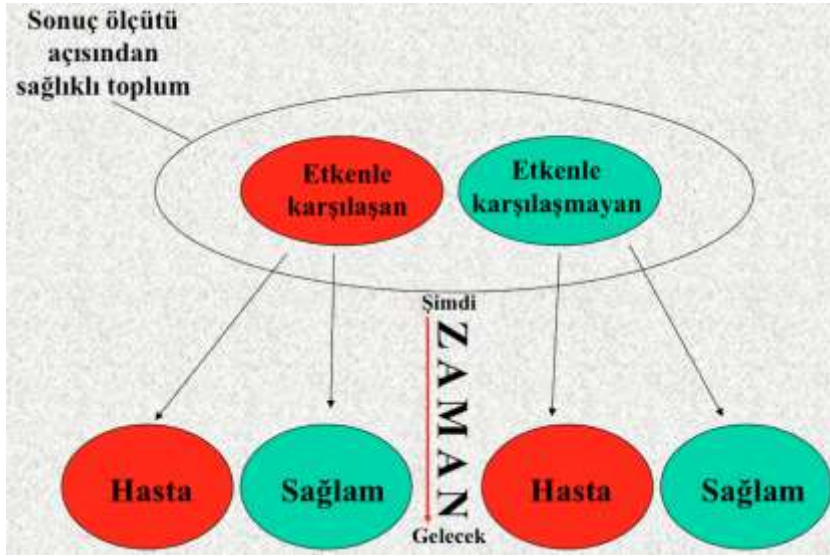
Araştırma çeşitlerini aşağıdaki gibi sınıflandırabiliriz:

- Gözlemsel araştırmalar
 - Vaka sunumu (case report)
 - Prevalans (yaygınlık)
 - Vaka kontrol (case control) (Şekil 2)
 - Kalitatif araştırmalar
 - Kohort çalışmaları (Şekil 3)
- Deneysel araştırmalar
 - Randomize kontrollü çalışmalar (Şekil 4)
 - Saha deneyleri
- Metodolojik

- Geçerlilik, güvenilirlik çalışması
- Özgüllük, duyarlılık çalışmaları



Şekil 2: Vaka Kontrol Çalışması

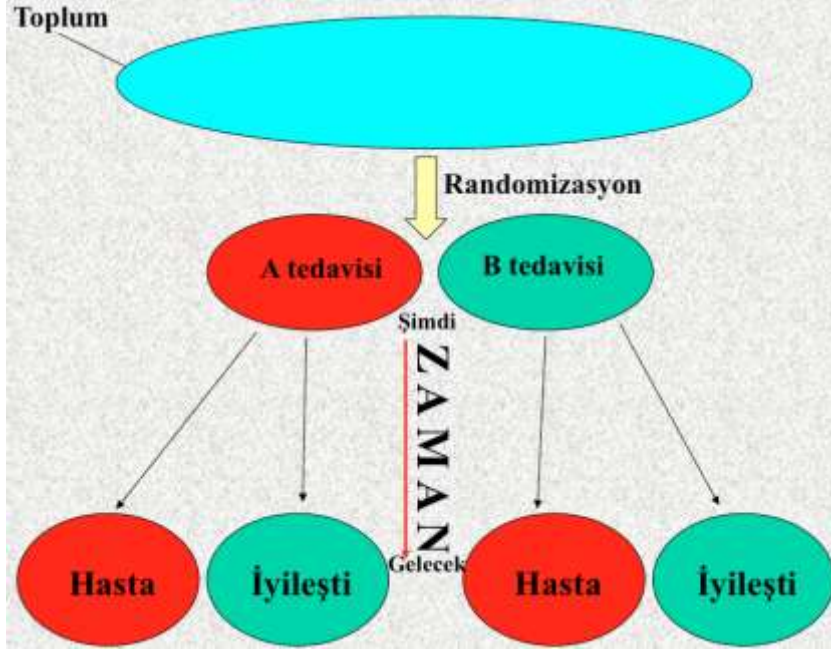


Şekil 3: Kohort çalışması

Vaka kontrol çalışmalarında araştırmaya incelenen hastalığa yakalanmış (sonuç ölçütü ortaya çıkmış) kişilerle (vakalarla) başlanır. Bu hastaların ve hasta olmayan kontrollerin riskle karşılaşma durumları geriye doğru incelenir.

Kohort araştırmasında ise sonuç ölçütü açısından sağlıklı kişiler araştırmaya dahil edilir ve gelecekte riskle karşılaşp karşılaşmama durumlarına göre hastalığın (sonuç ölçütünün) ortaya çıkıp çıkmayacağı incelenir.

Deneysel araştırmalarda, belirlenen dahil edilme ve hariç tutma ölçütleri çerçevesinde araştırmaya alınan kişiler rastgele (random) gruplara ayrılır. Katılımcılarda araştırılmak istenen sonuç ölçütü araştırmanın başında ve belli bir müdahale yapıldıktan sonra ölçülerek karşılaştırma yapılır.



Şekil 4: Randomize Kontrollü Çalışma

İstatistik Analiz

Araştırma tasarımıımızı belirleyip verilerimizi topladıktan sonra sıra verilerin bilgisayara girilmesi ve analiz edilmesine gelecektir. Bu noktada verilerde bulunması gereken özellikler, veri çeşitleri ve istatistik önemlilik testleri (hipotez testleri) hakkında bilgi sahibi olmalıyız.

Verilerde bulunması gereken özellikler:

- Güncel
- Güvenilir
- Doğru
- Eksiksiz
- Kullanıma uygun
- Amaca uygun

Veri çeşitleri:

- Nümerik
- Kategorik

Bazı önemlilik testleri:

- Parametrik (verilerin belli bir olasılık dağılımından geldiğini varsayar ve dağılım parametreleri hakkında çıkarım yapmaya çalışır¹)
 - Student t-test
 - ANOVA
- Nonparametrik (parametrik olasılık dağılımlarına dayanmazlar ve olasılık dağılımı hakkında varsayımda bulunmazlar²)
 - Ki kare (Chi Square)
 - Mann-Whitney U testi
 - McNemar testi
 - Wilcoxon
 - Fisher exact test

¹ https://en.wikipedia.org/wiki/Parametric_statistics

² https://en.wikipedia.org/wiki/Nonparametric_statistics

Parametrik testlerin bazı varsayımları vardır:

- Evrenin normal dağılıma uyması
- Varyansların homojen olması
- Deneklerin evrenden rasgele seçilmiş olması
- Deneklerin birbirinden bağımsız olması
- Denek sayılarının yeterli olması

Bir örnek yapacak olursak:

Cevaplanabilir soru oluşturmak

- Örneğin: Üniversite eğitimine başlayan gençlerde aile planlaması yöntemi bilenlerle bilmeyenler arasında cinsel ilişkide bulunma açısından bir fark var mıdır?³

Soru daha önce cevaplanmış mı?

- Araştırma sorusu oluşturduktan sonra mutlaka literatüre bakmalıyız. Literatüre baktığımızda konuyla ilgili birçok araştırma bulabiliriz. Bunlardan birisi: Fantahun M, Chala F, Loha M. Knowledge, attitude and practice of family planning among senior high school students in north Gonder. Ethiop Med J. 1995 Jan;33(1):21-9.
 - Bu makalede şöyle denmektedir: “The most common reason for not using modern contraceptive methods among sexually active respondents was little or no knowledge of contraceptives” (Cinsel aktivitede bulunanların modern kontraseptif kullanmamalarının en önemli sebebi, kontraseptif bilgilerinin olmaması veya zayıf olmasıydı).

Hipotez oluşturma

- H_0 : Aile planlaması yöntemlerini bilen ve bilmeyen gençler arasında cinsel ilişkide bulunma açısından fark yoktur.

³ Doç. Dr. Turan Set. Üniversite çağındaki ergenlerin cinsel bilgi tutum ve davranışları. Uzmanlık Tezi, Edirne 2005.

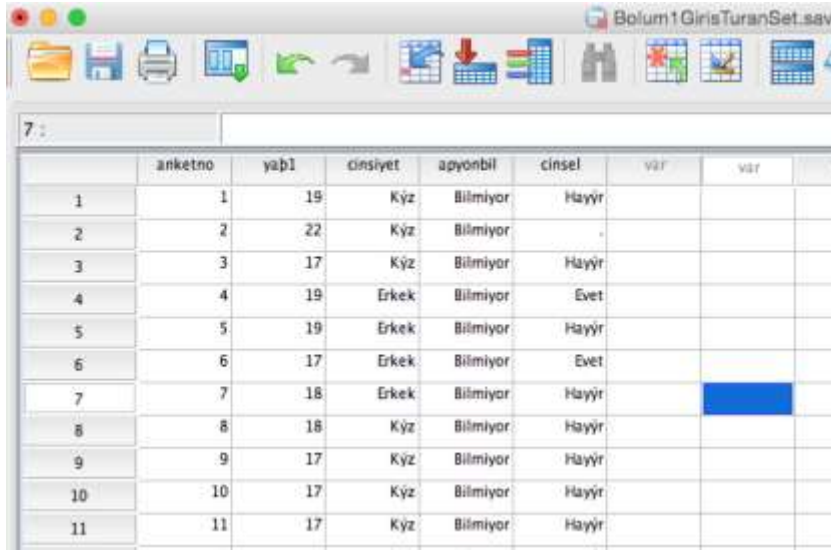
- H_1 : Aile planlaması yöntemlerini bilen ve bilmeyen gençler arasında cinsel ilişkide bulunma açısından fark vardır.

Yöntem

- Bu araştırma için **kesitsel tanımlayıcı** bir çalışma tasarımı uygundur.
- Üniversiteye kayıt için gelen gençlere bir anket doldurtulması
 - Bağımlı değişken (= ana sonuç ölçütü):
 - Cinsel ilişkide bulunmuş olma durumu
 - Bağımsız değişkenler
 - Yaş
 - Cinsiyet
 - Aile planlaması yöntemi bilme durumu

SPSS veri seti

Veriler bilgisayara girildi ve SPSS ile analiz etmeye hazır hale getirildi:



	anketno	yaş	cinsiyet	aileyonbil	cinsel	v1	v2	v3
1	1	19	Kız	Bilmiyor	Hayır			
2	2	22	Kız	Bilmiyor				
3	3	17	Kız	Bilmiyor	Hayır			
4	4	19	Erkek	Bilmiyor	Evet			
5	5	19	Erkek	Bilmiyor	Hayır			
6	6	17	Erkek	Bilmiyor	Evet			
7	7	18	Erkek	Bilmiyor	Hayır			
8	8	18	Kız	Bilmiyor	Hayır			
9	9	17	Kız	Bilmiyor	Hayır			
10	10	17	Kız	Bilmiyor	Hayır			
11	11	17	Kız	Bilmiyor	Hayır			

(Bolum1GirisTuranSet.sav veri setini şuradan indirebilirsiniz:

<https://drive.google.com/file/d/0B0ieO-Fy3z-5NktDS3FBblk0UWc/view?usp=>).

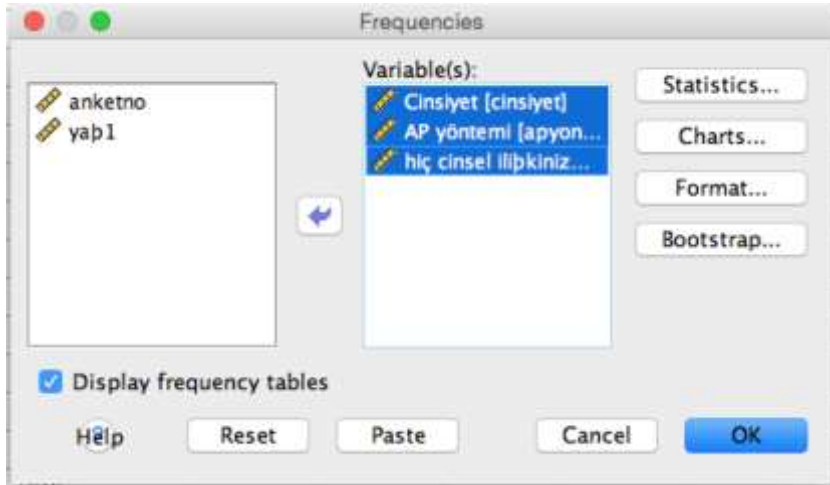
Analiz edeceğimiz verileri sınıflandıracak olursak:

- Bağımsız değişkenler
 - Yaş
 - Nümerik, Sürekli
 - Cinsiyet
 - Kategorik, Nominal⁴
 - AP yöntemi bilme durumu
 - Kategorik, Nominal
- Bağımlı değişkenler
 - Cinsel ilişkide bulunmuş olma
 - Kategorik, Nominal

Yapacağımız istatistik analizleri tanımlayıcı istatistikler ve hipotez testleri olmak üzere iki gruba ayırabiliriz.

Tanımlayıcı istatistikler

[Bolum1GirisTuranSet.sav](#) verisetini SPSS’de açınız. Analyze>Descriptive Statistics>Frequencies [cinsiyet, apyonbil ve cinsel değişkenlerini Variable(s) kısmına geçirelim]>OK:



⁴ İng. “Nomination” = Atanmış

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
YAP1	16051	16	46	18,55	1,57
Valid N (listwise)	16051				

Cinsiyet

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Erkek	8859	51,4	52,9	52,9
	Kız	7899	45,8	47,1	100,0
	Total	16758	97,2	100,0	
Missing	System	479	2,8		
Total		17237	100,0		

AP yöntemi

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Bilmiyor	12791	74,2	82,9	82,9
	Biliyor	2633	15,3	17,1	100,0
	Total	15424	89,5	100,0	
Missing	System	1813	10,5		
Total		17237	100,0		

hiç cinsel ilipkiniz oldu mu?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Evet	4497	26,1	27,8	27,8
	Hayır	11706	67,9	72,2	100,0
	Total	16203	94,0	100,0	
Missing	System	1034	6,0		
Total		17237	100,0		

Önemlilik (Hipotez) testi

Bu değişkenlerin karşılaştırılması için Ki Kare analizi uygundur (Sayfa 293).

Analyze>Descriptive Statistics>Crosstabs [apyonbil değişkenini Row(s), cinsel değişkenini Column(s) kısmına geçirelim]>OK:

Crosstabs

Row(s):
AP yöntemi [apyonbil]

Column(s):
hiç cinsel ilipkiniz oldu mu? [cinsel]

Layer 1 of 1

Previous Next

☐ Display clustered bar charts
☐ Suppress tables

Help Reset Paste Cancel OK

			hiç cinsel ilipkiniz oldu mu?		Total
			Evet	Hayır	
AP yöntemi	Bilmiyor	Count	3169	8762	11931
		% within AP yöntemi	26.6%	73.4%	100.0%
	Biliyor	Count	900	1653	2553
		% within AP yöntemi	35.3%	64.7%	100.0%
Total	Count	4069	10415	14484	
	% within AP yöntemi	28.1%	71.9%	100.0%	

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	78,644 ^a	1	,000	,000	,000
Continuity Correction ^b	78,215	1	,000		
Likelihood Ratio	75,920	1	,000		
Fisher's Exact Test					
Linear-by-Linear Association	78,639	1	,000		
N of Valid Cases	14484				

a. Computed only for a 2x2 table

b. 0 cells (,0%) have expected count less than 5. The minimum expected count is 717.22.

Yorum ve raporlama

- Aile planlaması bilenlerin %35,3'ü cinsel ilişkide bulunmuş iken bilmeyenlerin ise 26,6'sı cinsel ilişkide bulunmuştur. Cinsel ilişki ile aile planlaması yöntemi yöntemi bilme arasında iki yönlü anlamlı bir ilişki vardır (Ki kare=78,6; $p<0,001$)
- Aile planlaması yöntemi bilme durumu her iki grupta da düşüktür.
- Gençlerin aile planlaması yöntemi bilme durumlarına göre gelecekteki davranışlarını araştıran kohort çalışmaları daha yararlı bilgi sağlayacaktır.

Bölüm 2: İstatistiksel Testler ve Kullanım Yerleri – Akış Şemaları

Zekeriya AKTÜRK

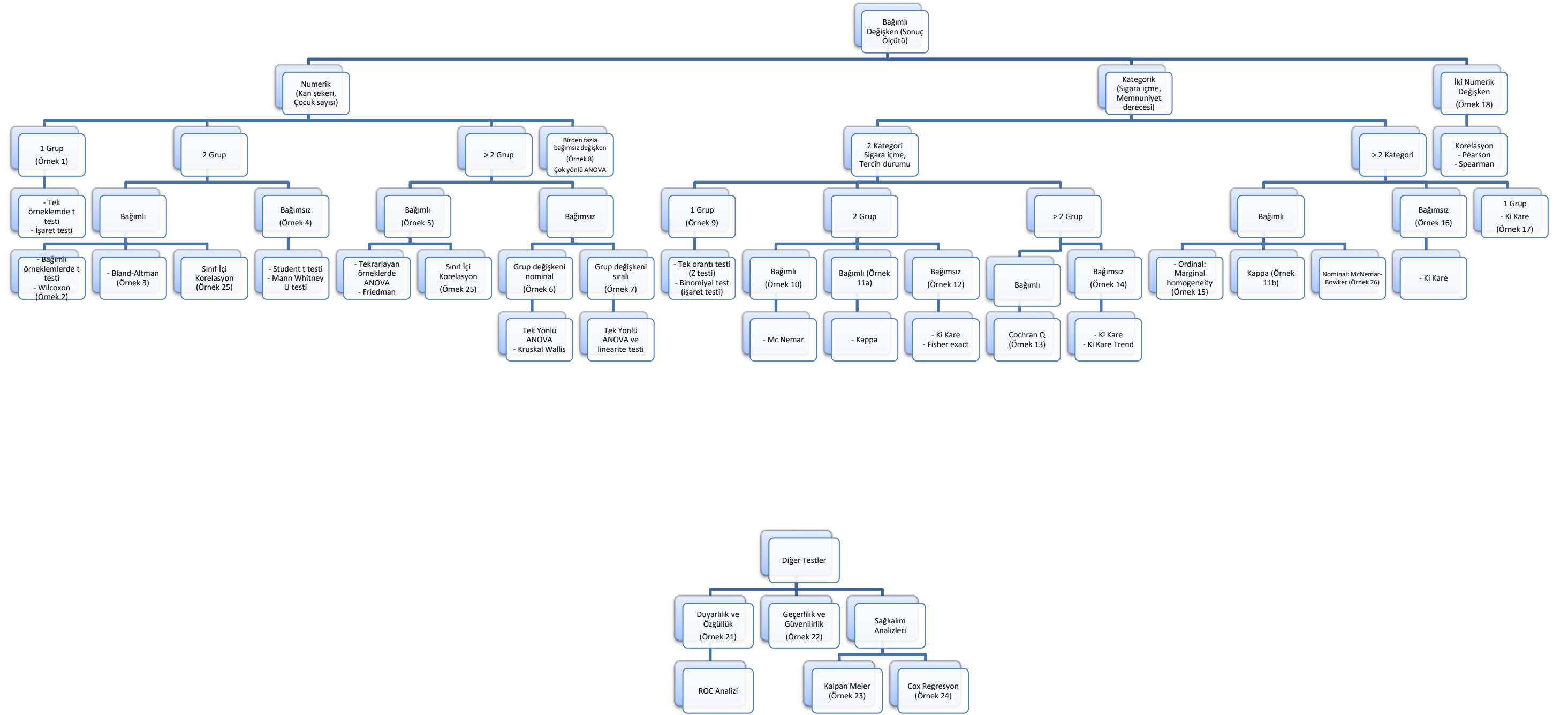
Amaç: Bu konu sonunda okuyucuların sağlık çalışmalarında kullanılan temel istatistik önemlilik testleri ve kullanım yerleri konusunda bilgi sahibi olmaları amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Temel hipotez testlerini ve kullanım yerlerini %80 doğrulukla belirtebilmeleri
 - Tek örnekleme t testi
 - İşaret testi
 - Bağımlı örneklemlerde t testi
 - Bağımsız örneklemlerde t testi
 - Wilcoxon testi
 - Mann Whitney U testi
 - Tek yönlü ANOVA
 - İki yönlü ANOVA
 - Tekrarlayan ölçümlerde ANOVA
 - Friedman analizi
 - Z testi
 - Ki kare testi
 - McNemar testi
 - Ki kare trend testi
 - Fisher exact testi
 - Kappa uyum analizi
 - Cochran Q testi
 - Pearson korelasyon analizi
 - Spearman korelasyon analizi
 - Lojistik regresyon analizi
 - Lineer regresyon analizi
 - Kaplan Meier sağkalım analizi
 - Cox regresyon analizi
 - ROC analizi
 - Geçerlilik ve güvenilirlik analizi
- Parametrik ve nonparametrik test ayrımını açıklayabilmeleri

Klinisyen için farklı bir dünya ifade eden istatistik analizlere girmeden önce bu dünyanın boyutlarıyla ilgili bir fikir verilmesi yararlı olacaktır. İstatistiksel analizlerde en çok tanımlayıcı istatistikleri (ortalamaların ve frekansların

verilmesi, grafiklerle gösterme vs.) ve hipotez testlerini (t-testi, Ki kare vb.) kullanıyoruz. Hangi hipotez testinin nerede kullanılacağını bilmek temel araştırmalar yapacak bir araştırmacı için çok önemlidir. Bu amaçla hazırlanan akış şemaları iyice incelenmeli ve pratik kazanılıncaya kadar araştırmacının başucunda asılı bulunmalıdır.



Örnek Problemler

Örnek 1

(Sayfa 196)

Nümerik Değişken Tek Grup

Parametrik: Tek örnekleme t testi; Non parametrik: İşaret testi

Mahallemizdeki bireylerin boylarını ölçtüğümüzü düşünelim [mahalle=tek grup; boy=nümerik veri]. Sonuçlarımızı il geneli ile karşılaştırmak istiyoruz. Daha önceki araştırmalardan ilimizdeki bireylerin boy ortalamasının 161,5 cm olduğunu biliyoruz.

Örnek 2

(Sayfa 207)

Nümerik Değişken İki Bağımlı Grup

Parametrik: Bağımlı örneklemelerde t testi; Non parametrik: Wilcoxon testi

Diyabetik hastalarımıza bir beslenme ve egzersiz programı uyguladığımızı varsayalım. Program sonunda katılımcıların kilo verip vermediklerini araştırmak istiyoruz.

Örnek 3

(Sayfa 324)

Nümerik Değişken İki Bağımlı Grup

Bland-Altman analizi

Hastalarımızın glomeruler filtrasyon hızlarını (GFR) [nümerik değişken] iki farklı yöntemle (A ve B yöntemi) ölçtük. Bu yöntemlerin birbirleri ile ne kadar uyumlu olduklarını araştırmak istiyoruz. Klasik olarak kliniğimizde yıllardır A yöntemi kullanılmaktadır. A yöntemi yerine yeni geliştirilen ve daha hızlı ve ekonomik olan B yöntemini kullanarak kullanamayacağımızı araştırmak istiyoruz. H_0 : GFR hesaplamada A yöntemi ile B yöntemi arasında fark yoktur.

Örnek 4

(Sayfa 220)

Nümerik Değişken İki Bağımsız Grup

Parametrik: Bağımsız örneklemelerde t testi (Student t testi); Non parametrik: Mann Whitney U testi

Diyabet hastalarımızın vücut kitle indekslerinin (VKİ) erkekler ve bayanlar arasında farklılık gösterip göstermediğini araştırmak istediğimizi varsayalım

Örnek 5

(Sayfa 231)

Nümerik Değişken İki Fazla Bağımlı Grup

Parametrik: Tekrarlayan ölçümlerde (repeated measures) ANOVA; Non parametrik: Friedman analizi

Senaryo 1, Tek faktör (grup içi karşılaştırma)

Diyabetik hastalarımıza ilaç, beslenme ve egzersizden oluşan tedavi programı uyguladığımızı varsayalım. Program süresince katılımcıların Hemoglobin A1c düzeylerini araştırmak istiyoruz (H_0 : Bireylerin uygulanan tedavi programı öncesi ve sonrasındaki Hemoglobin A1c düzeyleri arasında fark yoktur). Hemoglobin A1c ölçümleri tedavi programı uygulamadan önce, programın 3. ve 6. aylarında yapılmış olsun. Ölçmek istediğimiz değişken nümeriktir. Uygulama öncesi ve sonrasında elde edilen üç ölçümler ise bağımlıdır.

Senaryo 2, İki faktör (grup içi ve gruplar arası karşılaştırma)

Erkek ve kadın diyabetik hastalarımıza ilaç, beslenme ve egzersizden oluşan tedavi programı uyguladığımızı varsayalım. Program süresince cinsiyete ve zamana göre Hemoglobin A1c düzeylerini araştırmak istiyoruz (H_0 : Kadın ve erkek grupları arasında ve tedavi

programı öncesi ve sonrasındaki Hemoglobin A1c düzeyleri arasında fark yoktur ve cinsiyet*zaman etkileşimi (interaction) yoktur).

Örnek 6

(Sayfa 248)

Nümerik Değişken İkiden Fazla Bağımsız Grup

Parametrik: Tek yönlü ANOVA; Non parametrik: Kruskal Wallis analizi

Diyabet hastalarımızın geldikleri mahallelere göre boylarının farklılık gösterip göstermediğini araştırmak istediğimizi varsayalım (H_0 : Farklı mahallelerden gelen diyabet hastalarının boyları arasında fark yoktur). Ölçmek istediğimiz değişken (boy) nümerik sürekli bir değişkendir. Güney, kuzey, doğu, batı ve merkez olmak üzere beş grubumuz var.

Örnek 7

(Sayfa 248)

Nümerik Değişken İkiden Fazla Bağımsız Grup

Parametrik: Tek yönlü ANOVA ve lineerite testi

Hastalarımızın eğitim durumlarına göre yaşlarının farklılık gösterip göstermediğini araştırmak istiyoruz. Ayrıca arada bir ilişki varsa yaşın eğitim durumu arttıkça lineer bir şekilde değişip değişmediğini belirlemek istiyoruz.

(H_0 : Eğitim durumları farklı olan hastalarının yaşları arasında lineer bir ilişki yoktur). Ölçmek istediğimiz değişken (yaş) nümerik bir değişkendir. Eğitim durumu kategorik sıralı bir değişkendir.

Örnek 8

(Sayfa 262)

Nümerik Değişken İki veya Daha Fazla Kategorik Bağımsız Değişken

Parametrik: İki yönlü ANOVA

Diyabet hastalarımızın cinsiyet ve sağlık hizmetlerinden memnuniyet durumlarına göre açlık kan şekerlerinin farklılık gösterip göstermediğini araştırmak istiyoruz (H_0 : Cinsiyetleri ve memnuniyet durumları farklı olan diyabet hastalarının açlık kan şekerleri arasında fark yoktur). Ölçmek istediğimiz değişken (açlık kan şekerleri) nümerik sürekli bir değişkendir. Birinci kategorik bağımsız değişkenimiz cinsiyet, erkek ve kadınlardan oluşan iki bağımsız grubu bulunurken, ikinci kategorik bağımsız değişkenimiz olan memnuniyet durumu; Hiç memnun değil, az memnun, memnun ve çok memnun olmak üzere dört bağımsız kategoriden oluşmaktadır.

Örnek 9

(Sayfa 270)

Kategorik Değişken Tek Grup

Z testi

Diyabet hastalarında sigara içme durumunu araştırdığımızı düşünelim. Örneklemimizdeki sigara içme durumunun toplumdakinden farklı olup olmadığına bakmak istiyoruz

(Kategorik değişken tek gruba diğer bir örnek için Örnek 17'ye bakınız).

Örnek 10

(Sayfa 276)

Kategorik Değişken İki Bağımlı Grup

McNemar testi

Diyabetik hastalarımıza kan şekeri kontrolünün önemi ve gerektiğinde insülin kullanılması konusunda bir eğitim verdiğimiz düşünelim. Örneklemimizdeki bireylerin ilaç tercihlerinin eğitimden sonra değişip değişmediğini incelemek istiyoruz.

Örnek 11

(Sayfa 280)

Kategorik Değişken İki Bağımlı Grup

Kappa analizi

11a: Hastalarımızın EKG'lerini iki farklı hekim inceliyor (hekimler aynı kişinin EKG'sini incelediğinden bağımlı) ve patoloji olup olmadığını rapor ediyorlar (Var/Yok). Hekimler arasındaki uyumu araştırmak istiyoruz.

11b: Hastalarımızın EKG'lerini iki farklı hekim inceliyor (hekimler aynı kişinin EKG'sini incelediğinden bağımlı) ve patoloji olup olmadığını rapor ediyorlar (Var/Şüpheli/Yok). Hekimler arasındaki uyumu araştırmak istiyoruz.

Örnek 12

(Sayfa 293)

Kategorik Değişken İki Bağımsız Grup

Ki Kare testi, Fisher exact testi

Kliniğimizde takip edilen diyabet hastalarının ilaç tercihlerinde cinsiyet açısından bir farklılık olup olmadığını araştırmak istiyoruz. İncelediğimiz değişken (ilaç tercihi; 1-insülin, 2-oral antidiyabetik) kategorik bir değişkendir. Erkekler ve bayanlar olmak üzere iki grubumuz söz konusu.

Örnek 13

(Sayfa 311)

Kategorik Değişken İki Fazla Bağımlı Grup

Cochran Q testi

Diyabetik hastalarımıza kan şekeri kontrolünün önemi ve gerektiğinde insülin kullanılması konusunda bir eğitim verdik. Örneklemimizdeki bireylerin ilaç tercihlerinin eğitimlerden sonra değişip değişmediğini incelemek istiyoruz. İncelediğimiz değişken (ilaç tercihi; 1-

insülin, 2-oral antidiyabetik) kategorik dikotom bir değişkendir. Danışmanlık öncesi, danışmanlık sonrası 1. ayda ve danışmanlık sonrası 6. aylarda aynı bireylerden veri topladık; üç bağımlı grup söz konusudur.

Örnek 14

(Sayfa 299)

Kategorik Değişken İki Fazla Bağımsız Grup

Ki Kare testi, Ki Kare Trend testi

1: Farklı mahallelerden gelen diyabet hastalarımızın medeni durumları arasında fark olup olmadığını incelemek istiyoruz.

2: Diyabetik bireylerde eğitim durumu ile sigara içme durumu arasında bir trend olup olmadığını incelemek istiyoruz. Eğitim arttıkça içme yüzdesinin azaldığını varsayıyoruz.

Örnek 15

(Sayfa 315)

Kategorik Değişken İki Fazla Bağımlı Kategori

Marginal Homogeneity testi

Diyabetik hastalarımıza hizmet veren sağlık kuruluşu ile ilgili bir kalite iyileştirme programının uygulandığını varsayalım. Program uygulanmadan önce ve uygulandıktan sonra hasta memnuniyeti "hiç memnun değilim, az memnunum, memnunum ve çok memnunum" şeklinde sorulmuştur. Dolayısıyla ordinal kategorik veri elde edilecektir.

Programdan sonra hasta memnuniyetinde anlamlı bir değişim olup olmadığı araştırılmak isteniyor.

Örnek 16

(Sayfa 319)

Kategorik Değişken İki Fazla Bağımsız Kategori

Ki Kare testi

Örneklelimizdeki bireylerin medeni durumlarının (evli, bekar, boşanmış = ikiden fazla kategori) antidiyabetik ilaç tercihleri (oral antidiyabetik, insülin = iki grup) ile ilişkili olup olmadığını incelemek istiyoruz. Bağımsız değişkeni-miz ikiden fazla kategoriden oluştuğundan bu örnekte de örnek 12 ve 14'te olduğu gibi Ki Kare testi uygulanmalıdır.

Örnek 17

(Sayfa 321)

Kategorik Değişken İkiden Fazla Kategori, 1 Grup

Ki Kare testi

Örneklelimizdeki bireylerin kullandıkları tedaviler arasında anlamlı bir fark olup olmadığını incelemek istiyoruz. "Aldığı tedavi" (therapy) 6 kategorisi olan kategorik bir değişkendir.

Başka bir örnek verecek olursak: Kliniğimizde dört farklı doktor hizmet veriyor. Hastalara bir anket uygulayıp hangi doktordan daha fazla memnun olduklarını soruyoruz. Sonuçlar şöyle olsun: Dr1 %22, Dr2: %24, Dr3: %30, Dr4: %24. Görünüşe göre doktor 3 daha fazla tercih edilmiş. Acaba aradaki fark istatistiksel açıdan anlamlı mı?

Binomiyal test

Kategorik değişken tek gruba diğer bir örnek de tercih durumlarıdır (Sayfa 272)

Örneğin diyabetli hastalarımıza iki farklı biguanid türü ilacı verdiğimiz ve kullanım kolaylığı açısından hangi ilacı tercih ettiklerini sorduğumuzu düşünelim. Bir sınıfta oylama yaptığımızı ve evet oylarının hayırlara göre %50'den farklı olup olmadığını incelemek istediğimizde yine binomiyal test yaparız.

Örnek 18

(Sayfa 333)

İki Nümerik Değişken

Pearson ve Spearman Korelasyon Analizi

Diyabetli hastalarımızda diyabet süresi (x) uzadıkça HbA1c düzeylerinin (y) de arttığından şüpheleniyoruz. Diyabet süresi ile HbA1c arasında doğrusal bir ilişki olup olmadığına bakmak istiyoruz.

Başka bir örnek: gelir düzeyi (nümerik sürekli değişken) azaldıkça yıllık hekime başvuru sayılarının (nümerik tam sayı değişken) arttığı yönünde bir hipotezimiz var. Bu hipotezi test etmek istiyoruz.

Örnek 19

(Sayfa 351)

Çoklu Çözümlemeler

Lojistik Regresyon Analizi

Diyabet hastalarımızın sigara içme durumunu (iki kategorili kategorik değişken) belirleyen faktörleri araştırmak istiyoruz. Böylece hangi hastalarımızın sigara içmeye daha meyilli olduğunu saptayabilir, gerekli klinik yaklaşımı sunabiliriz.

Örnek 20

(Sayfa 339)

Çoklu Çözümlemeler

Lineer Regresyon Analizi

Kişilerin vücut ağırlıklarını [nümerik değişken] bazı bağımsız değişkenleri kullanarak tahmin etmek istiyoruz. Araştırmamızda kişilerin yaşları, boyları, cinsiyetleri gibi değişkenler [hepsi nümerik] hakkında veri toplanmış. Bağımsız değişkenlerden faydalanarak kişilerin vücut ağırlıklarını tahmin edebilir miyiz?

Örnek 21

(Sayfa 392)

Duyarlılık - Özgüllük

ROC Analizi

Hastalarımızın kan şekeri kontrollerinin ne kadar kontrol altında olduğunu takip etmek için hemogloblin A1c (HbA1c) değerlerini ölçüyoruz. HbA1c değerlerine göre bireyleri 1-iyi kontrollü ve 0-KontROLSÜZ olarak sınıflandırdık (HbA1cGrup değişkeni). Diğer taraftan vücut kitle indeksi (VKİ) arttıkça bireylerin şeker kontrollerinin zorlaştığını biliyoruz.

Acaba VKİ değerlerine bakarak bireylerin kan şekerlerinin kontrol altında olup olmadığını tahmin edebilir miyiz? Böyle bir tahmin yapmak istesek hangi HbA1c değerini (kesme değeri) sınır olarak almalıyız?

Örnek 22

(Sayfa 399)

Metodolojik Çalışmalar

Geçerlilik ve Güvenilirlik Analizi

Yabancı dilde hazırlanmış bir ölçeği Türkçe'ye uyarlamak istiyorsunuz veya belli bir durumu ölçmek için siz bir psikometrik ölçek hazırlamak istiyorsunuz.

Örnek 23

(Sayfa 361)

Kaplan Meier Sağkalım Analizi

*Diyabetli hastalarımızda 60 ay boyunca **retinopati** gelişimini takip ettiğimizi ve cinsiyete göre (erkek/kadın) retinopati görülme durumu (var/yok) açısından fark olup olmadığını araştırdığımızı varsayalım (H_0 : Takip edilen süre içerisinde erkek ve kadınlarda retinopati gelişme durumu açısından fark yoktur). [Sonuç ölçütümüz (olay) retinopati olup olmaması iki kategorisi olan kategorik bir değişkendir. Zaman (nümerik) ve kategorik bir grup değişkeni (cinsiyet) değişkenleri söz konusudur].*

Örnek 24

(Sayfa 372)

Cox Regresyon Analizi

Diyabetli hastalarımızda 60 ay boyunca retinopati gelişimi açısından takip ettiğimizi ve cinsiyet ve eğitim durumunun retinopati görülmesi üzerindeki etkisini araştırdığımızı varsayalım. (H_0 : 60 aylık takip süresinde cinsiyet ve/veya eğitim durumunun retinopati gelişimi üzerinde etkisi yoktur).

Örnek 25

(Sayfa 379)

Nümerik Veri Uyum Analizi

(Sınıf İçi Korelasyon Katsayısı; Intraclass Correlation Coefficient – ICC)

Kliniğimizde takip edilen diyabet hastalarının tedaviye uyumları farklı asistanlar tarafından değerlendirilmiştir. Hastaların tedavi uyumu 0-10 puan arasında puanlanmıştır. Aynı hastaları değerlendiren asistanlar arasında uyumu araştırmak istiyoruz. İncelediğimiz değişken nümerik bir değişkendir. (H_0 : İki asistanın hastaların tedaviye uyum değerlendirmeleri sıfırdan farklı değildir).

Örnek 26

(Sayfa 379)

McNemar Bowker Testi

(İkiden fazla nominal kategorisi olan bağımlı iki ölçüm)

Kliniğimizde takip edilen diyabet hastalarının aldıkları tedavi çeşitlerinde bir ay öncesine göre anlamlı bir değişim olup olmadığını araştırmak istiyoruz. Her hastanın aldığı tedavi nominal olarak (tedavi almıyor/sadece sulfonilüre/sulfonilüre+biguanid/insülin) kodlanmış ve birer ay arayla iki kez kaydedilmiştir. (H_0 : Diyabet hastalarının aldıkları tedavi çeşitlerinde bir ay öncesine göre anlamlı bir değişim olmamıştır).

Bölüm 3: Veri Çeşitleri

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun veri çeşitleri hakkında bilgi sahibi olması amaçlanmaktadır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Aşağıdaki terimleri tanımlayabilmeli
 - Veri
 - Değişken
 - Parametre
 - Toplum
 - Örneklem
 - İstatistik
 - Censored (durdurulmuş) veri
- Verileri sınıflandırabilmeli
- Aşağıdaki türetilmiş veri çeşitlerini açıklayabilmeli
 - Yüzde
 - Oran
 - Orantı
 - Hız
 - Puan

*2500 öğrencisi olan bir okuldaki [evren ya da **toplum**] öğrencilerden rastgele seçilen 250 öğrencide [**örneklem**] boy ve ağırlık **değişkenleri** hakkında **veri** toplamaya karar verdik. Öğrencilerin boy ve kilolarını inceleyerek şişmanlık durumlarının **istatistik** analizini yapmayı hedefliyoruz.*

Araştırma yapmadaki öncelikli amaç bir konuyla ilgili veri (data) toplamaktır. Türkçe sözlükte “veri”, “Bir araştırmanın, bir tartışmanın, bir muhakemenin temeli olan ana öge, muta, done” olarak tanımlanmaktadır.

Verilerimizi elde ettiğimiz kaynağın tamamına “**toplum**” veya “**evren**” denir. Toplumun hakkında veri topladığımız özelliklerine ise “**değişken**” adı verilir.

İnceleyeceğimiz değişkeni toplumun tamamında araştırmak için kaynak, zaman ve imkanlarımız çoğunlukla yetersiz olacağından genelde toplumu temsil edebilecek bir grup birey yani “**örneklem**” hakkında veri toplarız. Eğer incelediğimiz değişken hakkında toplumun tamamından veri toplamışsak o

zaman “değişken” ifadesi yerine “**parametre**” kullanmak daha uygun olur. Makale yazımında buna dikkat etmeliyiz. Değişken yerine uygunsuz bir şekilde parametre terimini kullanmak sık yapılan yanlışlardır.

Elde ettiğimiz verileri yöntemli bir biçimde toplayıp sayı olarak belirtme, özetleme, analiz etme ve sonuç çıkarma işine ise “**istatistik**” denir.

Değişkenler **kategorik** ve **nümerik** olarak ikiye ayrılır.

Kategorik değişkenlerde her birey belli sınıflardan birine dahil edilebilir. Bu değişken türü de kendi içerisinde **Nominal** (Atanmış) ve **Ordinal** (Sıralı) olmak üzere ikiye ayrılır. Nominal değişkenlerde hangi kategorinin önce geleceği araştırmacının atamasına bağlıdır (1=bayan, 2=erkek olabileceği gibi 1=erkek, 2=bayan da olabilir). Sıralı değişkenlerde ise küçükten büyüğe, ya da büyükten küçüğe doğru bir sıra vardır (1=zayıf, 2=normal, 3=kilolu, 4=şişman, 5=aşırı şişman gibi).

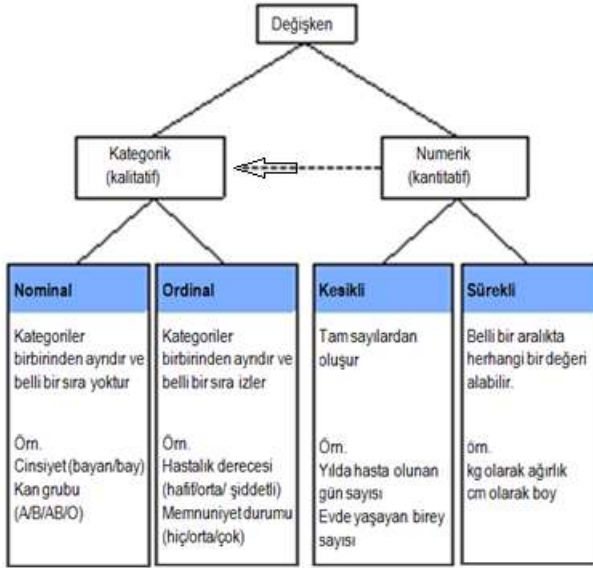
Bir kategorik değişkenin sadece iki özelliği olması halinde ona binary (**ikili**) veya dichotomous (**dikotom**) değişken denir (evet/hayır; canlı/ölü gibi).

Değişkenimiz, sınıflardan oluşmak yerine belli sayıları alan nümerik bir değişken de olabilir. Bu durumda da alınan değer **tam sayılı** (discreet) (Kaç kez doktora gittiniz? [4 kez] Hastane kaç yataklı? [60 yataklı] gibi) veya **sürekli** (continuous) (Boy [165,5 cm], Hemoglobulin değeri [14,3 mg/dl] gibi) olabilir.

Değişken çeşitlerini ayırt edebilmek, kullanılacak istatistiksel yöntemi seçerken ve SPSS’te analiz yaparken önemli olacak. Kategorik değişkenler sayı ve yüzde olarak gösterilirken nümerik değişkenler ortalama ve standart sapma gibi ölçütlerle gösterilir.

Kategorik ve nümerik verilerin ayırımı genellikle kesin olmakla birlikte bazen sıralı değişkenlerde kategori sayısı fazla olabilir. Bu durumlarda sıralı değişkenler kategorikmiş gibi muamele görebilir. Örneğin ağrı şiddetini araştırdığımız bir ankette 1’den 7’ye kadar bir görsel analog skala kullanmış olabiliriz. Tanım olarak bu, sıralı kategorik bir değişken olmakla birlikte çeşitli gruplar arasında ağrı şiddetini karşılaştırmak için ortalamaları almamızda bir sakınca yoktur.

Diğer taraftan, değişkenler SPSS tarafından ordinal, nominal ve nümerik olarak ayırt edilir (tam sayılı ve sürekli nümerik veriler aynı şekilde işlem görür).



Şekil 2: Veri çeşitleri.

Veri çeşitlerini ayırt etmedeki önemli bir nokta da veri toplama sırasında karşımıza çıkmaktadır. Değişkenimizin nümerik olması halinde veriyi de nümerik olarak toplamalıyız. Sınıflandırma yapmak istersek daha sonra kategorilere ayırabiliriz. Örneğin, yaş değişkenini sorgulayan bir ankette açık uçlu bir şekilde

“Kaç yaşındasınız?: _____” sorulmalıdır.

Sorumuz “Kaç yaşındasınız? a) 20’den az, b) 20-40, c) 41-60 d) 60’dan fazla”

şeklinde olmamalıdır.

Türetilmiş Veriler

Tıpta yüzde, oran, orantı ve hız gibi türetilmiş verilerle de sıkça karşılaşırız.

Hastalarımızın durumlarındaki değişimleri takip etmek için bulgularımızı genellikle **yüzde** olarak ifade ederiz. Örneğin

“Astımlı hastamızın FEV1 değerlerinde tedavi sonrasında %30 iyileşme oldu”

deriz.

Bir bütünün iki parçasının birbirine bölünmesiyle **oran** (ratio) elde edilir. Örneğin,

“Araştırmamızdaki bayan/erkek oranı 225/212 idi”.

Bütünün parçalarından birinin bütüne bölünmesine ise **orantı** (proportion) denilir. Örneğin,

“Araştırmamıza katılanların 225/437’si bayandı”. Orantıyı yüzde olarak ta ifade edebiliriz. Bu durumda “Araştırmamıza katılanların %51,5’i (225/437) bayandı.” deriz.

Orantının zamanla ilişkili olarak ifade edilmesi durumunda **hız**’dan (rate) bahsederiz. Hız, epidemiyolojik çalışmalarda sıkça kullanılan bir terimdir.

***Neonatal ölüm hızı** (bir yıl içerisinde 0-28 günlükken ölen bebek sayısı/ bir yıldaki toplam doğum sayısı), **kaba doğum hızı** (her 1.000 kişilik nüfus başına yıllık doğum sayısı)*

hız terimine örnek olarak verilebilir. Diğer taraftan dilimizde “hız” olarak ifade edilen *anne ölüm hızı* (belli bir dönemde meydana gelen anne ölümlerinin aynı dönemdeki canlı doğumlara bölünmesi) aslında bir orandır.

Durdurulmuş (Censored) Veri

Bazen verilerimizin kesin durumu hakkında bilgimiz olmayabilir. Bu durumda durdurulmuş (censored, gizli) veriden bahsedebiliriz. Verilerimizin durdurulmuş olması istatistik açıdan istenmeyen bir durumdur. Mümkünse verinin gerçek değerini elde etmeliyiz veya sonuçlarımızı yazarken bu durumu belirtmeliyiz. İki şekilde durdurulmuş veri söz konusu olabilir:

- Ölçümle elde edeceğimiz veriler için ölçüm aracımızın belli bir değerini üzerini veya aşağısını saptayamaması durumunda. Laboratuvarımızın sonuçları bildirirken aşırı yüksek ve aşırı düşük değerleri “hi” veya “low” olarak ifade edip gerçek değerini veremediği dikkatinizi çekmiştir. Parmaktan kolesterol ölçümü yapan cihazların genelde 70 mg/dl ve altındaki değerleri okuyamayıp “low” olarak belirtmesini örnek olarak verebiliriz.
- Bazen de araştırmamız tamamlanmadan çeşitli nedenlerle (ölüm, yer değiştirme, araştırmaya katılmamaya karar verme...) hastalarımız ayrılabilirler. Bu durumda ölçmek istediğimiz değişkenin son durumu hakkında veri elde edemeyebiliriz. Örneğin, iki farklı girişimin kilo vermedeki etkisini araştırdığımızı ve hastalarımızın her ay kilolarını ölçtüğümüzü düşünelim. Belli bir aşamada hastanın ayrılması durumunda araştırmamızın sonlanma noktasındaki veriyi elde edemeyeceğiz.

Bölüm 4: Veri Girişi

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun bilgisayara veri girişi konusunda bilgi sahibi olması ve SPSS kullanarak veri girebilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Veri girişi öncesi hazırlık yapmanın önemini kavramalı
- SPSS'te bir veri seti oluşturabilmeli
- ASCII terimini açıklayabilmeli
- Veri setleri arasında alışverişin önemini açıklayabilmeli
- Evet/Hayır değişkenlerini kodlamada sıfırın önemini açıklayabilmeli
- Birden fazla seçeneği olan (Multi coded) değişkenler için SPSS'te alan oluşturabilmeli
- Missing (eksik) verileri nasıl kodlayacağını açıklayabilmeli

Günümüzde araştırma verileri mutlaka bilgisayara girilerek analiz edilmektedir. Veri toplama aşamasından başlayarak bilgisayara girinceye kadar bir sistematik takip edilmeli ve belli kurallara uyulmalıdır. Aksi halde analiz sırasında güçlükler ve hatalar söz konusu olabilecektir.

Veri Girişi Yöntemleri

Veri girişi yapılırken daha sonra çeşitli yazılımlar arasında veri alışverişinin yapılması gerekebileceği akılda tutulmalıdır. Genelde araştırmacının kullanacağı istatistik yazılımı önceden bellidir. Ancak, farklı görüş almak veya dergi editörleri tarafından istenmesi halinde verilerin ikinci kişilere gönderilmesi gerekebilir. Bu amaçla veri girişi sırasında standart yöntemlerin kullanılması iyi olacaktır. En basit standart İngiliz alfabesine dayanan ve 128 karakterden oluşan ASCII (*American Standard Code for Information Interchange*) formatı veya metin dosyası olarak veri girişi yapılmasıdır. Burada her bireyin verileri ayrı bir satıra yazılır ve değişkenler arasında boşluk, tab veya virgül gibi ayraçlar kullanılır.

Günümüzdeki istatistik programlarının çoğu metin (text) veya Excel dosyası şeklinde veri çıkarmaya izin verdiklerinden veri dönüştürme ve ihraç etme konusunda genelde sorun yaşanmamaktadır.

SPSS 21'de Excel, Text, dBase, SAT, Data ve Lotus formatları da dahil olmak üzere 10'un üzerinde farklı dosya formatından veri alınabilmektedir [File>Open>Data>File type].

Ayrıca dBase, Access ve Excel formatında veri ihracı (export) yapılabilir [File> Save as>>File type>Excel 2007(.xlsx)].*

Veri Girişinin Planlanması

Veri girişinin, daha araştırmanın başlangıç aşamasında, veri toplama formu (anket) hazırlanırken planlanması gerekir. Verileri toplarken bilgisayara da nasıl gireceğimizi düşünmeliyiz: örn:

Yaş:	
Cinsiyet:	() Bayan () Erkek
Öğrenim durumu:	() İlkokul () Ortaokul () Lise () Üniversite
Hemoglobin değeri:	
Sigara içme durumu	() İçiyor () İçmiyor
Hangi ev aletlerine sahip?	() Buzdolabı () Çamaşır makinesi () Fırın () Bulaşık makinesi

Yerine

S1 Yaş:	
S2 Cinsiyet:	1() Bayan 2() Erkek
S3 Öğrenim durumu:	1() İlkokul 2() Ortaokul 3() Lise 4() Ünivers.
S4 Hemoglobin değeri (mg/dl):	
S5 Sigara içme durumu	0() İçmiyor 1() İçiyor
Hangi ev aletlerine sahip?	S6 Buzdolabı: 0 () Yok, 1 () Var S7 Çamaşır makinesi: 0 () Yok, 1 () Var S8 Fırın: 0 () Yok, 1 () Var S9 Bulaşık makinesi: 0 () Yok, 1 () Var

Şeklinde bir veri toplama formu hazırlanmalı. Böylece bilgisayarda her kategoriye karşılık hangi sayısal değer girileceği önceden belirlenmelidir. Aksi halde özellikle birden çok veri gircisinin (sekreterin) bulunduğu araştırmalarda veri girişi ve analiz aşamalarında sıkıntılarla karşılaşılabilir.

Kategorik Veriler

İstatistik programlarının çoğu nümerik verilerle işlem yapar. Kategorik verileri bilgisayara girerken düz metin yerine ilgili kategorileri temsil eden sayıları girmek gerekir.

Cinsiyet: Erkek

yerine

Cinsiyet: 1 (erkek)

girmek daha doğru olacaktır.

Aynı zamanda bu şekilde veri girişi daha hızlı olacaktır.

Evet/hayır şeklindeki iki kategorili (binary) veriler için “Hayır” seçeneğini “0”, Evet seçeneğini ise “1” olarak kodlamak daha doğru olacaktır. Aksi halde bazı analizlerde (örn. Lojistik regresyon analizi) bilgisayar ikili kategorik verileri (biz ne girersek girelim) 0 ve 1 olarak algılayacağından sonuçların yorumlanması sırasında kafa karışıklığı ortaya çıkabilir.

Değişkenimizde birden fazla seçeneğin olması durumunda:

- Az sayıda seçenek vardır ve bunlardan çoğunun seçilmesi söz konusu olabilir. Bu durumda seçenekleri tek tek evet/hayır sorusuna dönüştürmek daha doğru olur. Örn:

Hastanın solunum sistemiyle ilgili hangi şikayetleri var?

S1 Öksürük 0 () Yok, 1 () Var

S2 Nefes darlığı 0 () Yok, 1 () Var

S3 Hemoptizi 0 () Yok, 1 () Var

S4 Balgam 0 () Yok, 1 () Var

- Çok sayıda seçenek vardır ama bunlardan sadece birkaçının seçilmesi söz konusudur. Bu durumda farklı nominal kategorik değişkenler oluşturmak daha uygun olacaktır. Örn.

Hastanın hangi şikayetleri var?

S1 Belirti 1:.....

S2 Belirti 2:.....

S3 Belirti 3:.....

S4 Belirti 4:.....

Nümerik Veriler

Daha önce de belirtildiği gibi, nümerik veriler ölçüldükleri şekilde kaydedilmelidir. Birimlerin aynı olmasına dikkat edilmelidir:

“Kaç yaşındasınız? a) 20’den az, b) 20-40, c) 41-60 d) 60’dan fazla” şeklinde sorulmamalıdır.

“Kaç yaşındasınız?: _____ (yıl)” şeklinde sorulmalıdır.

Katılımcı 6 aylık bir bebekse veri “6 ay” olarak değil, “0,5 yıl” olarak girilmelidir.

Anketlerin Numaralandırılması

Bazen aynı birey için birden fazla anket bulunabilir. Örn. Kişinin kendi kendine doldurduğu demografik bilgiler anketi ve laboratuvar tarafından doldurulan kan tahlilleri sonuçları. Karışıklığı önlemek için her bireye bir numara verilmesi gerekir. Bu numara anketlerin üzerine yazılmalı, bilgisayara girilirken de önce anket numarası (birey kodu) girilmelidir. Böylece verilerle ilgili bir sorun olduğunda anket tekrar bulunup kontrol edilebilir, bireye ulaşılabilir, gerekirse ölçüm tekrarlanır.

Boş Verilerin Kodlanması

Bazen araştırmamızdaki soruların tamamı cevaplanmamış olabilir. Analiz sırasında boş alanların katılımcının dikkatinden kaçtığı veya veri toplamadaki eksiklikten mi kaynaklandığını yoksa bireyin bilerek cevaplamaması veya sorunun uygun olmamasından mı olduğunu bilmemiz gerekir. Bu nedenle boş veriler için özel bir kod seçebiliriz (Genelde 9, 99 veya 999 seçilir). Örn.

Hiç bağımlılık yapıcı madde kullandınız mı? 0 () Hayır, 1 () Evet
yerine

Hiç bağımlılık yapıcı madde kullandınız mı? 0 () Hayır, 1 () Evet 9 () CY
veya

S1 Bu giysi için ödediğiniz paraya değdi mi?

- 1 () Kesinlikle evet,
- 2 () Evet,
- 3 () Emin değilim,
- 4 () Hayır,
- 5 () Kesinlikle hayır

yerine

S1 Bu giysi için ödediğiniz paraya değdi mi?

- 1 () Kesinlikle evet,
- 2 () Evet,
- 3 () Emin değilim,
- 4 () Hayır,
- 5 () Kesinlikle hayır
- 9 () Uygun değil/cevapsız [kişi giysiye para vermemiş olabilir]

Aıştırma 1

1. Aşağıdaki anketi inceleyiniz. Anket hakkında düşününüz ve görüşlerinizi arkadaşlarınızla paylaşınız. Bu anket için bir SPSS veri seti hazırlayınız.

Diyabet Araştırma Anketi

1. Yaşınız:
2. Cinsiyetiniz:
 - a. Bayan
 - b. Erkek
3. Kaç yıldır diyabet hastasıınız?:
4. En son HbA1c değeriniz kaçtı?
5. Son iki hafta içerisinde hastalığınız nedeniyle günlük işlerinizi yapmakta ne kadar zorlandınız?
 1. () Hiç zorlanmadım
 2. () Biraz zorlandım
 3. () Epey zorlandım
 4. () Çok zorlandım
 5. () Hiçbir iş yapamadım
6. Diyabet için hangi ilaçları kullanıyorsunuz?
Yazınız.
 - a.....
 - b.....
 - c.....
 - d.....
 - e.....
 - f.....
7. Muayenenizde diyabete bağlı herhangi bir komplikasyon saptandı mı?
 - a. Evet
 - b. Hayır

2. Aşağıdaki doldurulmuş anketleri SPSS'e giriniz. Veri setinizi arkadaşlarınızla paylaşınız.

1

1. Yaşınız: 55
2. Cinsiyetiniz:
 - a. Bayan
3. Kaç yıldır diyabet hastasıınız?: 10
4. En son HbA1c değeriniz kaçtı? %8,9
5. Son iki hafta içerisinde hastalığınız nedeniyle günlük işlerinizi yapmakta ne kadar zorlandınız?
 - c. Epey zorlandım
6. Diyabet için hangi ilaçları kullanıyorsunuz? Yazınız.
 - a.Diamicron MR
 - b.....
 - c.....
 - d.....
 - e.....
7. Muayenenizde diyabete bağlı herhangi bir komplikasyon saptandı mı?
 - a. Evet

2

1. Yaşınız: 35
2. Cinsiyetiniz:
 - b. Erkek
3. Kaç yıldır diyabet hastasıınız?: 1
4. En son HbA1c değeriniz kaçtı? %7,3
5. Son iki hafta içerisinde hastalığınız nedeniyle günlük işlerinizi yapmakta ne kadar zorlandınız?
 - a. Hiç zorlanmadım
6. Diyabet için hangi ilaçları kullanıyorsunuz? Yazınız.
 - a.Diyet yapıyorum
 - b.....
 - c.....
 - d.....
 - e.....
7. Muayenenizde diyabete bağlı herhangi bir komplikasyon saptandı mı?
 - b. Hayır

3

1. Yaşınız: 65
2. Cinsiyetiniz:
 - b. Erkek
3. Kaç yıldır diyabet hastasıınız?: 15
4. En son HbA1c değeriniz kaçtı? %8,3
5. Son iki hafta içerisinde hastalığınız nedeniyle günlük işlerinizi yapmakta ne kadar zorlandınız?
 - c. Epey zorlandım
6. Diyabet için hangi ilaçları kullanıyorsunuz? Yazınız.

- a. İnsülin
- b. Diamicron MR
- c.....
- d.....
- e.....

7. Muayenenizde diyabete bağlı herhangi bir komplikasyon saptandı mı?

- a. Evet

4

1. Yaşınız: 47
2. Cinsiyetiniz:
 - a. Bayan
3. Kaç yıldır diyabet hastasıdır?: 3,5
4. En son HbA1c değerinizi kaçtı? %8
5. Son iki hafta içerisinde hastalığınız nedeniyle günlük işlerinizi yapmakta ne kadar zorlandınız?
 - d. Çok zorlandım
6. Diyabet için hangi ilaçları kullanıyorsunuz? Yazınız.
 - a. Glucophage
 - b. Diamicron MR
 - c.....
 - d.....
 - e.....
7. Muayenenizde diyabete bağlı herhangi bir komplikasyon saptandı mı?
 - a. Hayır

5

1. Yaşınız: 48
2. Cinsiyetiniz:
 - b. Erkek
3. Kaç yıldır diyabet hastasıdır?: 1
4. En son HbA1c değerinizi kaçtı? %7,5
5. Son iki hafta içerisinde hastalığınız nedeniyle günlük işlerinizi yapmakta ne kadar zorlandınız?
 - d. Hiç zorlanmadım
6. Diyabet için hangi ilaçları kullanıyorsunuz? Yazınız.
 - a. Glucophage
 - b.....
 - c.....
 - d.....
 - e.....
7. Muayenenizde diyabete bağlı herhangi bir komplikasyon saptandı mı?
 - a. Hayır

Çözüm için bakınız S. 464

Bölüm 5: Hata Ayıklama ve Uç Değerler

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucu veri girişinde hatalar ve uç değerler hakkında bilgi sahibi olmalıdır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- verilerle çalışmaya başlamadan önce hata kontrolü yapmanın önemini kavramalı
- verileri hatalara karşı tarama yöntemlerini sayabilmeli
 - frekans dağılımı
 - dağılım genişliği
- uç değerleri saptayabilmeli
 - frekans
 - grafik
- uç değerlerin nasıl kullanılacağını açıklayabilmeli

Araştırmalarımızda veri toplanması veya bilgisayara girilmesi aşamalarında hatalar söz konusu olabilir. Veri girişindeki kuralların dikkatle uygulanması haline hata olasılığı da azalacaktır. Bununla birlikte, hatalı veriler açısından veritabanımızın gözden geçirilmesi ve analiz aşamasına bundan sonra başlanması çok önemlidir. Analizimizi yapıp makalemizi yazdıktan sonra bazı verilerin yanlış girildiğini veya ölçüm sırasında hata yapıldığını bir düşününüz. Bu tür durumlarda analizleri tamamen yeniden yapmak bile gerekebilir...

En fazla hataya daktilo işlemi sırasında verileri bilgisayara girerken rastlıyoruz. Uzun veri formlarında daha hızlı veri girişi yapabilmek için bazen bilgisayar ekranına bakılmadan sırayla veriler girilebilir. Bu durumda bir değişken alanının atlanması halinde geriye kalan tüm veriler kaydırılmış olacaktır. Diğer bir hata da aynı tuşa birden fazla kez basılmasıyla olabilir. Bu durumda da 1 yerine 11 veya 111 gibi değerler girilmiş olabilir.

Hataları önlemek için verileri iki kişinin ayrı ayrı girmesi ve veritabanlarının karşılaştırılması yapılabilir.

Veri girişi sırasında azami dikkati gösterdikten sonra girilen verileri hatalar açısından yine de kontrol etmeliyiz.

Hata arama

Girilebilecek değerler sınırlı olduğundan kategorik değişkenleri hatalar açısından kontrol etmek nispeten daha kolaydır. Nümerik değişkenleri kontrol etmek daha zordur. Burada da verilerin belli aralıklarda olması gerektiğinden sıralama yaparak o aralığı aşan veri olup olmadığına bakılabilir. Örn.:

Verileri gözle tek tek kontrol edebiliriz. Büyük veritabanlarında bu zaman alacaktır. 1 Evet ve 2 Hayır olarak kodlanmış bir değişkende 1 ve 2 dışındaki değerleri bulmak kolaydır. Nümerik değişkeninde de uygunsuz verileri tahmin edebiliriz. Lise öğrencilerinde yaptığımız bir araştırmada yaş aralığının 14-20 yaş, hemoglobin değerinin 10-16 mg/dl arasında olmasını bekleriz.

SPSS'te hatalı verileri kolayca bulabilmek için büyükten küçüğe sıralama, frekans dağılımı ve dağılım genişliği özelliklerini kullanabiliriz. Örn.:

<https://drive.google.com/file/d/0B0ieO-Fy3z-5UTNWejZUU0EwLU0/view?usp=sharing>

adresinden indirebileceğiniz hataayıklama.sav veritabanında hata ayıklaması yapalım. Veriler “id” (anket no) değişkenine göre sıralanmış durumda. “age” (yaş) değişkenine bakacağız. Verileri bu değişkene göre sıralayarak kontrol edebiliriz:

Data>Sort cases>[“age” değişkenini “sort by” alanına geçirelim]>ok
Age değişkeninin 22 ile 99 arasında dağıldığını görüyoruz. Bu değerler normal olabilir. 99 yaşındaki bireyin anketini bulup (340 nolu anket) yaşını kontrol edebiliriz.

Başka bir yöntem “age” değişkeninin dağılım genişliğine bakmak olabilir:

Analyze>Descriptive Statistics>Descriptives>[“age” değişkenini “Variable(s)” alanına geçirelim]>ok

Age değişkenine 439 birey için veri girildiğini, en küçük yaşın 22, en büyük yaşın 99 olduğunu görüyoruz.

Bir de kategorik bir değişkene bakalım. “sex” (cinsiyet) değişkeni için de aynen yukarıdaki örnekte olduğu gibi sıralama yaparak kontrol edebiliriz. Başka bir yöntem de bu değişkenin frekans dağılımına bakmak olabilir:

Analyze>Descriptive Statistics>Frequencies>[“sex” değişkenini “Variable(s)” alanına geçirelim]>ok

Cinsiyet için 440 veri girildiğini, 241 adet 1 (Male), 195 adet 2 (Female), bir adet 3, bir adet 4, bir adet 11 ve bir adet te 22 girildiğini görüyoruz. 11 olarak girilen verinin 1 (Male), 22 olarak girilen verinin de 2 (Female) olma ihtimali yüksektir. Bu 4 veriyi de anket numaralarını bularak kontrol etmeli, hatayı bulup düzeltmeliyiz.

Eksik veriler (missing) varsa ne yapacağız?

Eksik verilerin birkaç nedeni olabilir: 1- birey reddettiği için veri alınamamıştır (örn. birey alkol kullanma durumunu belirtmek istemeyebilir), 2- bireye uygun olmadığı için veri alınamamıştır (örn. erkek katılımcı “doğum kontrol hapı kullanıyor musunuz?” sorusunu boş bırakacaktır), 3- veri alınmıştır ama bilgisayara girilmemiştir (sekreter hatası). Eksik verinin sebebi ne olursa olsun istenmeyen bir durumdur. Hatta veri eksikliği olan değişken esas araştırma konumuz (main outcome measure/ana sonuç ölçütü) ise bu durum daha da ciddidir. Hatta bazı analizler yapılamaz veya sonuçların güvenilirliği etkilenir. Örnek:

Hastalık şiddetini değerlendirmek için bir ölçüm yaptığımızı düşünelim. Eğer bazı bireylerin verilerinin eksik olduğunu fark edersek, bu hastaların muayeneye gelemeyecek kadar şiddetli hasta oldukları spekülasyonunu yapabiliriz.

Uç değerler

Uç değerler, diğer verilerle karşılaştırıldığında veri setine uygun olmadığı düşünülen aşırı değerlerdir. Bu aşırı değerler hatalı olabileceği gibi gerçeği de yansıtabilir. Bu nedenle hata ayıklaması yapılmalı ve doğruluğu kontrol edilmelidir. Örnek:

Bir bayanın 190 cm boyunda olması bir uç değerdir. Ancak, nadir de olsa bu durum mümkündür. Varsa bu bireyin yaş ve ağırlık verilerine de bakarak yorum yapabiliriz.

Uç değerlerin gerçeği yansıttığına karar verdiğimizde bu değerleri korumalıyız. Bir uç değer ancak şüpheli bulunması halinde silinmelidir.

Uç değerlerin fazla olması yapacağımız istatistiksel analizleri de etkileyebilir. Bu durumda verilere dönüştürme uygulayabiliriz veya nonparametrik bir test seçmemiz gerekir.

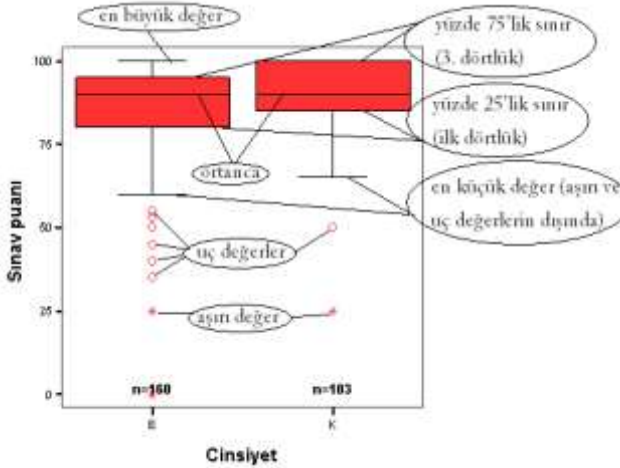
Uç değerlimizin olup olmadığını verileri sıralayarak gözümüzle kontrol edebiliriz. Bir yöntem de saplı kutu grafikleri (box plots) yaparak bakmaktır:

Açık olan veri setimizde; Graphs>Interactive>Boxplot [Y eksenine “Weight”, X eksenine “Marital status” değişkenlerini koyalım]>OK
Saplı kutu grafiklerinde sap kısmının dışında işaretlenen bireyler uç değerleri temsil etmektedir.

Saplı kutu grafikleri değişkenin 5 özelliği hakkında bilgi verir: en küçük değer (minimum), değerlerin %25’inin başladığı sınır (first quartile), ortanca (median) ve değerlerin %75’inin (third quartile) başladığı sınır. Örnek:

<https://drive.google.com/open?id=0B0ieO-Fy3z-5V0FmX2RZSmV4aUE> adresinden saplikutugrafik.sav dosyasını SPSS

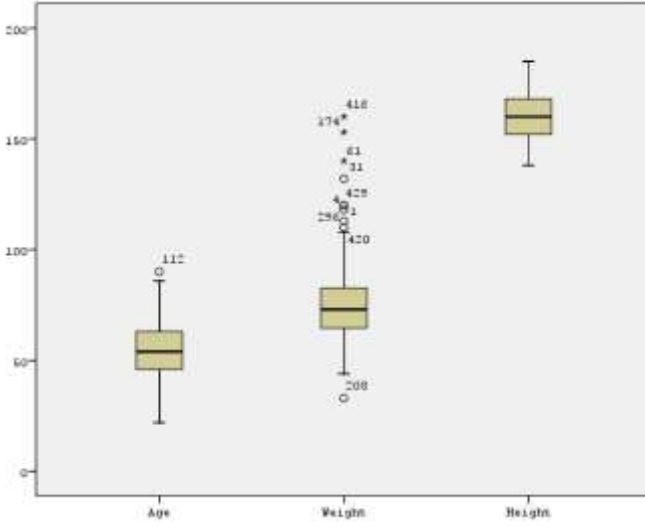
ile açınız. *Graphs>Interactive>Boxplot* [Y eksenine “Sınav puanı”, X eksenine “Cinsiyet” değişkenlerini koyalım]>OK



SPSS’te saplı kutu grafikleri çizildiğinde kutunun sap kısmının dışında uç değerler ve aşırı değerler de gösterilir. Veri kutudan 1,5 dördlük çeyrek değerler genişliği (interquartile range) uzaktaysa “**uç değer**” (outlier), 3 veya daha fazla dördlük çeyrek değerler genişliği uzaktaysa “**aşırı değer**” (extreme) olarak tanımlanır. SPSS çıktısında uç değerler daire ile, aşırı değerler ise yıldızla görülmektedir.

SPSS’te birden fazla değişken için aynı anda saplı kutu grafikleri çizdirip uç değerlere bakmak mümkündür. <https://drive.google.com/open?id=0B0ieO-Fy3z-5M0o1WV92NFpBWmM> veri setinde “age”, “weight” ve “height” değişkenlerinin saplı kutu grafiklerini çizdirelim:

Graphs>Legacy Dialogues>Boxplot [Simple, Summaries of separate variables seçip Define butonuna tıklayınız]>[“age”, “weight”, “height” değişkenlerini “Boxes represent:” alanına geçiriniz]>OK. Aşağıdaki grafiği elde ederiz:



Görüldüğü üzere “age” değişkeninde 112. bireyin yaşı (90 yaşında) uç değer olarak belirtilmiştir. “Weight” değişkeninde ise hem uç, hem de aşırı değerler vardır. Height değişkeninde aşırı ve uç değer saptanmamıştır.

Alıştırma 2

Bu alıştırma tartışmaya açmak istediğimiz noktalar okuduğumuz ilk 3 konuyla ilgili. <https://drive.google.com/file/d/0B0ieO-Fy3z-5bnplbGpRLVNJWGM/view?usp=sharing> <https://drive.google.com/file/d/0B0ieO-Fy3z-5bnplbGpRLVNJWGM/view?usp=sharing> odev2.sav adresindeki SPSS dosyasını indiriniz. Bu veriler "Suudi Arabistan'da 50 bin nüfuslu bir şehirde yaşayan işçilerin hepatit B bilgisini ve aşılı olma durumlarını araştırmak" amacıyla yapılmış bir araştırmadan alındı. Bu verileri

- "veri çeşitleri",
- "veri girişi" ve
- "hata ayıklama ve uç değerler"

açısından değerlendiriniz. Değerlendirirken sistematik olmaya çalışınız. İşlediğimiz konulara uygun yapılanları ve daha iyi nasıl olabileceğini, siz olsaydınız neleri farklı yapacağınızı belirtiniz.

Çözümü için bakınız S. 465

Bölüm 6: Verilerin Grafiklerle Gösterilmesi

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucu verilerin grafiklerle gösterilmesi konusunda bilgi sahibi olmalı ve SPSS kullanarak verileri grafiklerle gösterebilmelidir.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Temel grafik çeşitlerini sayabilmek
 - Çubuk (Bar) grafik
 - Pasta dilimi
 - Histogram
 - Nokta veya çizgi grafik
 - Saplı kutu grafiği (Boxplot, asansör grafiği)
 - Saçılma grafiği (scatter diagram, nokta grafikler)
- Verilere uygun grafik seçebilmek
 - Tek değişken olduğunda
 - Çok değişken olduğunda
 - Nümerik veri olduğunda
 - Kategorik veri olduğunda
- Temel grafikleri SPSS kullanarak yapabilmek
- Frekans dağılımının yönünü belirleyebilmeli
 - Sağa eğimli
 - Sola eğimli

Verilerimizi bilgisayara girdikten sonra öncelikli olarak yapmak isteyeceğimiz şey onları özetlemek ve bir şekilde “hissedilebilir” hale getirmektir. Bunun en güzel yolu da tablolar ve grafikler hazırlayarak özet istatistikler oluşturmaktır. Grafiklere bakarak daha hipotez testlerini uygulamadan ve önemlilik durumuna bakmadan durum hakkında bir fikir edinebiliriz. Bu bölümde beş farklı grafik çeşidi gösterilecektir.

- **Bar veya sütun grafikleri:** her bir kategori için ayrı bir yatay veya dikey sütun çizilir. Sütunun boyu değişkenin ilgili kategorideki sıklığı (frekansı) ile orantılıdır. Sütunlar arasındaki boşluklar değişkenlerin kategorik veya tam sayılı olduğunu gösterir. Örn.:

<https://drive.google.com/open?id=0B0ieO-Fy3z-5a3cydj1PQ0YtYk0>
dosyasını SPSS ile açınız. “marital status değişkeninin frekans dağılımına bakalım. sütun grafiğini çıkaralım. Analyze>Descriptive

Statistics>Frequencies>[“Marital status” değişkenini “Variable(s)” alanına geçirelim]>ok

Şu çıktıları elde ederiz:

Statistics

Marital status

N	Valid	420
	Missing	10

Bu değişken için 420 veri girilmiş. 10 adet de boş veri bulunuyor.

Marital status

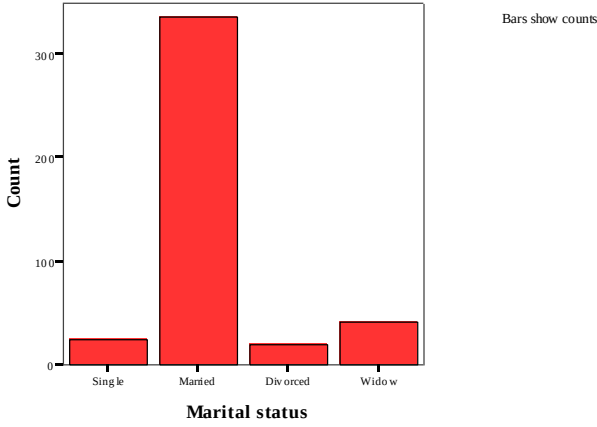
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Single	24	5,6	5,7	5,7
	Married	335	77,9	79,8	85,5
	Divorced	20	4,7	4,8	90,2
	Widow	41	9,5	9,8	100,0
	Total	420	97,7	100,0	
Missing	No response	3	,7		
	System	7	1,6		
	Total	10	2,3		
Total		430	100,0		

24 kişi bekarmış (Single). Bu 24 kişi toplamın %5,6’sını, eksik veriler çıkarıldığında geri kalanların ise %5,7’sini oluşturuyor. 3 kişi anketin bu sorusuna cevap vermemiş (no response). SPSS’te “No response” kategorisinin “Missing” başlığı altında 9 olarak kodlandığına dikkat ediniz. Ayrıca 7 cevapsız daha var. Bunlar “System missing” olarak belirtilmiş. Bu verilerin anketin uygulanması aşamasında veya bilgisayara girilmesi aşamasında unutulduğunu varsayabiliriz.

Şimdi aynı frekans dağılımını sütun grafiği ile gösterelim:

Graphs>Interactive>Bar [X eksenine “Marital status” değişkenini koyalım]>OK

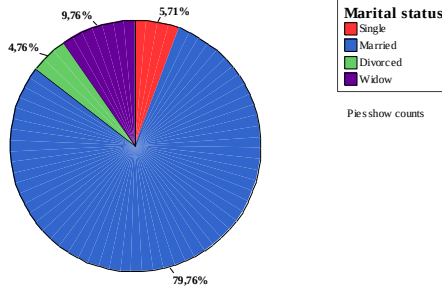
Aşağıdaki grafiği elde ederiz:



Grafiğimizde sütunların üzerinde sayı ve/veya yüzdelerin görülmesini istersek OK butonunu tıklamadan önce sekmelerden gerekli tercihleri yapabiliriz veya daha sonra grafik çıktısının üzerine çift tıklayarak da ayarları değiştirebiliriz.

- **Pasta grafiği:** daire şeklindeki bir pasta her bir dilimi değişkenin ilgili kategorisinin frekansını temsil edecek şekilde dilimlere ayrılır. Örn.:
<https://drive.google.com/open?id=0B0ieO-Fy3z-5M0o1WV92NFpBWmM> (diyabet.sav) dosyasını SPSS ile açınız.
“marital status değişkeninin pasta grafiğini yapalım.

Graphs>Interactive>Pie>Simple [Slice by kutusuna “Marital status” değişkenini sürükleyelim. Üstteki “Pies” sekmesini tıklayıp “Percent” kutucuğunu işaretleyelim]>OK. Aşağıdaki grafiği elde ederiz:



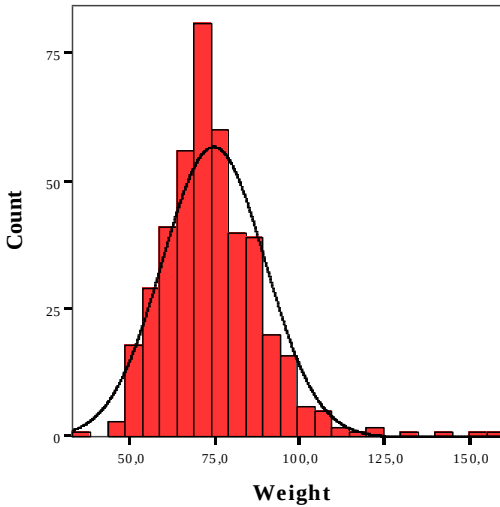
- **Histogram:** bu grafik de sütun grafiğine benzer. Ancak, sütunlar arasında boşluk yoktur. Sütunlar bir değeri değil, belli aralıktaki değerleri temsil eder. Bu grafik daha çok nümerik değişkenlerde verilerin dağılım şeklini incelemek için kullanılır. Örn.:

<https://drive.google.com/open?id=0B0ieO-Fy3z-5M0o1WV92NFpBWmM> (diyabet.sav) dosyasını SPSS ile açınız.

“Weight” değişkeninin histogram grafiğini çizelim.

Graphs>Interactive>Histogram [X eksenine “Weight” değişkenini sürükleyelim. Üstteki “Histogram” sekmesini tıklayıp “Normal curve” kutucuğunu işaretleyelim]>OK

Aşağıdaki grafiği elde ederiz:



Verilerimizin 75 kg civarında yoğunlaştığını, dağılımın çan eğrisine benzediğini, çan eğrisinin sağ taraftaki kuyruğunun biraz daha uzun olduğunu görüyoruz.

Frekans dağılımının şekli

Hangi istatistiksel testi uygulayacağımız büyük oranda verilerimizin dağılımına bağlıdır. Genelde verilerimizin dağılımı **unimodaldır**, yani tek bir pik noktası vardır. Bazen **bimodal** (iki ayrı pik noktası) ve **uniform** (tüm veriler eşit) dağılımlar da söz konusu olabilir. Unimodal verilerde verinin en büyük ve en küçük değerler arasında nasıl dağıldığını görmek önemlidir. Verilerimizin özellikle bakmak istediğimiz özellikleri şunlardır:

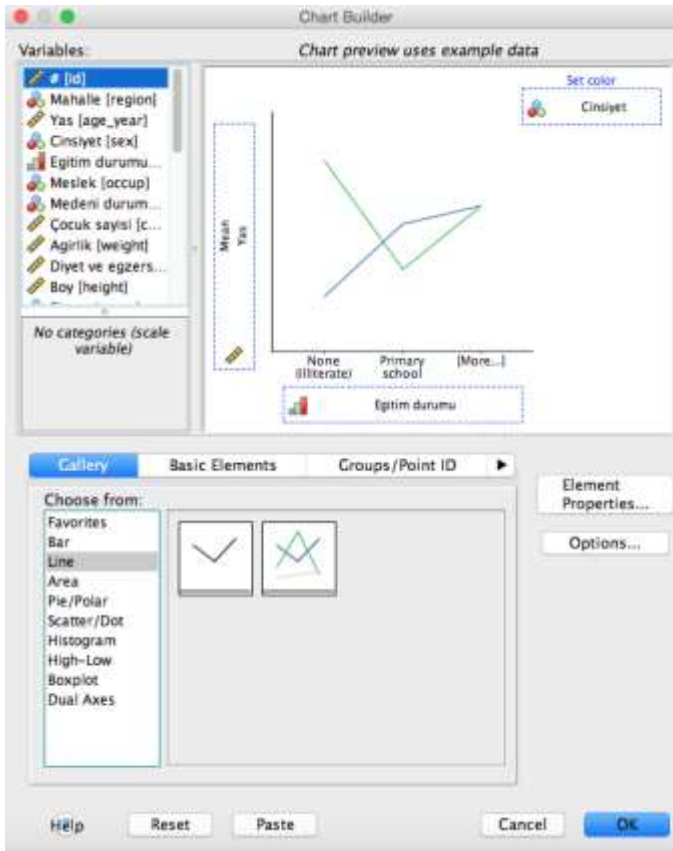
- **Simetrik dağılım:** bir orta değer etrafında eşit olarak dağılır. Histogram grafiğinde çan eğrisinin iki tarafı da simetriktir.
- **Sağa eğimli (pozitif eğimli):** histogram grafiğinde çan eğrisinin kuyruğu sağa doğrudur.

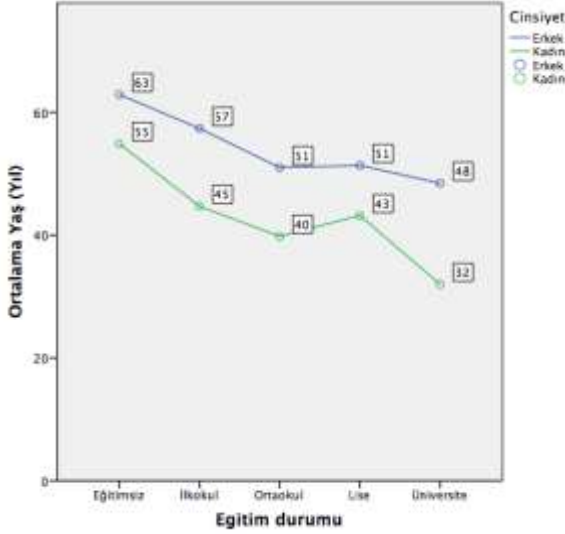
- **Sola eğimli** (negatif eğimli): histogram grafiğinde çan eğrisinin kuyruğu sola doğrudur.

- **Çizgi veya nokta grafik:** Noktalardan oluşur. Noktaların birleştirilmesiyle çizgiler elde edilir.

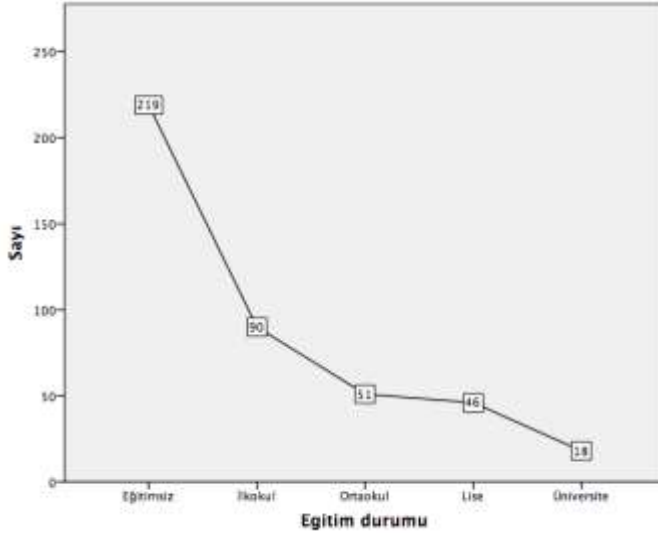
<https://drive.google.com/open?id=0B0ieO-Fy3z-5M0o1WV92NFpBWmM> (diyabet.sav) dosyasını SPSS ile açınız.
“Yaş” değişkeninin çizgi grafiğini çizelim.

Graphs>Chart Builder> [Line seçeneğini tıklayın. “Eğitim durumu” değişkenini X eksenine, “Yaş” değişkenini Y eksenine, “Cinsiyet” değişkenini Set color alanına sürükleyin.]>OK





Bu grafikte nümerik bir değişken olan yaşın ortalama değerlerinin sıralı kategorik bir değişken olan eğitim durumu açısından dağılımı cinsiyet açısından verilmektedir. Çizgi grafikler nümerik değişkenlerin ortalama değerlerini göstermek için uygundur. Ancak, aşağıdaki örnekte olduğu gibi, kategorik değişkenlerin kategorilerindeki birey sayılarını göstermek için de kullanılabilirler.

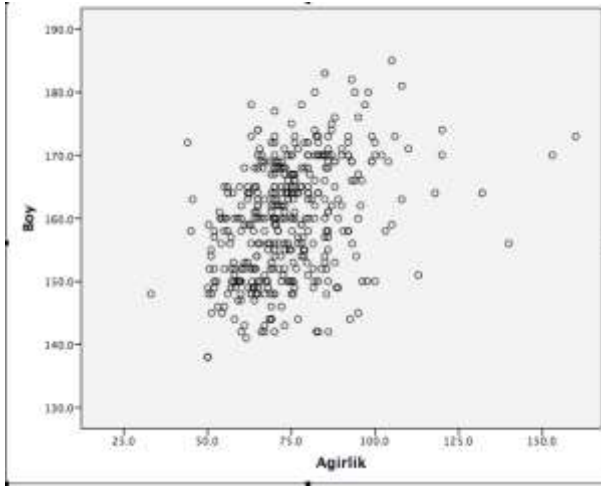


- **Boxplot (saplı kutu grafiği):** Asansör grafiği, kutu ve sakal grafiği (box and whisker plot) de denir. “Hata ayıklama ve uç değerler” konusuna bakınız.
- **Saçılma grafiği** (scatter diagram, nokta grafikler): iki ayrı nümerik veya ordinal değişkeni grafikte göstermenin bir yolu nokta grafikleridir. Örn.:

<https://drive.google.com/file/d/0B0ieO-Fy3z-5M0o1WV92NFpBWmM/view?usp=sharing> adresinden indirebileceğiniz diyabet.sav dosyasını SPSS ile açınız. “Height” ve “Weight” değişkenlerinin saçılma grafiğini çizelim.

Graphs>Interactive>Scatterplot [X eksenine “Height” değişkenini, Y eksenine de “Weight” değişkenini sürükleyelim]>OK

Aşağıdaki grafiği elde ederiz:



Grafikten boy arttıkça ağırlığın da arttığını görebiliyoruz. Gerekirse “Fit” sekmesinden iki değişken arasındaki ilişkiyi gösteren eğriyi de çizdirebiliriz.

Bölüm 7: Verilerin Tablolarla Gösterilmesi

Memet IŞIK

Amaç: Bu konu sonunda okuyucuların bilimsel çalışmalarda verilerin tablolarla gösterilmesi konusunda bilgi sahibi olmaları amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Bilimsel tablolarda olması gereken özellikleri sayabilmeli
- Bilimsel tablolarda sık yapılan hataları sayabilmeli
- SPSS kullanarak bilimsel tablo yapabilmeli

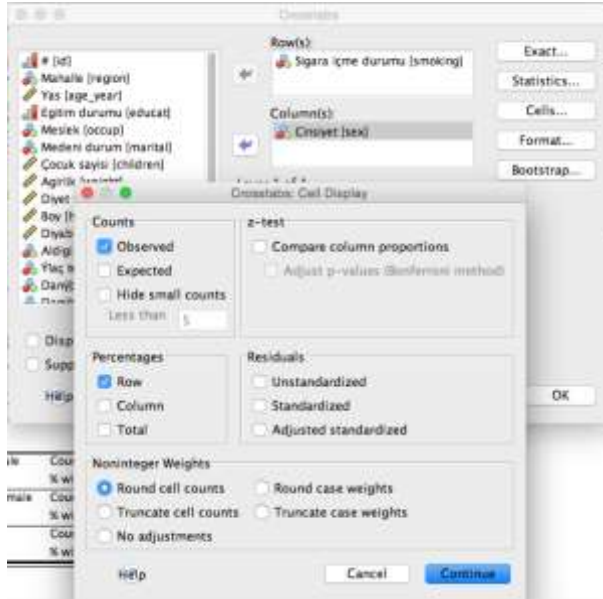
Genel prensipler

Tablo, ağırlıklı olarak nicel verilerin sütun ve satırlar halinde düzenlenmesiyle oluşturulan analitik bir gösterim şeklidir. Bu düzenleme çok sayıdaki verinin sınıflandırılmasını kolaylaştırdığı gibi veri grupları arasında karşılaştırma yapılması imkânını da sağlamaktadır. Tablolar dışında kalan grafik, fotoğraf, çizim vb. gösterimler ‘**şekil**’ olarak tanımlanmaktadır. Şekiller, çoğunlukla bütün bir şablonu betimlemek, unsurlar arasındaki ilişkileri ve kavramları açıklamak için kullanılır. Metin içerisinde, tablo ve şekillerde yer alan tüm bilgi ve bulgular tekrar edilmemelidir, sadece dikkat çekilmesi gereken noktalar vurgulanmalıdır. Bu nedenle tablo ve şekillerin kendi başlarına anlaşılabilir ve sade bir biçimde hazırlanmasına özen gösterilmelidir.

Verilerimizi tablo olarak sunmak için her zaman gerekmebilir. Anlaşılmayı kolaylaştırmak ve tekrarları önlemek için tablolara ihtiyaç duyulur. Mümkün olduğunca tablolar kolay anlaşılır ve sade sunulmalıdır. Gereksiz satır ve sütunlardan kaçınılmalıdır. Verilen bilginin üstten aşağı tarzda değil soldan sağa doğru bir okuma oluşturacak şekilde tablo düzenlenmelidir. Bunun sağlanması için bağımlı değişkenin sütunlara konulması ve yüzde olarak da satır yüzdelerinin sunulması gerekir.

Örneğin: Cinsiyete (bağımsız değişken) göre sigara içme (bağımlı değişken) durumlarının tablosunu yapmak istesek;

Analyze>Descriptive Statistics>Crosstabs[Rows kutusuna “smoking”, Columns kutusuna “sex” değişkenini alalım]>Cells>[Row kutucuğunu işaretleyelim]>Continue>OK



Aşağıdaki çıktıyı elde ederiz:

			Cinsiyet		
			Male	Female	Total
Sigara içme durumu	Yes	Count	33	1	34
		% within Sigara içme durumu	97.1%	2.9%	100.0%
	No	Count	187	186	373
		% within Sigara içme durumu	50.1%	49.9%	100.0%
Total		Count	220	187	407
		% within Sigara içme durumu	54.1%	45.9%	100.0%

Dikkat edilirse, bu tabloda bağımlı değişken satırlarda gösterilmiş. Yüzde olarak da satır yüzdeleri hesaplanmış (satır yüzde toplamı 100 yapıyor).

Bunun yerine tablonun aşağıdaki gibi olması önerilmektedir:

			Sigara içme durumu		Total
			Yes	No	
Cinsiyet	Male	Count	33	187	220
		% within Cinsiyet	15.0%	85.0%	100.0%
	Female	Count	1	186	187
		% within Cinsiyet	0.5%	99.5%	100.0%
Total		Count	34	373	407
		% within Cinsiyet	8.4%	91.6%	100.0%

Şimdi tabloyu okuyacak olsak: erkeklerin 33'ü (%15'i) sigara içmekte, 187'si ise (%85'i) sigara içmemektedir. Kadınların ise 1'i (%0,5) sigara içerken 186'sı (%99,5) içmemektedir.

Tablo ve şekiller, raporda ilk değinildiği sayfada veya yer yeterli gelmiyorsa bir sonraki sayfada yer almalıdır. Raporda tüm tablo ve şekillere atıfta bulunulmalıdır. Tablo ve şekiller, sayfa kenar boşluklarına taşmayacak şekilde, mümkünse sayfa yazı alanının sağ ve sol kenarlarına yaslanacak genişlikte sunulmalıdır. Taşma durumunda olanlar 'EKLER' bölümüne 'Ek' olarak verilmelidir.

Tablo ve şekil gösteriminde esas olan sadeliktir. Dolayısıyla özellikle tablolarda, dikey çizgilerin kullanılmaması, yatay çizgilerin ise mümkün olduğunca az kullanılması tercih edilmelidir. Tablolarda yer alan yazılı metinler, tablo hücrelerinde sola yaslı olarak yazılmalıdır. Tablolarda yer alan sayılar ve özellikle ondalıkla sayıların yazımında ise bu sayıların aynı basamaklarının alt alta gelmesine dikkat edilmelidir. Ondalıklar Türkçe yazarken virgül ile İngilizce ise nokta ile ayrılmalıdır. Tablo içeriğinde bir kısaltma kullanılmış ise bu kısaltma ile ilgili açıklama, 10 punto yazı karakteri ile hemen tablonun altına dipnot şeklinde verilmelidir. Benzer şekilde araştırma raporunda yer alan tablo, bir alıntı ise alıntının yapıldığı kaynak ile ilgili açıklama, 10 punto yazı karakteri ile hemen tablonun altına dipnot şeklinde verilmeli veya referans kaynak numarası verilerek, kaynaklarda gösterilmelidir.

Tüm tablo ve şekillerin kendine ait bir numarası ve adı olmalıdır. Belirlenen ad, tablo ve şeklin bir açıklaması niteliğindedir ve olabildiğince öz olmalıdır. Tercih edilmemesine rağmen belirlenen adın bir satırı aşması hâlinde, ikinci ve diğer satırlar birinci satır başı ile aynı sütundan başlamalı, blok yazım yapılmalıdır. Tablo ve şekil açıklama ve adlandırılmaları sonuna nokta veya virgül konulmamalıdır. Numaralama rakamlarla yapılmalıdır.

Tablo ve şekillerin numaraları, adları ve içeriği 11 punto Times New Roman yazı karakteri ile sola yaslı olarak, koyu ve baş harfler büyük olacak şekilde yazılmalıdır. Tabloların ve şekillerin adlandırılması, bunların farklı yerlerinde belirtilir:

- Şekiller için ad ve numara, alt kısma bir satır aralığı kullanılarak yazılmalı, şekil açıklaması ile şeklin alt kenarı arasında 1,5 satır aralığı boşluk bırakılmalıdır.
- Tablolar için ise ad ve numara, üst kısma bir satır aralığı kullanılarak yazılmalı, çizelge açıklaması son satırı ile çizelgenin üst kenarı arasında da 1,5 satır aralığı kadar boşluk bırakılmalıdır.

Metin içerisinde kullanılan tablo ve şekiller bir kaynaktan alınmış ise alıntı yapılan kaynak, tablo ve şekillerin alt kısmına, 10 punto Times New Roman yazı tipinde ve kaynak gösterme kurallarına uygun olarak yazılmalıdır.

SPSS'de tablo ayarları

<https://drive.google.com/file/d/0B0ieO-Fy3z-5M0o1WV92NFpBWmM/view?usp=sharing> adresinden indirebileceğiniz diyabet.sav dosyasını SPSS ile açınız.

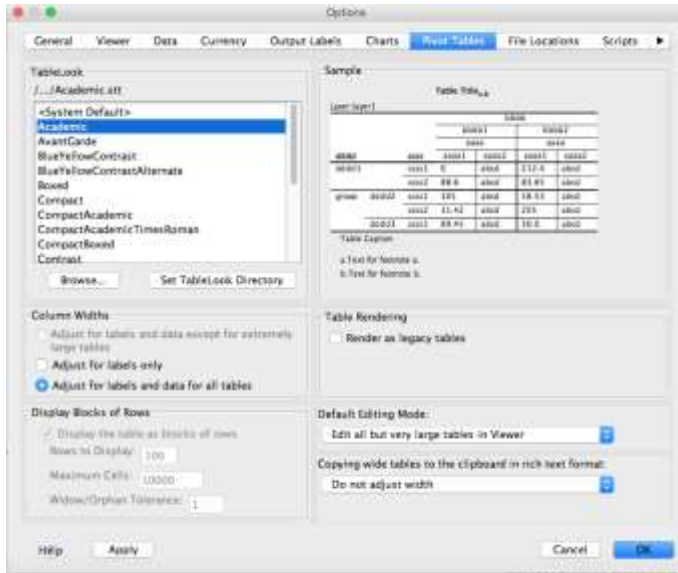
SPSS'de tablo yapmanın birden çok yolu vardır. Yapılan analizlerin tabloları otomatik olarak üretilmektedir. Bununla birlikte, özel tablolar yapılmak istenirse *Analyze>Tables>Custom Tables* yolundan yapılabilir.

SPSS'in başlangıç ayarlarında tablolarda hem yatay, hem de dikey çizgiler bulunmaktadır (Yukarıdaki örneğe bakınız). Tabloların görünümünü SPSS ayarlarından değiştirebiliriz. Şimdi tablo ayarlarımızı bilimsel görünüme dönüştürelim. Bunun için *SPSS Statistics>Preferences* sekmesini tıklayınız.



Sonra *Pivot Tables* sekmesinden *TableLook* ayarını *Academic* olarak değiştiriniz>OK.

Bundan sonra yapacağımız tablolar akademik görünümde olacaktır.



Yukarıda bağımlı değişken örneği için verdiğimiz tabloyu tekrar yapacak olursak şu şekilde bir çıktı elde ederiz:

Cinsiyet * Sigara içme durumu Crosstabulation					
		Sigara içme durumu			
		Yes	No	Total	
Cinsiyet	Male	Count	33	187	220
		% within Cinsiyet	15.0%	85.0%	100.0%
	Female	Count	1	186	187
		% within Cinsiyet	0.5%	99.5%	100.0%
Total		Count	34	373	407
		% within Cinsiyet	8.4%	91.6%	100.0%

SPSS'de özel tablolar

Analyze>Tables>Custom Tables

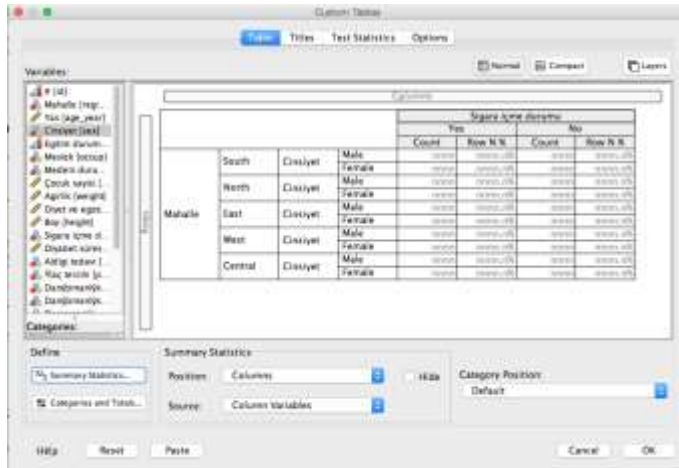
Gelen uyarı kutusunda tablo yapımının optimum verimde olabilmesi için başlangıçta SPSS verisetinin düzgün tanımlanmış olmasının öneminden bahsedilmektedir. Yani nümerik değişkenler, sıralı ve nominal kategorik değişkenler doğru olarak belirtilmiş olmalıdır:



Bu kutucuğu OK tıklayarak geçebilirsiniz.

Gelen ekranda sürükleyip bırakarak dilediğimiz şekilde tablolar yapabiliriz. Örneğin, sigara içme durumunu mahalleler ve cinsiyet açısından (iç içe geçmiş şekilde) göstermek istesek aşağıdaki şekilde yapabiliriz:

[Bağımlı değişkenin üzerine tıkladık, sonra Summary Statistics'i tıkladık. Buradan Count'a ek olarak "Row N%" istatistiğini de Display kutusuna aktardık]>Apply to Selection



>OK

Aşağıdaki çıktıyı elde ederiz:

				Sigara içme durumu			
				Yes		No	
				Count	Row N %	Count	Row N %
Mahalle	South	Cinsiyet	Male	11	12.1%	80	87.9%
			Female	1	1.0%	95	99.0%
	North	Cinsiyet	Male	0	0.0%	0	0.0%
			Female	0	0.0%	0	0.0%
	East	Cinsiyet	Male	4	10.0%	36	90.0%
			Female	0	0.0%	0	0.0%
	West	Cinsiyet	Male	0	0.0%	0	0.0%
			Female	0	0.0%	0	0.0%
	Central	Cinsiyet	Male	18	20.2%	71	79.8%
			Female	0	0.0%	91	100.0%

Bu şekilde yapacağımız tablolarda başlangıçta veri setinde değişkenlerin özellikleri düzgün tanımlanmışsa kategorik değişkenler için n ve %, nümerik değişkenler için ise ortalama ve standart sapma gibi tanımlayıcı istatistikleri içeren tablolar üretebiliriz.

Tablolarınızın üzerinde çalışmak isterseniz çift tıklamak suretiyle gerekli değişiklikleri yapabilirsiniz. Ayrıca tablonuzu kopyalayıp Microsoft Excel'e yapıştırarak burada daha farklı şekillere de dönüştürebilirsiniz.

Bölüm 8: Merkezi Eğilim (Yığılma) Ölçütleri

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucu merkezi eğilim ölçütleri hakkında bilgi sahibi olmaları ve SPSS kullanarak merkezi eğilim ölçütlerini hesaplayabilmeleri amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Merkezi eğilim ölçütlerini sayabilmeli
 - Aritmetik ortalama
 - Ortanca
 - Mod
 - Geometrik ortalama
 - Ağırlıklı ortalama
- SPSS kullanarak merkezi eğilim ölçütlerini hesaplayabilmeli
- SPSS kullanarak histogram grafiği çıkarabilmeli ve dağılımın durumunu değerlendirebilmeli
- Sağa eğimli dağılıma sahip değişkenlere logaritmik dönüşüm uygulayabilmeli
- Ağırlıklı ortalamayı açıklayabilmeli
- SPSS kullanarak ağırlıklandırma yapabilmeli

Elde ettiğimiz verileri bir şekilde özetlemeden verinin geneli hakkında fikir sahibi olamayız. Grafiklerle gösterme iyi bir özetleme yöntemi olabilir. Verinin belli özelliklerini hesaplayarak da genel bir fikir sahibi olabiliriz. Verilerimizi temsil edece bir değer bulabilir ve verilerin bu değer etrafındaki dağılımını hesaplayabiliriz. “merkezi eğilim ölçütleri” ve “yaygınlık ölçütleri” olarak ifade edilen bu ölçütler bu bölümde ve sonraki bölümde işlenecektir.

Aritmetik ortalama

Aritmetik ortalamaya sadece “ortalama” da denir; bütün verilerin toplanması ve veri sayısına bölünmesiyle elde edilir.

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

Formülde aritmetik ortalama x üzerinde çizgi ile gösterilir. Bu formül Yunancadaki Sigma toplama işareti kullanılarak

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad \text{veya} \quad \bar{x} = \frac{\sum x_i}{n} \quad \text{şeklinde gösterilebilir. Örn.:}$$

Birey no	1	2	3	4	5	6	7	8	9
Yaş	10	13	15	16	16	18	19	20	20

Yukarıdaki veri setinde yaşın aritmetik ortalaması
 $(10+13+15+16+16+18+19+20+20)/9 = 16,33$ 'tür.

Ortanca

Verilerimizi büyükten küçüğe doğru sıraladığımızda ortadaki değere **ortanca** denir. Eğer veri adedimiz çift sayı ise ortadaki iki değer ortalaması alınır. Örn.:

Birey no	1	2	3	4	5	6	7	8	9
Yaş	10	13	15	16	16	18	19	20	20

Yukarıdaki veri setinde yaşın ortanca değeri 5 numaralı bireyin yaşı olan 16'dır.

Birey no	1	2	3	4	5	6	7	8	9	10
Yaş	10	13	15	16	16	18	19	20	20	21

Yukarıdaki veri setinde ise 10 adet veri olduğundan ortanca 5 ve 6. Verilerin toplanıp ikiye bölünmesiyle bulunur: $(16+18)/2=17$

Mod

Mod, sık kullanılmayan bir merkezi eğilim ölçütüdür. Veri setinde en fazla tekrar eden değişkene mod denir. Veri setinde birden fazla mod olabilir. Her verinin sadece bir kez tekrarlaması halinde ise mod yoktur. Örn.:

Birey no	1	2	3	4	5	6	7	8	9	10
Yaş	10	13	16	16	16	18	19	20	20	21

Yukarıdaki veri setinde mod 16'dır.

Geometrik ortalama

Geometrik ortalamam için aşağıdaki adımların takip edilmesi gerekir:

1. Her birey için ilgili değişkenin **logaritması**nın alınması
2. Elde edilen logaritma değışlerinin **aritmetik ortalaması**nın alınması
3. Ortaya çıkan aritmetik ortalamanın **antilog**'unun hesaplanması

Verilerimizin eğimli olması halinde aritmetik ortalamayı kullanmak uygun değildir. Verilerin sağa eğimli olması haline (histogram grafiğinde çan eğrisinin kuyruğu sağa doğru) verilerin tek tek logaritmasını (10 tabanına veya e tabanına göre) alırsak elde edeceğimiz yeni veri seti simetrik hale gelebilir. Bu logaritma değerlerinin aritmetik ortalamasını alabiliriz. Tabii ki verinin orijinal birimine geri dönmesi için tekrar dönüşüm yapmak (antilog) gerekir. Bu yeni değere ise **geometrik ortalama** denir. Genelde geometrik ortalama ortancaya yakın ve aritmetik ortalamadan daha küçük bir değer olur. Örn.:

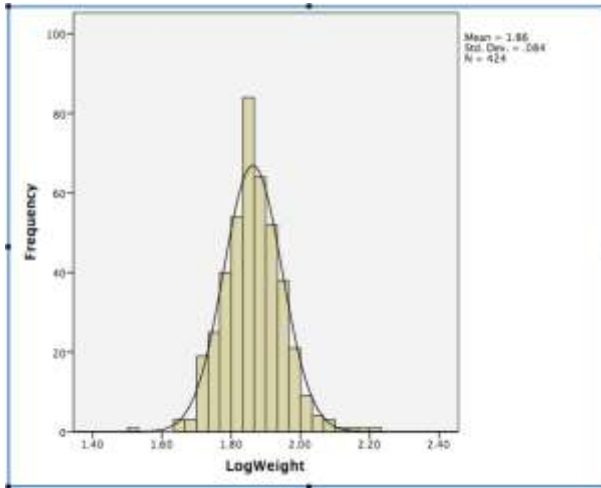
diyabet.sav veri setinde “Weight” değişkenini incelediğimizde histogram grafiğinde çan eğrisinin kuyruğunun sağa doğru olduğunu (sağa eğimli) görmüştük. Şimdi “weight” değişkeninin logaritmasını alarak yeni bir değişken oluşturalım:

Transform>Compute variable>[“Target Variable” alanına “LogWeight”, “Numeric Expression” alanına ise “LG10(weight)” yazalım]>OK

SPSS veri setimizde “LogWeight” adında yeni bir değişken oluşacaktır. Şimdi bu değişkenin histogram grafiğine bakalım:

Graphs>Interactive>Histogram [X eksenine “LogWeight” değişkenini sürükleyelim. Üstteki “Histogram” sekmesini tıklayıp “Normal curve” kutucuğunu işaretleyelim]>OK

Aşağıdaki grafiği elde ederiz:



Görüldüğü gibi çan eğrisi simetrik hale gelmiştir.

“Weight” değişkenimizin aritmetik ortalamasını hesaplayalım:

Analyze>Descriptive Statistics>Descriptives [“Weight” değişkenini “Variable(s)” alanına geçirelim]>OK

Aşağıdaki çıktıyı elde ederiz:

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Weight	424	33,0	160,0	74,266	15,1381
Valid N (listwise)	424				

Görüldüğü gibi “Weight” değişkeninin aritmetik ortalaması 74,266 kg’dır.

Şimdi de “LogWeight” değişkeninin aritmetik ortalamasını alalım:

Analyze>Descriptive Statistics>Descriptives [“LogWeight” değişkenini “Variable(s)” alanına geçirelim]>OK

Aşağıdaki çıktıyı elde ederiz:

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
LogWeight	424	1,52	2,20	1,8625	,08428
Valid N (listwise)	424				

Görüldüğü gibi “LogWeight” değişkeninin aritmetik ortalaması 1,862’dir.

Şimdi elde ettiğimiz değeri klinik açıdan yorumlayabilmek için Weight değişkeninin birimine geri çevirmemiz gerekiyor. 1,862 değerinin anti logaritmasını almalıyız. $\text{Antilog}(1,862) = 10^{1,862} = 72,777 \text{ kg}$.

“Weight” değişkeninin ortanca ve mod değerlerini SPSS ile hesaplamak için:

Analyze>Descriptive Statistics>Frequencies [“Weight” değişkenini “Variable(s)” alanına geçirelim. “Statistics” butonuna tıklayalım. “Central tendency” başlığının altından “Median” ve “Mode” kutucuklarını işaretleyelim]>Continue>OK

Not: aritmetik ortalamayı yukarıdaki menüde “Mean” kutucuğuna işaret koyarak da hesaplayabiliriz. Geometrik ortalamayı ise; Analyze>Compare Means>Means> menüsünden hesaplayabiliriz.

Aşağıdaki çıktıyı elde ederiz:

Statistics

Weight

N	Valid	424
	Missing	6
Median		72,750
Mode		70,0

Ağırlıklı ortalama

Bir değişkenin bazı değerlerinin diğerlerinden daha önemli olması halinde **ağırlıklı ortalama** kullanırız. Örneklemimizdeki her bir değere bir katsayı veririz. Her bir değerle katsayısını çarpıp bunları toplarız. Sonra katsayıların toplamına böleriz.

$$\frac{w_1x_1 + w_2x_2 + \dots + w_nx_n}{w_1 + w_2 + \dots + w_n} = \frac{\sum w_i x_i}{\sum w_i}$$

Örn. İlimizdeki hastanelerin günlük taburcu sayılarını inceliyoruz. Değişkenimiz “Bu gün hastanenizden kaç hasta taburcu oldu?” şeklinde olsun. İlimizdeki 3 hastane için elde ettiğimiz veriler aşağıdadır:

	Hastane 1	Hastane 2	Hastane 3
Taburcu edilen hasta	20	5	50

3 nolu hastanenin en fazla hasta taburcu ettiğini görüyoruz. Ortalama taburcu sayısı 25’dir. Bu hastanelerin yatak kapasitelerini bilmeden iş yoğunlukları hakkında fikir sahibi olamayız. Hasta kapasiteleri şu şekilde olsun:

	Hastane 1	Hastane 2	Hastane 3
Taburcu edilen hasta	20	5	50
Yatak kapasitesi	50	50	400

Taburcu sayılarını yatak kapasitesine göre ağırlıklandırmamız halinde daha geçerli bir fikir elde edebiliriz. Formülü uygulayacak olursak:

$(20 \times 50 + 5 \times 50 + 50 \times 400) / (50 + 50 + 400) = 42,5$ taburcu.

	Hastane 1	Hastane 2	Hastane 3	Ortalama
Taburcu edilen hasta	20	5	50	25
Yatak kapasitesi	50	50	400	166,6

Yatak kapasitesine göre ağırlıklandırılmış taburcu	66,6*	16,6	20,8	42,5

* $20 \times 166,6 / 50 = 66,6$. Burada basit orantı kurulmaktadır: [50 yatak kapasitesi olduğunda 20 hasta taburcu oluyorsa 166,6 yatak kapasitesi olsaydı kaç hasta taburcu olmalıydı?]

Sonuç olarak, bu üç hastanenin ağırlıklı taburcu ortalamasının 42,5 kişi olduğunu ve yatak kapasitesine göre ağırlıklandırma yapıldığında 3 nolu hastanenin zannedildiği gibi en yoğun çalışan hastane olmadığını, yatak kapasitesine göre en yoğun çalışan hastanenin 1 numaralı hastane olduğunu görüyoruz.

Şimdi de ağırlıklandırmayı SPSS ile yapalım. Çocuk sayısını [nümerik değişken] nümerik bir değişken olan yaş açısından ağırlıklandıracağız:

Çocuk sayısının yaş faktöründen etkilendiğini ve yaş ilerledikçe daha fazla çocuk sahibi olunacağını biliyoruz. diyabet.sav veri setinde “children” (çocuk sayısı) değişkenini yaşa göre ağırlıklandıralım. Ağırlıklandırma yapmadan önce “children” değişkeninin aritmetik ortalamasının 6,38 olduğuna dikkat ediniz.

Data>Weight Cases>Weight cases by>[“Age” değişkenini “Frequency Variable” alanına geçiriniz]>OK.

Şimdi çocuk sayısının (“children”) aritmetik ortalamasını aldığımızda 6,61 olduğunu görüyoruz.

Bu işleme “Yaşa göre düzeltilmiş çocuk sayısı” da denir. Uluslararası istatistiklerde ölüm oranları gibi istatistikler nüfusa veya başka değişkenlere göre düzeltilerek (ağırlıklandırılarak) verilir.

SPSS’de katagorik değişkenler için ağırlıklandırma yapmak biraz daha karmaşıktır. Bu konuya burada değinilmeyecektir.

Diğer Merkezi Eğilim Göstergeleri

Yukarıda belirtilenlerin dışında, sağıkta sık kullanılmayan diğer merkezi eğilim göstergeleri de vardır. Harmonik ortalama “değerlerin terslerinin aritmetik ortalamasının tersi” olarak tanımlanır ve daha çok orantılarla çalışıldığında

kullanılır. Örneğin, 1, 2 ve 4’ün harmonik ortalaması $\frac{3}{\frac{1}{1} + \frac{1}{2} + \frac{1}{4}}$ yani 12/7 olarak hesaplanır. Diğer taraftan, aritmetik ortalamasının uç değerlerden etkilenmesi

dikkate alınarak “istatistikte altın oran” gibi farklı hesaplama yöntemleri de önerilmiştir (Bakınız Günver MG, İstatistikte Altın Oran. Seçkin Yay. 2015).

Merkezi eğilim ölçütlerinin (MEÖ) olumlu ve olumsuz yönleri aşağıdaki tabloda görülmektedir:

MEÖ türü	Olumlu	Olumsuz
Aritmetik ortalama	- tüm değerler kullanılır - cebirsal olarak tanımlanmış ve matematiksel olarak kullanılabilir - bilinen örneklem dağılımı (bakınız Bölüm: veri dönüştürme)	- uç değerlerden etkilenir - eğimli verilerden etkilenir
Medyan (Ortanca)	- uç değerlerden etkilenmez - eğimli (skewed) veriden etkilenmez	- bilginin büyük bir kısmı ihmal edilir - cebirsal olarak tanımlanmamıştır - örneklem dağılımından etkilenir
Mod	kategorik veriler için kolayca saptanabilir	- bilginin büyük bir kısmı ihmal edilir - cebirsal olarak tanımlanmamıştır - örneklem dağılımı bilinmez
Geometrik ortalama	- geri dönüştürme yapılmadan önce aritmetik ortalamayla aynı avantajlara sahiptir - sağa doğru eğimli veriler için uygundur	sadece log dönüşüm simetrik bir dağılım oluşturuyorsa işe yarar

Ağırlıklı ortalama	• aritmetik ortalamayla aynı avantajlara sahiptir • her bir gözleme nisbi bir önem verilir • cebirsel olarak tanımlanmıştır	• ağırlık bilinmeli veya hesap edilmelidir

Aıştırma 3

<https://drive.google.com/file/d/0B0ieO-Fy3z-5T3ZzMmZaUFdtblk/view?usp=sharing>

adresinden odev3.sav veri setini açınız.

1. Bu verileri kullanarak SPSS'te "tkol" değişkeninin histogram ve saplı kutu (boxplot) grafiklerini çıkarınız.
2. Bu grafiklerde gördüklerinizi yorumlayınız.
3. " tkol " değişkenini log10 tabanına göre dönüştürerek yeni bir değişken (değişken "x") oluşturunuz.
4. Değişken "x" in histogram grafiğini çizin ve " tkol " değişkeninin histogram grafiği ile karşılaştırınız.
5. tkol değişkeninin aritmetik ortalamasını alınız.
6. tkol değişkeninin geometrik ortalamasını alınız.
7. Kilo değişkeninin aritmetik ortalaması, medyan (ortanca) ve mod değerleri kaçır? (analyze-dercriptive statistics-frequencies sekmesinden). Ayrıca ortancaya verileri sıralayarak da bakınız.
8. Boya göre ağırlıklandırılmış (düzeltilmiş) kilo değişkeninin aritmetik ortalaması kaçır?
9. Araştırmada kaç erkek, kaç kadın vardır? Boy değişkeninin aritmetik ortalaması kaçır? Cinsiyete göre ağırlıklandırma yapınız. Şimdi boy değişkeninin aritmetik ortalaması kaç oldu?

Çözüm için bakınız s. 466

Bölüm 9: Yaygınlık Ölçütleri

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun yaygınlık ölçütleri hakkında bilgi sahibi olması ve SPSS kullanarak yaygınlık ölçütlerini hesaplayabilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Yaygınlık ölçütlerini sayabilmek
- SPSS kullanarak
 - aralık (range)
 - persantil
 - varyans
 - standart sapma hesaplayabilmek
- Varyans ve standart sapma formülünü söyleyebilmek
- Birey içi ve bireyler arası varyasyonu açıklayabilmek

Bir nümerik değişken hakkında biri merkezi eğilim diğeri de yaygınlık ölçütü olmak üzere iki özelliğini belirtmemiz halinde verilerimizin yapısını yeterince özetlemiş oluruz. Bir önceki bölümde merkezi eğilim ölçütlerinden bahsedilmişti. Şimdi sıra yaygınlık ölçütlerinde.

Aralık (range)

Verilerimizin en büyük ve en küçük değeri arasındaki farka **range** denir. Aralık yerine genelde en küçük (min.) ve en büyük (max.) değerler verilir. Uç değerlerimizin fazla olması halinde aralık ölçütünün yeterince güvenilir olmayacağına dikkat edilmelidir.

Persantil aralıkları

Verilerimizi küçükten büyüğe doğru sıraladığımızda veri adedinin %1'inin bulunduğu kısma 1. persantil, yüzde 50'sinin bulunduğu sınıra 50. persantil denir.

1. Çeyreğin yeri veriler sıralandığında: $(n+1)/4$ formülünden, 3. çeyreğin yeri ise $3*1$. Çeyrek) olarak hesaplanır.

Örn.:

8 bireyin boy ölçümlerinden oluşan bir veri setimiz olsun:

No	Boy
1	145
2	148

3	154
4	160
5	166
6	170
7	176
8	182

Bu veri setinde 1. çeyreğin yeri: $(8+1)/4=2.25$. değerdir.

Yani 1. çeyrek, ikinci ve üçüncü değerler arasında yer almaktadır. Tam değerini hesaplamak için 2. değer ile 3. değer arasındaki farkı 0,25 ile çarpıp 2. değere eklememiz gerekir:

$$1. \text{ çeyrek} = [148 + ((154-148)*0,25)] = 149,5$$

$$3. \text{ çeyreğin yeri} = 3*1. \text{ çeyreğin yeri} = 3*2,25 = 6,75. \text{ değer}$$

$$3. \text{ çeyrek} = [170 + ((176-170)*0,75)] = 174,5$$

Tam %50 sınırındaki değere “ortanca” denir. 25-75. Persantiller arasına ise **interquartile range** (dörtlük çeyrek değerler genişliği-DÇDG) denir. DÇDG, veriler sıralandığında orta kısımda kalan %50’lik bölümü gösterir. Verimizin toplumu temsil edecek kadar büyük bir örneklemden gelmesi halinde her iki uçtaki %2,5’lik kısmın içerisinde kalan değerlere **referans aralığı**, **referans genişliği** veya **normal aralık** denir. Laboratuvar vs. ölçümlerde değerimizin toplumla karşılaştırıldığında normal olup olmadığına bu aralığa bakarak karar veririz.

Varyans

Verilerin dağılımını ölçmenin bir yolu, her bir gözlemin aritmetik ortalamadan ne kadar sapma gösterdiğine bakmaktır. Elde edeceğimiz değerlerin ortalamasını alamayız zira artı taraftakiler eksilerle yaklaşık aynı olacağından birbirini götürür. Bunun yerine her bir değer aritmetik ortalamadan olan uzaklığının karesini alarak bir hesap yaparız. Bu değerleri toplar ve (örneklem sayısı $[n]-1$)’e böleriz. Buna **varyans** hesabı denir. Varyans s^2 ile gösterilir.

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

Varyansı hesaplarken aritmetik ortalamadan farklı olarak $(n-1)$ ’e bölüyoruz. Bunun nedeni, evrenin tamamında değil, belli bir örneklem üzerinde çalışmamızdır. Bu durumda teorik olarak toplum değerine yakın bir varyans elde ettiğimiz gösterilmiştir.

Standart sapma

Standart sapma varyansın kareköküdür.

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

Standart sapmayı verilerin ortalamadan sapma dereceleri olarak düşünebiliriz.

Standart sapmayı ortalamaya böler ve bunu yüzde olarak ifade edersek, **varyans katsayısını** (coefficient of variation) buluruz. Varyans katsayısının avantajı değişkenin biriminden etkilenmemesidir (% olarak ifade edilmesi) ancak, teorik dezavantajları nedeniyle genelde tercih edilmez.

Bir veri seti ile örneklendirecek olursak:

Birey No	Yaş (x _i)	Ort ($\bar{x}$)	x _i - $\bar{x}$	(x _i - $\bar{x}$) ²
1	14	30	-16	256
2	25	30	-5	25
3	38	30	8	64
4	41	30	11	121
5	22	30	-8	64
6	41	30	11	121
7	25	30	-5	25
8	26	30	-4	16
9	33	30	3	9
10	35	30	5	25
Toplam	300	300	0	726

$$n - 1 = 10 - 1 = 9$$

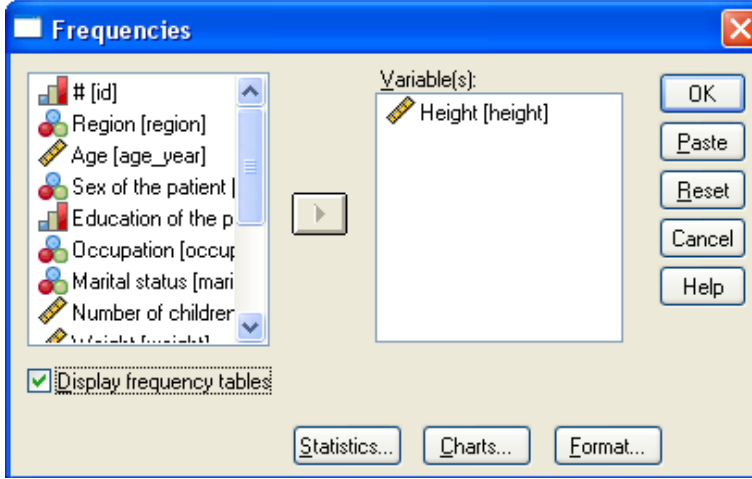
$$\text{Varyans} = 726 / 9 = 80,67$$

$$\text{Standart sapma} = \sqrt{80,67} = 8,98$$

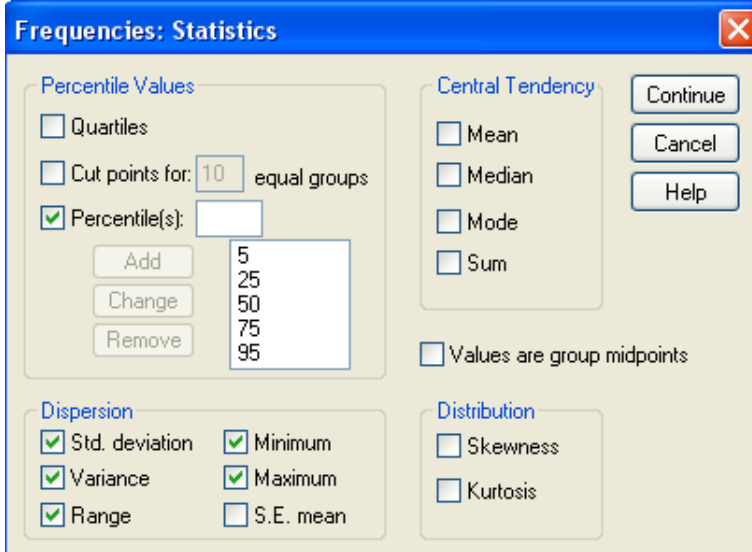
SPSS ile hesaplama

Diyabet.sav (<https://drive.google.com/open?id=0B0ieO-Fy3z-5a3cydjIPQ0YtYk0>) veri setini kullanarak “Height” değişkeni için yaygınlık ölçütlerini hesaplayalım:

Analyze>Descriptive Statistics>Frequencies>[“Height” değişkenini “Variable(s)” alanına geçirelim]



>Statistics>”Percentile(s)” kutucuğunu işaretleyelim. 5, 25, 50, 75 ve 95 persantil değerlerini tek tek kutuya yazıp her seferinde “Add” yapalım > “Std. deviation”, “Variance”, “Range”, “Minimum” ve “Maximum” kutucuklarını işaretleyelim



> Continue > “Display frequency tables” kutucuğundan işareti kaldıralım > ok

Aşağıdaki çıktıyı elde ederiz:

Statistics**Height**

N	Valid	406
	Missing	24
Std. Deviation		9,2030
Variance		84,696
Range		47,0
Minimum		138,0
Maximum		185,0
Percentiles	5	145,000
	25	152,000
	50	160,000
	75	168,000
	95	174,000

“Height” değişkenimizin varyansı 84,6 cm², standart sapması 9,2 cm, en küçük değeri 138 cm, en büyük değeri 185 cm, aralığı 47 cm’dir. Birinci çeyrek 152 cm’de, 3. çeyrek ise 168 cm’dedir.

Birey içi ve bireyler arası farklılıklar

Aynı bireyde birden çok ölçüm yapmamız halinde farklı sonuçlar elde edebiliriz (**birey içi farklılık**). Bu fark, bireyin her defasında aynı cevabı vermemesi veya ölçüm hatasından kaynaklanabilir. Bununla birlikte, birey içi farklılık aynı ölçümün bir gruptaki farklı bireylerde yapılan ölçümler arasındaki farklılığa (**bireyler arası farklılık**) göre daha azdır. Araştırma tasarımı sırasında bu farklılıklar önem arz edecektir.

Yaygınlık ölçütü	Olumlu yönleri	Olumsuz yönleri
Aralık (Range)	Kolayca saptanabilir	- Sadece iki gözlemi kullanır - Uç değerlerden etkilenir - Örneklem sayısı arttıkça artma eğilimindedir

Persantillere dayalı aralık	• Genelde uç değerlerden etkilenmez • Örneklem sayısından bağımsızdır • Eğimli veriler için uygundur	• Hesaplanması hantaldır • Küçük örnekler için hesaplanamaz • Sadece iki gözlem kullanır • Cebirsel olarak tanımlanmamıştır
Varyans	• Her gözlemi dikkate alır • Cebirsel olarak tanımlanmıştır	• Ölçüm birimi ham verinin karesidir • Uç değerlerden etkilenir • Eğimli veriler için uygun değildir
Standart sapma	• Varyansla aynı avantajlara sahip • Ölçüm birimi ham verininkiyle aynıdır • Kolayca yorumlanabilir	• Uç değerlerden etkilenir • Eğimli veriler için uygun değildir

Bölüm 10: Teorik Dağılımlar: Normal Dağılım

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun yaygınlık teorik dağılımlar ve normal dağılım hakkında bilgi sahibi olması amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Empirik (deneysel) dağılım ve probability (ihtimal, olasılık) dağılımını açıklayabilmeli
- Olasılık çeşitlerini açıklayabilmeli
 - Subjektif
 - Frequentist (sıklığa göre)
 - A priori (önsel)
- Olasılık kurallarını açıklayabilmeli
 - Toplama kuralı (mutually exclusive) (VEYA kuralı)
 - Çarpma kuralı (VE kuralı)
- Normal dağılımı açıklayabilmeli
- Standart normal dağılımı açıklayabilmeli

Daha önce ampirik (deneysel) frekans dağılımının nasıl oluşturulacağını görmüştük. Şimdi teorik dağılımlardan bahsedeceğiz.

İstatistikte yüzlerce teorik dağılım tanımlanmıştır. Biz bunların içerisinde en sık kullanılanları inceleyeceğiz.

Teorik dağılımlar matematiksel bir modelle ifade edilir. Elde ettiğimiz verilerin frekans dağılımının teorik dağılımlardan birisine benzemesi halinde o teorik dağılımla ilgili teorik bilgileri kullanarak elde ettiğimiz veriler hakkında karar veririz. İstatistiksel testlerimizin çoğu teorik dağılımlarla ilgili bilgilerimize dayanır. İstatistiksel önemlilik vs. kararlarımızı verirken hep teorik dağılımlardan yararlanırız. Teorik dağılımları ve olasılık hesaplarını anlamamız istatistiksel testlerin nasıl çalıştığını anlamamıza yardımcı olacaktır.

Olasılık hesapları

Olasılık hesabı belirsizlikleri tahmin etmede kullanılır. Dolayısıyla istatistiğin kalbinde yer alır. Olasılık, bir olayın olma veya olmama ihtimalini verir; sıfır ile bir arasındadır. Eğer olasılık *sıfır* ise olay gerçekleşemez. Olasılığın *bir* olması halinde ise olay kesinlikle gerçekleşir. Bir de **şarhli olasılık** (conditional probability) vardır. Şarhli olasılık, önceki bir olayın gerçekleşme durumuna göre şimdiki olayın olasılığını hesaplar.

Olasılığı 3 şekilde hesaplarız:

- **Subjektif:** olayın gerçekleşme durumuyla ilgili kişisel inancımızdır. 2012 yılında kıyametin kopacağına inanmak gibi.
- **Frekans hesabı:** deneyimizi tekrarlamamız halinde olayın meydana gelme olasılığıdır. Örneğin bir parayı 1000 kez attığımızda yazı gelme olasılığı.
- **A priori (önsel):** olasılık dağılımının önceden bilinen bir modeline göre hesaplanır. Kalıtım teorisiyle ilgili oluşturacağımız modellere dayanarak mavi gözlü bir anne ve kahverengi gözlü bir babadan olacak çocuğun göz rengini tahmin edebiliriz.

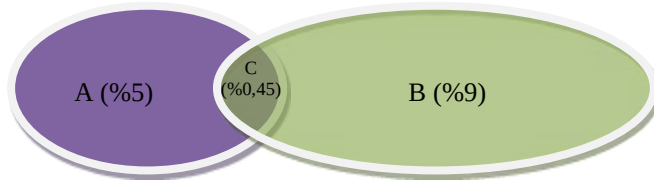
Olasılık kuralları

Olasılık hesabında çarpma ve toplama kuralları uygulanır.

- **Toplama kuralı:** Ya A olayı gerçekleşecektir veya B. Örn.
Bir kişide uyku apne sendromu (UAS) görülme olasılığı %5, diyabet görülme olasılığı ise %9'dur.
Bu iki durumun bir arada olma ihtimali olmasaydı (yani A ve B olayları birbirini dışlayan (mutually exclusive) olsalardı) bu kişide UAS veya diyabet görülme olasılığı $\%5 + \%9 = \%14$ olurdu.
- **Çarpma kuralı:** A ve B olaylarının birbirinden bağımsız olması durumunda geçerlidir. A olayı ve B olayının bir arada gerçekleşme olasılığıdır. Yukarıdaki örneklerle göre yapacak olursak:

Bir kişide UAS ve diyabetin birarada olma olasılığı ise $\%5 * \%9 = \%0,0045$ olacaktır.

Bu iki hastalığın birarada olma ihtimali de olduğuna göre bir kişide UAS veya diyabet görülme olasılığı $\%5 + \%9 - \%0,0045 = \%13,99$ olacaktır.



Benzer bir olasılık örneğini kan tahlilleri için verebiliriz: Bir kan testinin (A testi) hatalı olması olasılığı yaklaşık %3 olsun. Aynı hastadan ikinci bir test istesek (B testi) ve bu testin de hatalı olma olasılığı %3 olsa A veya B testinin hatalı olma olasılığı $= (\%3 + \%3) - (\%3 * \%3) = \%5,99$ olacaktır.

Olasılık dağılımlarının teorisi

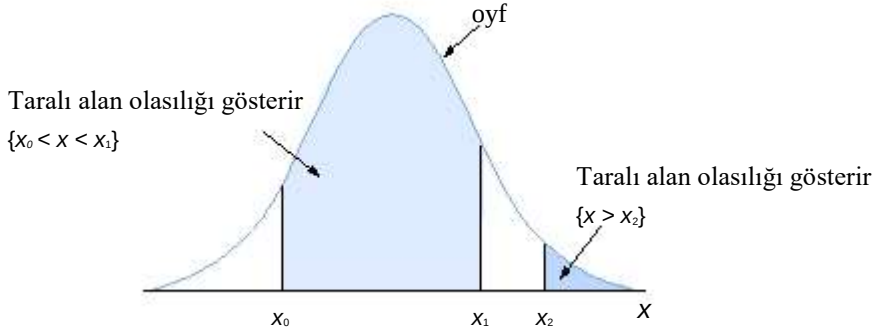
Bir **rastgele değişken**, belli olasılık çerçevesinde birbirinden farklı çeşitli değerleri alabilir. **Olasılık dağılımı** ise bu rastgele değişkenin alabileceği tüm değerlerin olasılıklarını verir. Bu dağılım matematiksel olarak ifade edilebilen teorik bir dağılımdır. Teorik dağılımın da ampirik dağılımdaki gibi ortalama ve standart sapması vardır.

Her bir olasılık dağılımı belirli **parametrelerle** (örn. Ortalama, varyans..) tanımlanır. Bu parametreleri bilmemiz halinde ilgili teorik dağılımı tanımlayabiliriz. İşte istatistikte örneklemimizdeki bu parametreler tahmin edilerek hesap yapılır.

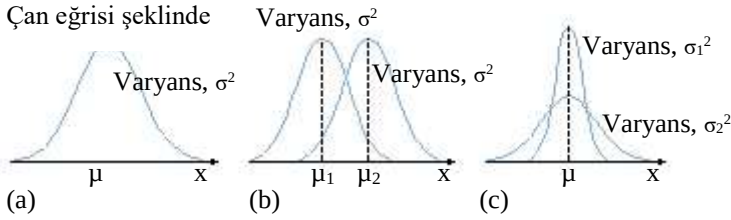
Değişkenimizin tam sayı veya sürekli olmasına göre teorik dağılımlar da tam sayı veya sürekli olabilir.

- **Tam sayı (discrete) teorik dağılımlar:** rastgele değişkenin mümkün olan her bir değeri için olasılıklar hesaplayabiliriz. Bu olasılıkların toplamı birdir. Binomiyal ve Poisson dağılımı bu tür teorik dağılımlardandır.
- **Sürekli (continuous) teorik dağılımlar:** x değişkeni sonsuz sayıda değer alabileceğinden ancak x 'in belli aralıktaki değerleri için olasılığı hesaplayabiliriz. x değişkeninin alacağı değerleri yatay ekseninde gösterdiğimizde **olasılık yoğunluğunu** (probability density function) gösteren bir eğri çizebiliriz. Bu eğrinin altındaki alan bire eşittir. x 'in belli bir aralıkta olma olasılığı, olasılık eğrisinde o aralığa denk gelen alana eşittir. Kolaylık olsun diye teorik dağılımlar için tablolar geliştirilmiştir. Ampirik dağılımımızın bir teorik dağılıma benzemesi halinde o teorik dağılımın olasılık tablolarına bakarak karar verebiliriz. Özellikle güven aralıkları ve hipotez testleri için bu tablolar çok yararlıdır. Bilgisayarla yapılan analizlerde değerler otomatik olarak hesaplandığından teorik dağılımların olasılık tablolarına gerek kalmamaktadır.

Eğrinin altındaki toplam alan = 1 (veya %100)

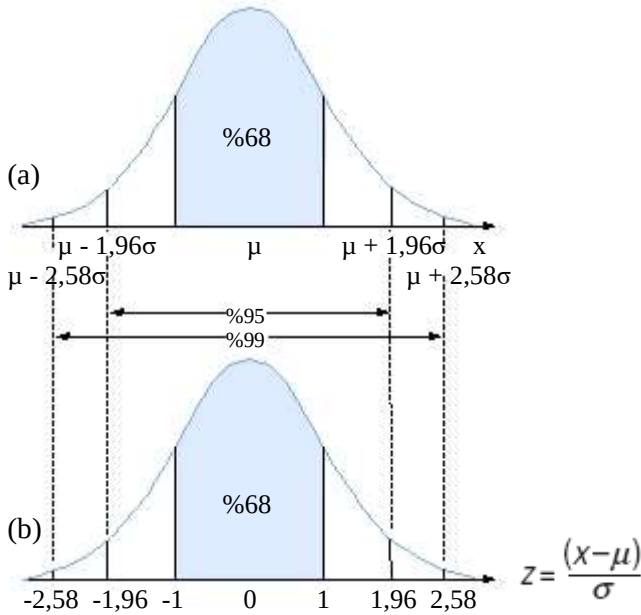


Şekil: x için olasılık yoğunluğu fonksiyonu (oyf)



Şekil: x değişkeninin olasılık yoğunluğu fonksiyonu.

- (a) μ ortalamasının etrafında simetrik dağılıyor. Varyans = σ^2
- (b) ortalamayı $\mu_2 > \mu_1$ olacak şekilde değiştirdiğimizdeki durum
- (c) varyansı $\sigma_1^2 < \sigma_2^2$ olacak şekilde değiştirdiğimizdeki durum



Şekil: (a) ortalaması μ ve varyansı σ^2 olan normal dağılıma sahip x ve (b) z 'ye ait standart normal dağılım için eğri altında kalan alanlar (toplam olasılıkların yüzdeleri).

Normal dağılım (Gauss dağılımı)

İstatistikteki en önemli dağılımlardan biri **Normal dağılım**dır. Olasılık yoğunluk fonksiyonu aşağıdaki gibidir:

- İki parametre tarafından tamamen tanımlanabilir: *ortalama* (μ) ve *varyans* (σ^2);
- Çan eğrisi şeklindedir (unimodal)
- Ortalamanın etrafında simetrik dağılır;
- Ortalama arttığında sağa, ortalama azaldığında ise sola kayar (varyanslar sabit kaldığında);
- Varyans arttığında yassılaşıp, varyans azaldığında ise sivrileşir (ortalama sabit);

Diğer özellikleri:

- Normal dağılımın ortalaması ve ortancası eşittir;
- Ortalaması μ ve varyansı σ^2 olan normal dağılıma sahip bir x rastgele değişkeni için olasılıklar aşağıdaki gibidir:

- $(\mu - \sigma)$ ile $(\mu + \sigma)$ arasında olma olasılığı 0,68
- $(\mu - 1,96\sigma)$ ile $(\mu + 1,96\sigma)$ arasında olma olasılığı 0,95
- $(\mu - 2,58\sigma)$ ile $(\mu + 2,58\sigma)$ arasında olma olasılığı 0,99

Yukarıdaki kurallar kullanılarak **referans aralıkları** (normal değerler) hesaplanabilir.

Standart Normal Dağılım

μ ve σ değerlerine bağlı olarak sonsuz sayıda farklı normal dağılım olabilir. Standart normal dağılım olasılıkların tablolar halinde verildiği özel bir normal dağılımdır. Özellikleri şöyledir:

- Standart normal dağılımın ortalaması sıfır ve varyansı bire eşittir.
- Bir x değişkeni normal dağılıyorsa ortalaması μ ve varyansı σ^2 olduğunda **Standart Normal Sapma** (Standardized Normal Deviate)

$$Z = \frac{(x - \mu)}{\sigma}$$
 formülüyle hesaplanır ve standart normal dağılıma sahip rastgele bir değişkendir.

Alıştırma 4

1. Bir hastanede 20 depresyon hastası, 20 psikotik hasta ve 60'da başka hasta vardır. Bu hastaneden rastgele bir hasta seçsek depresyonlu veya psikozlu olma olasılığı nedir?
2. Bir serviste 6 kızamık ve 4 astım hastası yatmaktadır. Bu servisten sırayla (rastgele) iki hasta seçsek ikisinin de astımlı olma olasılığı nedir?
3. 1000 kişinin katıldığı sınavda öğrencilerin aldıkları puanların aritmetik ortalaması 60, standart sapması 5'tir. Histogram grafiğinde verilerin normal dağıldığı görülmektedir. Geçme notu 50 olduğuna göre bu öğrencilerin kaç geçmiştir?
4. z formülünü kullanarak şu soruya cevap veriniz: sağlıklı bireylerin sistolik kan basınçları daha önceden ölçülmüş ve normal dağılıma uyduğunu saptanmıştır. Sağlıklı bireylerin kan basıncının ortalaması 130, standart sapması da 10 mmHg bulunmuş. Verilen bu toplum değerlerine göre kan basıncı 150 mmHg bulunan bir bireyi nasıl değerlendirirsiniz?
 - a. z testine göre bu kişinin tansiyonu toplum dağılımının neresine düşmektedir? %5 yanılma payıyla bu kişiye tansiyonunun yüksek olduğunu söyleyebilir miyiz?
 - b. Kişinin kan basıncı 145 olsaydı yaklaşımımız nasıl olurdu?

z	2-yönlü P-değeri
0.0	1.000
0.1	0.920
0.2	0.841
0.3	0.764
0.4	0.689
0.5	0.617
0.6	0.549
0.7	0.484
0.8	0.424
0.9	0.368
1.0	0.317
1.1	0.271
1.2	0.230
1.3	0.194
1.4	0.162
1.5	0.134
1.6	0.110
1.7	0.089
1.8	0.072
1.9	0.057
2.0	0.046
2.1	0.036
2.2	0.028
2.3	0.021
2.4	0.016
2.5	0.012
2.6	0.009
2.7	0.007
2.8	0.005
2.9	0.004
3.0	0.003
3.1	0.002
3.2	0.001
3.3	0.001
3.4	0.001
3.5	0.000

%5 yanılma payına göre standart normal dağılım olasılıkları (Z tablosu)

Çözüm için bakınız s. 470.

Bölüm 11: Teorik Dağılımlar: Diğer Dağılımlar

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun normal dağılım dışındaki diğer sık kullanılan teorik dağılımlar hakkında bilgi sahibi olmaları amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- t dağılımını açıklayabilmek
- t-dağılımını açıklayan parametreleri sayabilmek
- t dağılımının kullanım alanını açıklayabilmek
- Ki kare dağılımını açıklayabilmek
- Ki kare dağılımını açıklayan parametreleri sayabilmek
- Ki kare dağılımının kullanım alanını açıklayabilmek
- F dağılımını açıklayabilmek
- F dağılımını açıklayan parametreleri sayabilmek
- F dağılımının kullanım alanını açıklayabilmek
- Lognormal dağılımı açıklayabilmek
- Binomiyal dağılımını açıklayabilmek
- Binomiyal dağılımını açıklayan parametreleri sayabilmek
- Binomiyal dağılımının kullanım alanını açıklayabilmek
- Poisson dağılımını açıklayabilmek
- Poisson dağılımını açıklayan parametreleri sayabilmek
- Poisson dağılımının kullanım alanını açıklayabilmek

Teorik dağılımların teorisini anlamakta zorlanırsanız endişelenmeyiniz. İstatistik uzmanı olmayabilirsiniz. Bu nedenle daha çok bu bilgilerin pratik uygulamasıyla ilgilenmeniz normaldir. Bu bölümde de verilen bilgilerin olabildiğince pratiğe dönük olmasına dikkat edilmiştir.

Diğer sürekli teorik dağılımlar

Bu dağılımlar rastgele sürekli değişkenlere dayanır. Olasılık yoğunluğu grafiğinin altındaki toplam alan tüm olasılıkları ifade eder ve bire eşittir. Normal dağılım bir önceki konuda anlatılmıştı. Burada diğer sürekli teorik dağılımlardan bahsedilecek.

t-dağılımı

- William Sealy Gosset tarafından tanımlanmıştır. Kendisi ‘Student’ müstear ismiyle yazdığından dolayı dağılımın adı *Student t dağılımı* olarak bilinmektedir.

- t-dağılımını tanımlayan parametre **serbestlik derecesidir**. Dolayısıyla t-dağılımının formülünü ve serbestlik derecesini bilmemiz halinde olasılık yoğunluğu fonksiyonunu çizebiliriz. Serbestlik derecesi örneklem sayısıdır ve daha sonra tartışılacaktır.
- Şekli standart normal dağılıma benzer ama daha yayvandır ve kuyrukları daha uzundur. Serbestlik derecesi arttıkça şekli de normal dağılıma yaklaşır.
- Özellikle bir ya da iki ortalamayla ilgili güven aralıklarının hesaplanması veya hipotez testlerinin yapılmasında yararlıdır.

Ki Kare (χ^2) dağılımı

- Pozitif değerleri olan sağa eğimli bir dağılımdır.
- **Serbestlik derecesiyle** tanımlanır.
- Şekli serbestlik derecesine bağlıdır; serbestlik derecesi arttıkça daha simetrik ve normal dağılıma benzer hale gelir.
- Özellikle kategorik değişkenlerin analizinde yararlıdır.

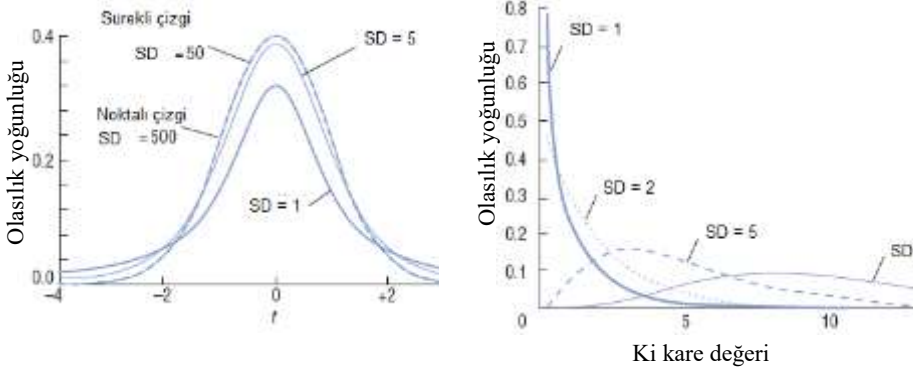
F-dağılımı

- Sağa eğimlidir.
- Bir oran tarafından belirlenir. Normal veriden hesaplanan iki tahmini varyansın birbirine oranı *F*-dağılımını verir.
- *F*-dağılımını tanımlayan iki parametre **pay ve paydasının serbestlik dereceleri**dir.
- *F*-dağılımı özellikle iki varyansın karşılaştırılmasında ve varyans analizini (ANOVA) kullanarak ikiden fazla ortalamayı karşılaştırmada yararlıdır.

Lognormal dağılım

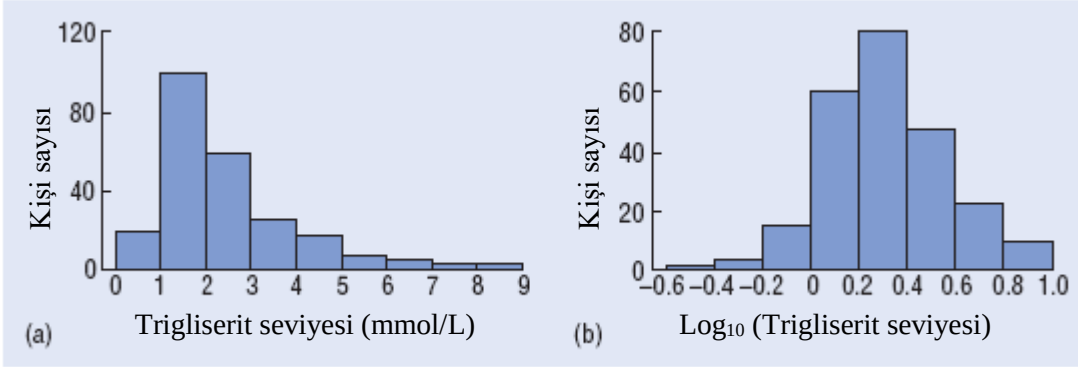
- 10 veya *e* tabanına göre logaritmasının normal dağılım gösterdiği rastgele bir değişkenin olasılık dağılımıdır.
- İleri derecede sağa eğimlidir.
- Eğer sağa doğru eğimli bir veriye logaritmik dönüşüm uyguladığımızda ortaya çıkan veri normal dağılıma yaklaşıyorsa verimizin lognormal dağılıma sahip olduğunu söyleriz.
- Tıptaki birçok değişken lognormal dağılıma sahiptir (Hastalarda yüksek değerlere daha çok rastlandığından). Bu tür verilere logaritmik dönüşüm uyguladıktan sonra normal dağılımla ilgili bilgilerimizi kullanarak çıkarımlar yapabiliriz.

- Veri setimizin lognormal dağılıma sahip olması halinde veriyi özetlemek için geometrik ortalamayı kullanmalıyız.

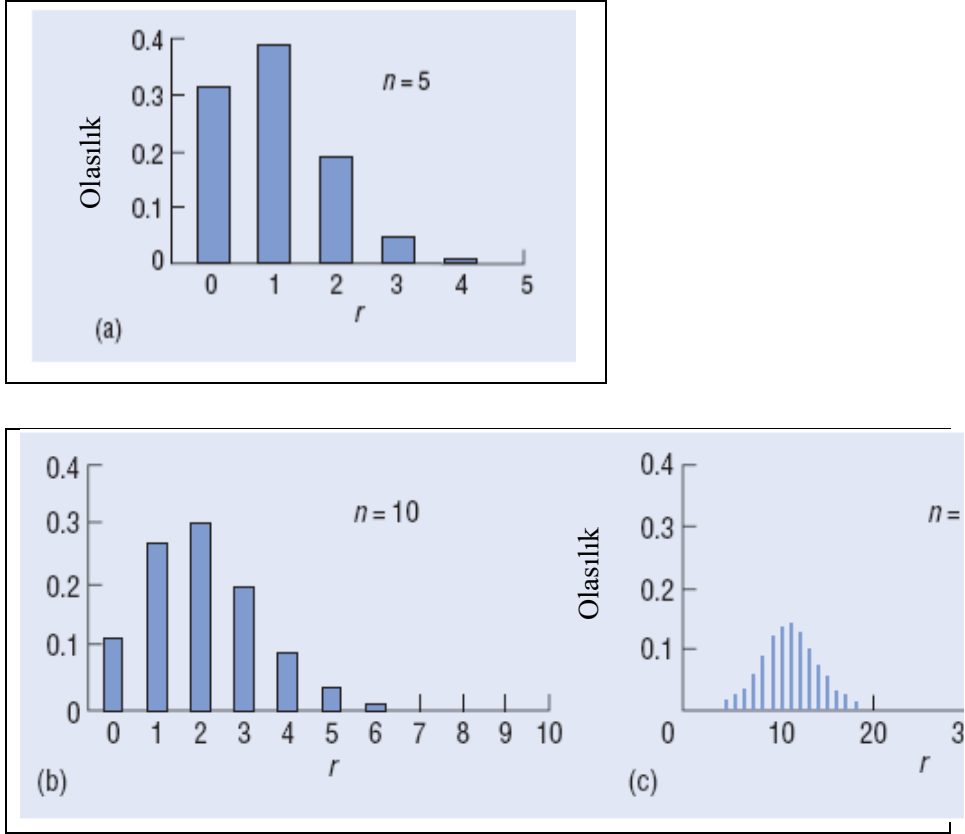


Şekil: Çeşitli serbestlik dereceleri (SD) olan t -dağılımları.

Şekil: Çeşitli serbestlik dereceleri (SD) olan Ki kare dağılımları.



Şekil: (a) kalp hastalığı gelişen 232 erkeğin trigliserit seviyelerinin lognormal dağılımı. (b) $\log_{10}$ (trigliserit) değişkeninin yaklaşık olarak normal dağılımı.



Şekil: Başarı olasılığı r 'yi gösteren binomiyal dağılımlar. Başarı olasılığı $\pi = 0,20$; örneklem sayıları (a) $n = 5$, (b) $n = 10$, (c) $n = 50$. Görüldüğü üzere, örneklem sayısı arttıkça dağılımın şekli normale yaklaşmaktadır.

Tam sayılı teorik dağılımlar

Olasılık dağılımını belirleyen rastgele değişken tam sayılıdır. Birbirini dışlayan (mutually exclusive) olayların olasılık toplamı bire eşittir.

Binomiyal dağılım

- Bir durum için sadece başarı ve başarısızlık şeklinde iki sonuç olabileceğini varsayalım. Bir kadının in vitro fertilizasyon (IVF) sonrasında gebe kalması (başarı) veya kalmaması (başarısızlık) örnek olarak verilebilir. 100 adet gebe kalma olasılığı aynı olan kadına IVF yapılsın. Binomiyal rastgele değişkenimiz başarılı gebeliklerin sayısıdır.

- Binomiyal dağılımı belirleyen iki parametre örneklemdeki birey sayısı n (veya olay bir para atma deneyi ise paranın kaç kez atıldığı) ve her bir birey için olayın gerçek meydana gelme olasılığı π 'dir.
- **Ortalaması** (deneyi n kez tekrarladığımızda veya n bireyde uygulama yaptığımızda olayın meydana gelme durumu) $n\pi$ 'dir. **Varyansı** ise $n\pi(1-\pi)$ 'dir.
- n küçük olduğunda, dağılım $\pi < 0,5$ olması halinde sağa, $\pi > 0,5$ olması halinde ise sola eğimlidir. Örneklem sayısı arttıkça dağılım daha simetrik olur ve hem $n\pi$ hem de $n(1-\pi)$ 5'ten büyük olduğunda normal dağılıma yaklaşır.
- **Orantılarla** ilgili çıkarımlar yapmak istediğimizde Binomiyal dağılımın özelliklerini kullanabiliriz.

Poisson dağılımı

- Poisson rastgele değişkeni, zamandan ve mekandan bağımsız olarak μ hızında gerçekleşen olayların **sayısıdır**. Bir hastaneye günde yatırılan hasta sayısı tipik olarak Poisson dağılımı gösterir. Poisson dağılımı bilgilerimizi kullanarak bir hastaneye belli bir günde yatırılacak hastaların sayısını hesaplayabiliriz.
- Poisson dağılımını tanımlayan parametre **ortalama**dır (yani ortalama hız μ).
- Poisson dağılımında **ortalama** ve **varyans** eşittir.
- Ortalaması düşük olduğunda sağa eğimli bir dağılımdır ama ortalama arttıkça simetrikleşir ve normal dağılıma yaklaşır.

Bölüm 12: Veri Dönüştürme

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun uygun istatistik testleri yapabilmek için veri dönüştürme yöntemleri hakkında bilgi sahibi olması ve SPSS'i kullanarak veri dönüşümü yapabilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Verilere neden bazen dönüşüm uygulamak gerektiğini açıklayabilmeli.
- Veri dönüştürmenin veri üzerinde oluşturacağı etkileri sayabilmeli
- Veri dönüşümünün nasıl yapılacağını açıklayabilmeli
- Tipik veri dönüşümlerini ve bunların veriye nasıl etki edeceğini açıklayabilmeli
 - Logaritmik dönüşüm
 - Karekök
 - $1/y$
 - Kare
 - Logit dönüşüm
- SPSS ile tipik veri dönüşümlerini yapabilmeli

Buraya kadar verileri elde etmede, bilgisayara girmede, bilgisayara girdikten sonra hata ayıklama ve çeşitli yöntemlerle verilerimizi özetleme konularını işledik. Artık verilerimizin analizine geçtiğimizde ve önemlilik testlerini uygulamak istediğimizde uygulayacağımız testlerin bazı varsayımları karşımıza çıkacak. Verilerimizin bu varsayımları karşılamaması halinde söz konusu istatistik analizi uygulayamayacağımızdan, verimize dönüşüm uygulayarak testin varsayımlarını karşılamak bir çözüm olabilir.

Veriye dönüşüm uygulamanın sebepleri ve dönüşüm uygulandıktan sonra oluşan durumlar şunlardır:

- Verimiz normal dağılıma sahip değildir. Oysa birçok istatistik testte **dağılım varsayımı** olarak normal dağılım bir gerekliliktir.
- Gruplarımız arasında verilerin yaygınlığı (spread) aşırı farklılık gösterebilir. Oysa bağımsız gruplarda t-testi gibi bazı testlerde **parametre varsayımı** olarak varyansların eşit olması gerekmektedir.
- İki değişken lineer olarak ilişkili değildir. Oysa regresyon analizlerinde olduğu gibi **linearite** gerekli bir varsayımdır.

Veri dönüşümü nasıl yapılır?

Verisetimizdeki herhangi bir değişkene dönüşüm uygulamak istersek o değişkenin tüm verilerine aynı matematiksel işlemi uygulamamız gerekir. Uygulayacağımız matematiksel işlem önemli değildir. Örn.

Yaş değişkenimizle işlem yapmak istiyoruz ama verilerimizi incelediğimizde normal dağılıma uymuyor. Verisetimizde 100 kişinin “Yaş” değeri bulunsun. Her bir yaş değişkenine aynı işlemi uygulatalyız (örn. Karesini almak). Sonuçta dönüşüm uygulanmış yeni bir değişken (“YaşDön” gibi) ortaya çıkacaktır.

Dönüşüm uygulanmış değişkeni klinik olarak yorumlamak zor olacağından (yaşın karesini klinisyenin yorumlaması zordur) analizimizi yaptıktan sonra sonuçlarımızı rapor ederken geri dönüştürmemiz gerekir (karesini almışsak bu sefer karekökünü alırız). Geri dönüştürmede bazı verilerde sorun olabileceğine dikkat etmeliyiz (eksi değerlere sahip bir değişkene “kare dönüşümü” uyguladığımızda sonuçların karekökünü almak yanıltıcı olacaktır).

Tipik dönüşümler

Yukarıda belirtildiği gibi, verimize herhangi bir işlemi uygulayarak dönüşüm yapabiliriz. Ancak burada etkileri ve kullanım alanları bilinen başlıca beş tipik dönüşümden bahsedeceğiz.

Logaritmik dönüşüm, $z = \log y$

Logaritmik dönüşümü 10 tabanına ($\log_{10}$ veya log diye ifade edilir) veya e tabanına ($\log_e$ veya ln olarak ifade edilir) göre yapabiliriz. Negatif sayıların ve sıfırın logaritmasını alamayacağımıza dikkat ediniz. Logaritmanın geri dönüşümüne ise antilog denir. Örn. 100’ün logaritmasını alacak olursak;

$$\log_{10}(100) = 2$$

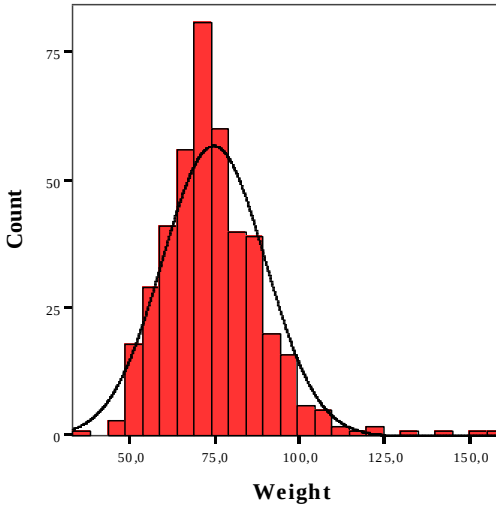
$$\text{Antilog}(2) = 10^2 = 100$$

- y sağa eğimli olduğunda “ $z = \log y$ ” genelde yaklaşık olarak **normal dağılıma** sahip olur. Bu durumda y’nin lognormal dağılıma sahip olduğundan bahsederez.
- y ile x değişkeni arasında üstel bir ilişki olduğunda yatay eksen x, dikey eksen y olacak şekilde grafik çizildiğinde yukarıya doğru eğimli bir grafik oluşuyorsa “ $z = \log y$ ” ile x arasında **lineer** bir ilişki olur.
- Farklı gruplarda sürekli bir değişken olan y’yi (örn. boy) ölçtüğümüzü varsayalım. y büyük olan gruplarda varyans da büyük olacaktır. Özellikle varyans katsayısının (standart sapma / ortalama) gruplar arasında eşit olması halinde “ $z = \log y$ ” değişkeni için **varyanslar eşit** olacaktır.

Yorumlanmasının kolay olması ve verilerin genelde sağa eğimli olması nedeniyle tıpta log dönüşümü sıklıkla kullanılır.

Örn.:

İleride göreceğimiz gibi, normal dağılıma uygunluk birçok parametrik hipotez testi için önemli bir varsayımdır. Diyabet.sav (<https://drive.google.com/open?id=0B0ieO-Fy3z-5a3cydj1PQ0YtYk0>) veri setinde “Weight” değişkenini incelediğimizde histogram grafiğinde çan eğrisinin kuyruğunun sağa doğru olduğunu (sağa eğimli) görürüz:



Weight değişkeninin normal dağılmadığını Kolmogorov Smirnov veya Skewness analizleri ile de gösterebiliriz:

SPSS ile diyabet.sav verisetini açınız. Analyze > Descriptive statistics > Frequencies [“Weight” değişkenini Variables kısmına geçirin, Statistics butonunu tıklayıp Skewness kutucuğunu seçiniz] > Continue > OK. Aşağıdaki çıktı oluşacaktır:

Statistics

Weight

N	Valid	424
	Missing	6
Skewness		1,329
Std. Error of Skewness		,119

Skewness değeri (1,329) Standart sapmasının (0,119) iki katından büyük ve pozitif olduğundan verimizin sağa doğru eğimli olduğunu söyleyebiliriz.

Analyze > Nonparametric Tests > 1- Sample K-S [“Weight” değişkenini Variables kısmına geçirin] > Continue > OK. Aşağıdaki çıktı oluşacaktır:

One-Sample Kolmogorov-Smirnov Test

		Weight
N		424
Normal Parameters(a,b)	Mean	74,266
	Std. Deviation	15,1381
Most Extreme Differences	Absolute	,094
	Positive	,094
	Negative	-,045
Kolmogorov-Smirnov Z		1,926
Asymp. Sig. (2-tailed)		,001

a Test distribution is Normal.

b Calculated from data.

Asymp. Sig. (2-tailed) (0,001) 0,05’ten küçük olduğundan verimizin normal dağılmadığı sonucuna varırız.

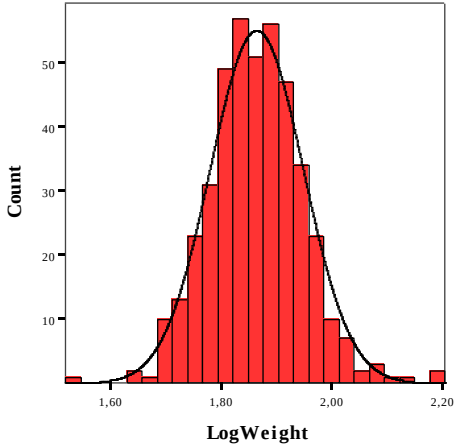
Şimdi “weight” değişkeninin logaritmasını alarak yeni bir değişken oluşturalım:

Transform>Compute variable>[“Target Variable” alanına “LogWeight”, “Numeric Expression” alanına ise “LG10(weight)” yazalım]>OK

SPSS veri setimizde “LogWeight” adında yeni bir değişken oluşacaktır. Şimdi bu değişkenin histogram grafiğine bakalım:

Graphs>Interactive>Histogram [X eksenine “LogWeight” değişkenini sürükleyelim. Üstteki “Histogram” sekmesini tıklayıp “Normal curve” kutucuğunu işaretleyelim]>OK

Aşağıdaki grafiği elde ederiz:



Görüldüğü gibi çan eğrisi simetrik hale gelmiştir.

Yeni değişkenimizin Skewness değerine baktığımızda:

Statistics

LogWeight

N	Valid	424
	Missing	6
Skewness		,248
Std. Error of Skewness		,119

Standart sapmasının iki katına yakındır.

Kolmogorov Smirnov ile bakıldığında ise:

One-Sample Kolmogorov-Smirnov Test

		lgweight
N		424
Normal Parameters(a,b)	Mean	1,86
	Std. Deviation	,084
Most Extreme Differences	Absolute	,053
	Positive	,053
	Negative	-,038
Kolmogorov-Smirnov Z		1,091
Asymp. Sig. (2-tailed)		,185

a Test distribution is Normal.

b Calculated from data.

p değerinin (0,185) 0,05'ten büyük olduğunu, yani normal dağıldığını görüyoruz.

Bu arada, istatistik hesaplarımızda ve makalelerimizde normal dağılımı kontrol etmemiz ve sonuçlarını bildirmemiz gerektiğini, ancak parametrik testlerin çoğunun da (ileride görüleceği gibi) normallikten hafif sapmalara karşı toleranslı olduğunu belirtmeliyiz.

Başka bir örnek de varyansların homojenliği için yapalım:

Bağımsız iki grubu nümerik bir değişken açısından karşılaştırmak için bağımsız örneklemelerde t testi (Student t test) yapmalıyız. Student t testi, hem verilerimizin normal dağılmasını, hem de iki grup arasında varyansların homojen olmasını gerektirir.

Diyabet.sav verisetinde çocuk sayısını (nümerik değişken) cinsiyet açısından (iki bağımsız grup) karşılaştırmak istediğimizde çocuk sayısı değişkeni açısından iki grubun varyanslarının homojen olmadığını görüyoruz (Levene $p=0,041$):

Independent Samples Test										
		Levene's Test for Equality of Variances				t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
Number of children	Equal variances assumed	4,199	,041	2,588	339	,010	1,077	,416	,259	1,896
	Equal variances not assumed			2,547	297	,011	1,077	,423	,245	1,909

Logaritmik dönüşüm yaparak "lognumchild" isimli yeni bir değişken oluşturduk. Bu değişken için bağımsız örneklemelerde t testi yaptığımızda varyansların homojen olduğunu görüyoruz (Levene $p=0,877$):

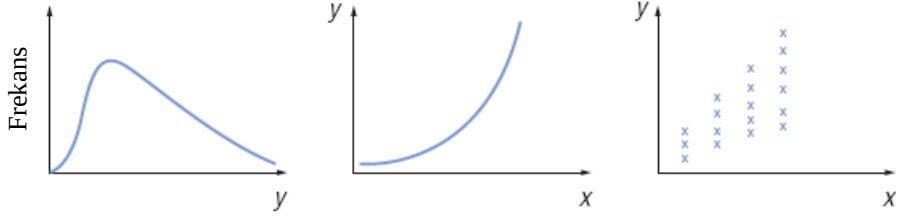
Independent Samples Test										
		Levene's Test for Equality of Variances				t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
lognumchild	Equal variances assumed	,024	,877	1,807	319	,072	,05105	,02825	-,005	,1066
	Equal variances not assumed			1,805	315	,072	,05105	,02828	-,005	,1067

Not: Varyansların homojen olmadığı durumda Student t testinin sonucunun gruplar arasında fark vardır şeklinde iken ($p=0,010$), logaritmik dönüşümden sonra aradaki farkın anlamsız çıktığına ($p=0,072$) dikkat ediniz.

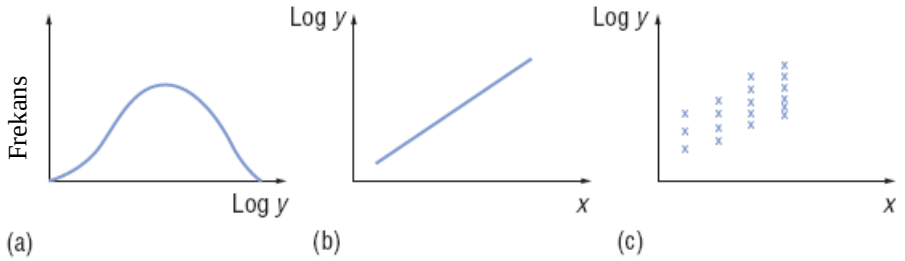
Karekök dönüşümü $z = \sqrt{y}$

Bu dönüşümün özellikleri log dönüşüme benzer. Ancak geriye dönüştürme sırasında verinin yorumlanmasında sorunlar olabilir. **Normalize** ve **linearize** edici özelliklerinin yanında y arttıkça varyansın artması halinde, (varyans/aritmetik ortalama sabit olduğunda) **varyansı stabilize edici** etkisi de vardır. Karekökü dönüşümü genelde Poisson türü değişkenler için kullanılır. Ayrıca, eksi sayıların karekökünün alınamayacağını da hatırlatmak gerekir.

Dönüşümden önce



Dönüşümden sonra

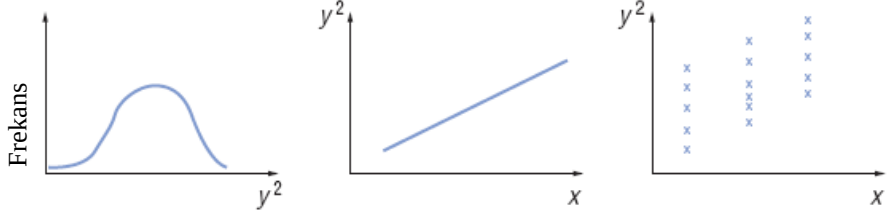


Şekil: Logaritmik dönüşümün etkileri: (a) normalize edici, (b) linearize edici, (c) varyans stabilize edici.

Dönüşümden önce



Dönüşümden sonra



(a)

(b)

(c)

Şekil: Kare dönüşümün etkileri: (a) normalize edici, (b) linearize edici, (c) varyans stabilize edici.

Resiprok dönüşüm, $z=1/y$

Sağkalım (survival) analizlerinde özel teknikler kullanmadığımız müddetçe resiprok dönüşüm uygulanır. Resiprok dönüşümün log dönüşüme benzer etkileri vardır. **Normalize** edici ve **linearize** edici özelliklerinin yanında **varyansı stabilize etmede** (y arttıkça varyansın çok hızlı artması halinde yani varyans/ortalama⁴ sabit olduğunda) log dönüşüme göre daha etkilidir. Sıfırın resiprokunun alınamayacağı hatırlatılmalıdır.

Kare dönüşümü, $z=y^2$

Kare dönüşüm log dönüşümün tersini yapar.

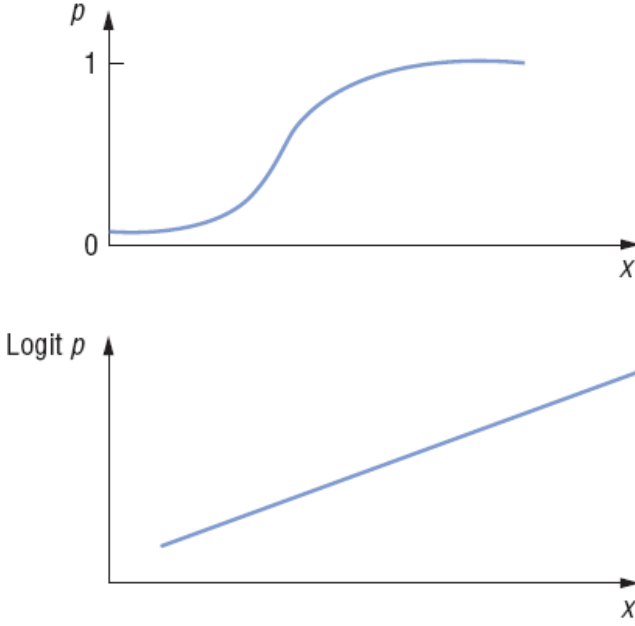
- y sola eğimli olduğunda $z=y^2$ genelde **normal** dağılır.
- x ve y arasındaki ilişkinin grafiği çizildiğinde aşağıya doğru bir eğri oluştuğunda $z=y^2$ ve x arasındaki ilişki yaklaşık olarak **lineer** olur.
- Sürekli bir değişken olan y'nin varyansı y arttıkça azalma meylindeyse kare dönüşümü $z=y^2$ varyansı stabilize eder.

Logit (logistic) dönüşüm

$$z = \ln \frac{p}{1-p}$$

Oranlarla çalıştığımızda p için bu dönüşümü uygularız. p 'nin sıfır veya bir olması halinde logit dönüşüm uygulayamayız.

Bu dönüşüm sigmoid bir eğriyi **linearize** eder.



Şekil: Logit dönüşümün sigmoid bir eğri üzerindeki etkileri.

Bölüm 13: Örneklem ve Örneklem Dağılımları

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun örneklem ve örneklem dağılımları hakkında bilgi sahibi olması amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Neden örneklem yaptığımızı açıklayabilmek
- Örneklem çeşitlerini açıklayabilmek
- SEM ve SEP formüllerini yazabilmek
- SD ve SEM kullanım alanlarını açıklayabilmek

İstatistik yaparken genelde bir toplumun tamamı hakkında bilgi toplamak ve yorum yapmak isteriz. Ancak, toplumun tamamından veri elde etmek gerek zaman, gerekse ekonomik açıdan genelde mümkün olmaz. Bu nedenle toplumu temsil edecek bir örneklem den veri toplar ve o verileri kullanarak toplum hakkında çıkarımlar yaparız.

Toplumdan bir örneklem aldığımızda örneklemimizin toplumu tamamen temsil edemeyeceğini tahmin edebiliriz. Toplumun sadece bir kısmını inceleyerek bir örneklem hatası oluşturmuş oluyoruz. Bu bölümde teorik dağılımları kullanarak bu hatayı hesaplamayı öğreneceğiz.

Örneklem in temsil edici olması

İdeal olanı **rastgele** bir örneklem yapılmasıdır. Toplumdaki bireylerin listesi çıkarılır ve bu listeden herkesin eşit katılım olasılığı olacak şekilde bir örneklem seçimi yapılır. Bazen toplumun listesini çıkarmak (örn. ilimizdeki tüm bireylerin TC kimlik numaralarını elde etmek) veya listeden rastgele yöntemle seçilen bireylere ulaşmak (örn. TC kimlik numaralarını bir şekilde bulsak ve bunların arasından rastgele seçsek bile her biri farklı adreslerdeki onlarca veya yüzlerce bireyi tek tek bulmak) zor olabilir. Bu durumda bize **rahat** gelen bir örneklem seçimi yaparız (convenience sample). Bir hastalık durumunu incelemek istediğimizde tüm hastanelerden örneklem almak yerine bir hastaneyi inceleyip ülke geneli hakkında tahminde bulunmak da buna örnek verilebilir.

Kota örneklem, küme örneklem, tabakalı örneklem, sistematik örneklem, kartopu örneklem ve karekök örneklem gibi başka örneklem yöntemleri de vardır.

Biz anlatacağımız konularda genelde örneklemimizin rastgele seçildiğini varsayacağız. Ancak, örneklemimizin **temsil edici** olması halinde örneklem yöntemlerinin hepsi kabul edilebilirdir.

Örneklemden elde ettiğimiz verilerle toplumla ilgili toplum ortalaması μ ve toplum standart sapması σ gibi parametreleri tahmin etmeye çalışırız.

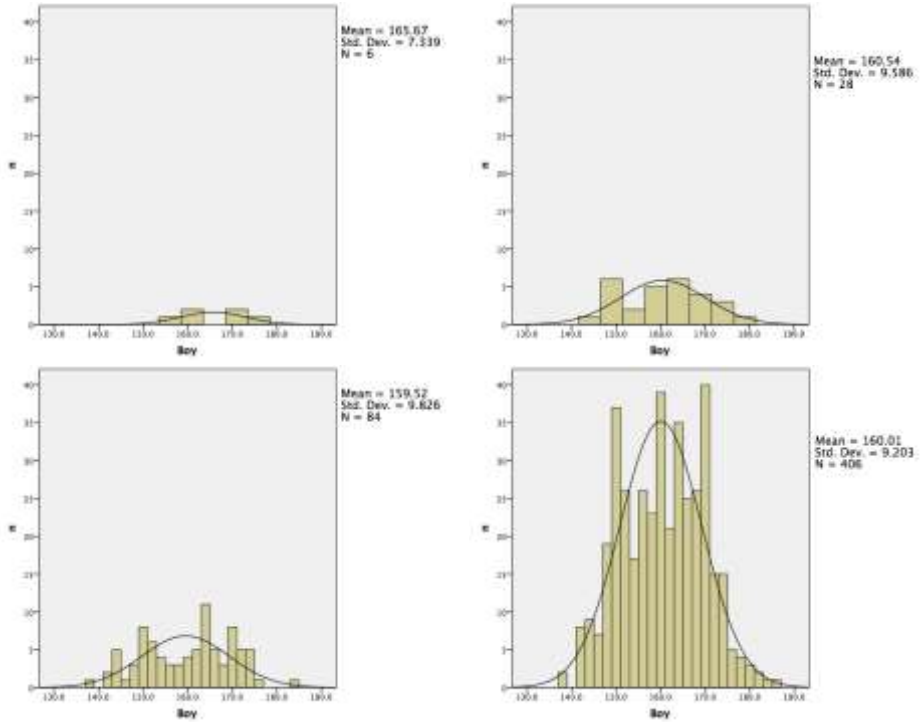
Örneklem varyasyonu

Aynı toplumdan aynı büyüklükte örneklem alsak bile μ ve σ gibi parametrelerde farklılıklar olacaktır. Halbuki biz toplumla ilgili tahminlerimizin gerçek değere yakın olmasını isteriz. Eğer bu farklılıkları rakama dökebilirsek, tahminimizin hassasiyeti konusunda bilgi sahibi olabiliriz ve böylece örnekleme hatamızın düzeyi hakkında fikrimiz olur.

Gerçekte toplumdan tek bir örneklem almamıza rağmen yine de teorik dağılımlarla ilgili bilgilerimizi kullanarak toplum geneli hakkında çıkarımlar yaparız.

Ortalamanın örneklem dağılımı

Toplum ortalamasını tahmin etmeye çalışıyoruz. Toplumdan n sayıda örneklem alıp bunların ortalamasını hesaplayabiliriz. Bu ortalamaların histogram grafiğini çıkarsak ortalamaların dağılımını görebiliriz. Buna **ortalamanın örneklem dağılımı** denir.



Şekil: Aynı toplumdaki çeşitli örneklem sayılarında dağılımların değişimi.

Bu dağılıma bakarak şunları söyleyebiliriz:

- Örneklem sayısı yeterince büyük olursa, asıl verinin nasıl dağıldığına bakılmaksızın ortalamaların dağılımı normal dağılıma yakın olur (*Central Limit Theorem*).
- Örneklem sayısı küçükse, asıl verinin normal dağılması halinde ortalamalar normal dağılıma yakın olur.
- Bu ortalamaların ortalaması **gerçek** (unbiased) toplum ortalamasını verir.
- Bu dağılımın variabilitesi alınan ortalamaların standart sapmasıyla hesaplanır. Buna **ortalamanın standart hatası** (SEM) denir. Toplumun standart sapmasını (σ) bilmemiz halinde ortalamanın standart hatası **SEM = $\sigma / \sqrt{n}$** formülüyle hesaplanır.

Genelde olduğu gibi, toplumdan tek bir örneklem almışsak, toplum ortalamasının en iyi tahmini örneklemimizin ortalaması olacaktır. Bu durumda toplumun standart sapmasını da bilmediğimizden ortalamanın standart hatasını **SEM = $s / \sqrt{n}$** formülüyle hesaplarız.

SEM, tahminimizin hassasiyeti konusunda bilgi verir. Tahminimizin ne kadar hassas olduğunu gösterir.

- Büyük bir standart hata, tahminimizin hassas olmadığını gösterir.
- Standart hatanın küçük olması ise tahminimizin hassas olduğunu gösterir.

Örneklem sayısının artırılması veya verilerin variabilitesinin daha az olması halinde standart hatayı küçültmek, yani daha hassas bir tahmin yapmak mümkündür.

Standart sapma mı standart hata mı?

Bu iki parametre birbirine benzer görünse de farklı amaçlarla kullanılırlar. Standart sapma verilerdeki varyasyonu (ortalamadan sapmayı) gösterir ve bu bilgiyi göstermek istediğimizde kullanılmalıdır. Buna karşın standart hata, örneklem ortalamasının hassasiyetini gösterir ve ölçümümüzün hassasiyetini vurgulamak istediğimizde kullanılmalıdır.

Orantının örneklem dağılımı

Araştırmamızda toplumdaki bir orantıyı incelediğimizi düşünelim. Toplumdan n sayıda örneklem almamız ve orantımızın p olması halinde toplum ortalaması π için en iyi tahmin

$$p = r / n$$

şeklinde hesaplanabilir (r , toplumda araştırdığımız özelliğe sahip kişilerin sayısı). Eğer toplumdan tekrarlayan n sayıda örneklem alsak ve orantılarımızın histogram grafiklerini çizsek, sonuçta ortaya çıkan **orantının örneklem dağılımının** ortalaması π olup normal dağılıma yakın olacaktır. Bu orantıların standart sapmasına **orantının standart hatası** $[SE(p)]$ denir. Tek bir örneklem aldığımızda şöyle hesaplanır:

$$SE(p) = \sqrt{\frac{p(1-p)}{n}}$$

Bu, tahmin ettiğimiz π değerinin hassasiyetini gösterir. **Küçük** bir standart hata daha hassas bir ölçüme işaret eder.

Ne kadar küçük “küçük”tür?

Standart hatanın küçük olması güven aralığının (Bakınız sayfa 127) dar olmasını sağlayacaktır. Bu ise ölçümün hassas olduğu anlamına gelecektir.

“Küçük” veya “dar” ifadeleri yoruma açıktır. Örneğin “Anticraving Effect of Baclofen in Alcohol-Dependent Patients” isimli makalelerinde (<http://www.ncbi.nlm.nih.gov/pubmed/26207768>) Imbert ve arkadaşları katılımcıların cevap verme oranlarını %38, %95 güven aralığını ise [7-76] arasında bildirmişlerdir. Dolayısıyla bu güven aralığının geniş olduğunu, ölçümün hassasiyetinin düşük olduğunu söylemek yanlış olmaz.

Post ve arkadaşlarının (<http://www.ncbi.nlm.nih.gov/pubmed/25010580>) makalelerinde ise hastaların %19,6’sının kilo verdikleri belirtilmiş ve %95 güven aralığı [18,0-21,2] arasında verilmiştir. Bu örnekte ise güven aralığının dar olduğunu, dolayısıyla ölçümün hassasiyetinin yüksek olduğunu söyleyebiliriz.

Alıştırma 5

<http://bmj.bmjournals.com/cgi/content/full/315/7119/1342> bağlantısından “Does malnutrition in utero determine diabetes and coronary heart disease in adulthood? Results from the Leningrad siege study, a cross sectional study” isimli makaleyi indiriniz.

Makalenin özellikle “Subjects and methods” ve “Results” bölümünü okuyunuz.

1. Bu araştırmada kullanılan verileri çeşitlerine göre sınıflandırınız?
2. Bu araştırmada sistolik kan basıncı ve vücut kitle endeksi arasındaki ilişki ne tür bir grafikte gösterilmiştir?
3. Araştırmanın sonuçlarından merkezi eğilim ölçütleri ve yaygınlık ölçütlerine örnekler veriniz.
4. Araştırmada belirtilen “adjusted for sex” ifadesinden ne anlıyorsunuz?
5. Araştırmada hangi veri dönüştürmeleri uygulanmıştır? Neden?
6. Yetkili araştırmacı siz olsaydınız bu araştırmanın yönteminde herhangi bir farklılık yapar mıydınız?
7. Başka bir makalede (<http://bja.oxfordjournals.org/cgi/content/full/90/4/514>) anestezi dergilerinde yayınlanan makalelerde SD (standart sapma) ve SEM (standart hata) nın yanlış kullanımları incelenmiş. Makaledeki Tablo 1 aşağıdadır. Tablodan göreceğiniz gibi araştırmalarda sıklıkla SD yerine SEM kullanılmaktadır. Sizce bu yanlış kullanım okuyucuda ne gibi yanlış anlamalara sebep olur? (Dilerseniz ilgili makalenin tartışma bölümünü okuyun)

Table 1 Frequency of use of standard error of the mean (SEM) and standard deviation (SD) in four anaesthesia journals, listed in order of decreasing percentage. Data are numbers of articles (%). *Some of these articles used both the SD and the SEM to describe the study sample

	Incorrect use of SEM/total*	Laboratory studies using SEM incorrectly	Correct use of SD
<i>Anaesthesia & Analgesia</i>	112 (27.7)	66 (59)	293 (72.3)
<i>British Journal of Anaesthesia</i>	31 (22.6)	15 (47)	106 (77.4)
<i>Anesthesiology</i>	48 (18.7)	43 (90)	209 (81.3)
<i>European Journal of Anaesthesiology</i>	7 (11.5)	3 (43)	54 (88.5)

Çözüm için bakınız s. 470

Bölüm 14: Güven Aralıkları

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun güven aralıkları hakkında bilgi sahibi olması amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Güven aralığını tanımlayabilmek
- Ortalamanın güven aralığını hesaplayabilmek
- Oranlar için güven aralığını hesaplayabilmek
- Güven aralığının teorik dağılımlarla ilişkisini tanımlayabilmek
- Farkların standart hatasını hesaplayabilmek
$$SEM = \sqrt{[(s_1^2/n) + (s_2^2/n)]}$$
- Relative Deviate'i (relative sapma) hesaplayabilmek
RS = (Ortalama 1 – Ortalama 2) / farkların standart hatası
- Güven aralığını yorumlayabilmek
 - Gruplar arası farkı incelerken güven aralığının 0'ı içermesi
 - Risk değerlendirmesinde (Odds ratio) güven aralığının 1'i içermesi
- Serbestlik derecesini açıklayabilmek
- SPSS kullanarak ortalamaların ve ortalamalar arasındaki farkın güven aralığını hesaplayabilmek.

Toplumdan aldığımız örnekleme değişkenimizin merkezi eğilim ve yaygınlık ölçütleri gibi özelliklerini hesapladık. Daha sonra tahminlerimizin hassasiyetini göstermek için standart hatayı hesaplamayı öğrendik. Çoğu kez standart hatayı yorumlamak zor gelir (örn. Elde ettiğimiz değer küçük mü? Yeterli hassasiyeti göstermesi için daha mı küçük olmalıydı?). Bu nedenle standart hatayı **aralık tahmini** (interval estimate) olarak kullandığımız yeni bir hesaplama yapabiliriz. Burada teorik dağılım bilgilerimizi kullanarak toplum parametresinin **güven aralığını** hesaplamaya çalışırız. Güven aralığı, ortalamanın iki tarafında standart hatanın bir katsayıyla çarpılması şeklinde bulunur.

Ortalamanın güven aralığı

Normal dağılım bilgilerimizi kullanarak

Daha önce örneklem sayımız arttıkça dağılımımızın normal dağılıma benzeyeceğinden bahsetmiştik. Bu bilgiyi kullanarak verilerimizin %95'inin toplum ortalamasından 1,96 standart sapma uzaklıkta olduğunu görmüştük. Ortalamamızın %95 olasılıkla hangi aralıkta olduğunu hesaplamak için ortalamanın standart hatası formülünden yararlanılır:

Ortalamanın %95 **güven aralığı** (GA), (ortalama – 1,96 x SEM) ile (ortalama + 1,96 x SEM) arasında olacaktır.

Deneyimizi tekrar edip değişkenimizin tekrar tekrar ortalamasını alsak, ölçümlerin %95’inde aritmetik ortalama yukarıda verilen aralıkta olacaktır.

%95 yerine %99 güven aralığını hesaplamak istersek, tabii ki katsayı olarak 1,96 yerine 2,58 almamız gerekecektir. Normal dağılım dışında örneğin t dağılımı söz konusu ise bu durumda da ilgili t değeri kullanılmalıdır. Güven aralığı İngilizce makalelerde Confidence Interval (CI) olarak geçmektedir.

Aslında toplum ortalamasının sabit bir değer olduğunu ve kendisine bir aralığın atfedilemeyeceğini biliyoruz. Ancak, bizler örneklemelerde çalıştığımızdan ve anlaşılması da kolay olduğundan güven aralığı hesaplaması sıkça kullanılmaktadır ve bilimsel yayınlarda da p değerinin yanında verilmektedir.

t-dağılımı bilgilerimizi kullanarak

Normal dağılım bilgilerimizi varyansın (σ^2) toplumdaki değerini bilmemiz halinde kullanmalıyız. Diğer taraftan, örneklem sayısı az olduğunda, verilerin ancak toplumda normal dağılması halinde ampirik dağılımımızın normal dağılıma benzeyeceği bilgisini de hatırlatmalıyız.

Verilerimizin normal dağılmaması ve/veya toplum varyansını bilmememiz halinde eğer verilerimiz t-dağılımına uygunsa örneklemimizin standart sapmasını kullanarak hesap yapabiliriz. Bu durumda:

Ortalamanın %95 güven aralığı (ortalama – $t_{0,05}$ x SEM) ile (ortalama + $t_{0,05}$ x SEM) arasında olacaktır.

$t_{0,05}$ değerini örneklemimizin serbestlik derecesine (SD) göre ilgili tablolardan bakabiliriz. Örn. 20 kişilik bir örneklem için $t_{0,05}$ 2,093 olur (SD=19).

t-tablolarından görüleceği gibi, örneklem sayısı arttıkça normal dağılım katsayısıyla t-tablosu değerleri arasındaki fark ihmal edilecek düzeye inmektedir. Bu nedenle, örneklem sayısı yeterince büyük olsa bile güven aralıklarının hesaplanmasında t-dağılımı bilgileri kullanılır.

SPSS ile ortalamanın güven aralığının hesaplanması

Diyabet.sav (<https://drive.google.com/open?id=0B0ieO-Fy3z-5a3cydjIPQ0YtYk0>) verisetinde yaş değişkeni için ortalamanın %95 güven aralığını hesaplayalım:

Analyze > Descriptive Statistics > Explore [“Dependent List” kutusuna “Age” değişkenini koyalım. “Display” kısmında “Statistics” işaretli olsun. >OK. Aşağıdaki çıktıyı elde ederiz:

			Statistic	Std. Error
Age	Mean		54,44	,603
	95% Confidence Interval for Mean	Lower Bound	53,26	
		Upper Bound	55,63	
	5% Trimmed Mean		54,33	
	Median		54,00	
	Variance		156,069	
	Std. Deviation		12,493	
	Minimum		22	
	Maximum		99	
	Range		77	
	Interquartile Range		18	
	Skewness		,157	,118
	Kurtosis		,048	,235

%95 güven aralıklarının alt ve üst sınırları 53,26 ile 55,63'tür.

Normal dağılım bilgilerimizi varyansın (σ^2) toplumdaki değerini bilmemiz halinde kullanmalıyız. Diğer taraftan, örneklem sayısı az olduğunda, verilerin ancak toplumda normal dağılması halinde ampirik dağılımımızın normal dağılıma benzeyeceği bilgisini de hatırlatmalıyız.

Orantılar için güven aralığı

Daha önce bahsedildiği gibi, yüzdelere çalıştığımızda binomiyal dağılım söz konusu olur. Bununla birlikte, örneklem sayısı n yeterince büyük olduğunda dağılımın normale benzeyeceği belirtilmişti. Örneklemimizde bir olayın görülme olasılığını $p = r/n$ şeklinde hesaplayabiliriz (r = araştırdığımız durumun örneklemde kaç kişide görüldüğü; n = örneklem sayısı). Bu olasılığın standart hatasının ise

$$SE(p) = \sqrt{\frac{p(1-p)}{n}}$$

formülünden hesaplandığını hatırlayınız.

Bu durumda orantının %95 güven aralığı

$$p - 1,96 \times \sqrt{\frac{p(1-p)}{n}} \quad \text{ile} \quad p + 1,96 \times \sqrt{\frac{p(1-p)}{n}}$$

arasında olacaktır.

Örneklem sayısının küçük olması halinde (Bu durumda np veya $n(1-p)$ 5'ten küçük olur) kesin güven aralıklarını hesaplayabilmek için Binomiyal dağılım tablosundan yararlanmalıyız. Bu arada, p 'nin % olarak ifade edilmesi durumunda formülde $(1 - p)$ yerine $(100 - p)$ kullanılması gerektiğinin belirtmeliyiz.

Ortanca için güven aralığı

Verilerimizin normal dağılmaması durumunda da ortancanın güven aralığını hesaplamak mümkündür. Yorumlanması aynı olmakla birlikte, burada güven aralıkları ortalamanın etrafında eşit mesafede olmayabilir. Hesaplama aşağıdaki formüllerden yararlanılır: %95 güven aralığının;

alt sınırındaki değerin sıra yerini belirlemek için $\frac{n}{2} - \frac{1.96\sqrt{n}}{2}$

üst sınırındaki değerin sıra yerini belirlemek için $1 + \frac{n}{2} + \frac{1.96\sqrt{n}}{2}$

Örnek verecek olursak: 11 bireyin olduğu bir örneklemde ortanca değer sıralamadaki 6. değer olacaktır. Güven aralığı ise $11/2 - (1.96 * \sqrt{11})/2$ ile $(1 + 11/2 + (1.96 * \sqrt{11})/2)$ yani **2,2** ile **9,7** değerler arasında olacaktır.

Onbir katılımcının boylarını ölçsek ve verilerimiz aşağıdaki gibi sıralansa:

1	2	3	4	5	6	7	8	9	10	11
155	159	163	165	166	166	169	175	177	185	192

Bu örneklem ortancası 166, %95 güven aralığı ise [159-185] olacaktır (2,2 ve 9,7. değer olmayacağı için 2. ve 10. değeri aldık).

Güven aralıklarının yorumlanması

- **Güven aralığı ne kadar geniş?**
Geniş bir güven aralığı, tahminimizin hassas olmadığını gösterir. Güven aralığının dar olması durumunda tam tersi geçerlidir. Güven aralığı standart hataya, o da standart sapma ve örneklem sayısına bağlı olduğundan, küçük bir örneklem ve variabilitesi yüksek bir değişken olması halinde güven aralığı da geniş olacaktır. (Sayfa 124'da "Ne kadar küçük "küçük"tür?" başlığına da bakabilirsiniz).
- **Hangi klinik çıkarımlar yapabiliriz?**
Güven aralığının üst ve alt sınırları klinik olarak yorumlanabilir. Güven aralığı (GA) genelde köşeli parantez içerisinde arasında çizgi ile gösterilir. Sistolik kan basıncı için GA [95 – 135] olması durumunda örneklemimizin bir kısmının klinik olarak prehipertansiyon sınırında olduğunu söyleyebiliriz.

• **Belli önemi olan değerleri içeriyor mu?**

Toplumdan alınan bir örneklemin ortalamasının toplumun güven aralığının içerisinde olması durumunda ilgili örneklemin toplumdan farklı olmadığı yorumunu yapabiliriz. Örn. Sistolik kan basıncı için GA'nın [95 – 135] olması durumunda sistolik kan basıncı ortalaması 120 olan örneklem %95 olasılıkla bu toplumdan farklı değildir.

Serbestlik derecesi

İstatistikte Serbestlik Derecesi (Degrees of Freedom = df) terimiyle sık karşılaşırız. Serbestlik Derecesi genelde SD ile gösterilir ve örneklem sayısı – örneklemde alacağı değer serbest olmayan birey sayısı şeklinde hesaplanır. Basit bir örnekle: 3 kişilik bir örneklemde boy ölçümü yaptığımızı varsayalım. Örneklemimiz $n = 3$ olacaktır. Bireylerimizin boylarının toplamı 500 cm olsun. İlk iki bireyimiz geniş bir yelpazede herhangi bir boy değerine sahip olabilir. Ancak, 1. ve ikinci bireylerimizin boylarını belirledikten sonra (burada 165 ve 175 olsun) 3. bireyimizin boyu artık serbest değildir. Bu bireyimiz 160 cm boyunda olmak zorundadır (500-165-175). Bu nedenle $SD = n - 1 = 3 - 1 = 2$ olacaktır.

Örnek

Aydemir Ç, Kasım İ, Cebeci S, Göka E, Tüzer V. Kronik böbrek yetmezliği hastalarının yakınlarında yaşam kalitesi ve psikiyatrik semptomlar. *Kriz Dergisi* 2002;10(2):29-40 (<http://dergiler.ankara.edu.tr/dergiler/21/103/785.pdf>)

Makalede Tablo 4'deki verilerden somatizasyon puanı için kadınlar ve erkeklerde güven aralıklarını hesaplayalım:

Tablo-1 : Çalışma ve kontrol gruplarının sosyo-demografik özellikleriyle ilgili veriler.

Özellik		1. Grup	2. Grup	Kontrol Grubu
Yaş		51,00±15,39	43,76±15,07	47,56±7,42
Cinsiyet	Kadın	20	18	16
	Erkek	5	7	9

Grup 1	Grup 2
- Ortalama yaş 1 = 51 - Standart sapma 1 = 15,39 $n_1=25$	- Ortalama yaş 2 = 43,76 - Standart sapma 2 = 15,07 $n_2 = 25$ kişi

• $SD = 24$; $p = 0,05$ için tablo t değeri = 2,064 • $SEM_1 = 15,39 / \sqrt{25} = 3,078$ • %95 GA 1 $[51 \pm 2,064 \times 3,078]$ • [44,65 – 57,35]	• $SD = 24$; $p = 0,05$ için tablo t değeri = 2,064 • $SEM_2 = 15,07 / \sqrt{25} = 3,014$ • %95 GA 2 $[43,76 \pm 2,064 \times 3,014]$ • [37,54 – 49,98]

Her iki grupta da yaşı %95 güven aralığının nispeten geniş olduğunu ve grupların güven aralıklarının kesiştiğini görüyoruz.

Benzer bir alıştırmayı aynı araştırmadaki Tablo 4'te kadınlar ve erkeklerin arasındaki somatizasyon puanları için yapacak olursak:

Tablo-4: Hasta yakınlarının cinsiyetleri ile KSE alt ölçekleri ve indekslerinin karşılaştırılması.

KSE Alt Ölçek ve İndeksleri	Kadın (N=38)	Erkek (N=12)	P
Somatizasyon	5,95±5,19	0,75±1,14	0,001
Obsessif-Kompulsif Bozukluk	5,42±3,85	2,08±1,68	0,005
Kişilerarası Duyarlılık	3,50±2,73	1,50±1,24	0,02
Rahatsızlık Ciddiyeti İndeksi	0,68±0,49	0,29±0,21	0,005
Belirti Toplamı İndeksi	19,39±9,90	12,42±8,18	0,013
Semptom Rahatsızlık İndeksi	1,70±0,58	1,19±0,22	0,003

Kadın	Erkek
• Ortalama 1 = 5,95 • Standart sapma 1 = 5,19 • $n_1 = 38$ • $SD = 37$; $p = 0,05$ için tablo t değeri ~ 2,02 • $SEM_1 = 5,19 / \sqrt{38} = 0,84$ • %95 GA 1 $[5,95 \pm 2,02 \times 0,84]$ • [4,25 – 7,65]	• Ortalama 2 = 0,75 • Standart sapma 2 = 1,14 • $n_2 = 12$ kişi • $SD = 11$; $p = 0,05$ için tablo t değeri = 2,201 • $SEM_2 = 0,75 / \sqrt{12} = 0,22$ • %95 GA 2 $[0,75 \pm 2,201 \times 0,22]$ • [0,26 – 1,23]

Farkların güven aralığı

Nümerik veriler için

Nümerik veri iki grupta farkların standart hatasını hesaplamak için şu formülü kullanırız: **SEM(fark)**= $\sqrt{[(s_1^2/n_1) + (s_2^2/n_2)]}$

Tablo 1'deki örneğimizi uygulayacak olursak;

$$SEM = \sqrt{[(15,39 \times 15,39/25) + (15,07 \times 15,07 / 25)]} = 4,30$$

$$\%95 \text{ GA} = \text{Ortalama 1} - \text{Ortalama 2} \pm t_{0,05} \times SEM$$

Burada serbestlik derecesi $SD = (n_1-1) + (n_2-1)$ olarak hesaplanır $(24+24=48)$. $\%5$ anlamlılık düzeyi için tablo t değeri yaklaşık 2,009'dur.

$$\%95 \text{ GA} = (51 - 43,76) \pm 2,009 \times 4,30$$

İki ortalama arasındaki farkın $\%95$ güven aralığı: **[-1,4 – 15,87]**

Not: İki ortalama arasındaki farkın güven aralığını yorumlarken güven aralığının sıfırı içermesi durumunda ortalamaların arasında anlamlı bir farkın olmadığı sonucuna varırız (yani iki ortalama arasındaki fark artı olabileceği gibi sıfır da olabilir, hatta eksi bile olabilir. Bu durumda ortalamalardan birinin ötekinden büyük olduğunu iddia edemeyiz).

SPSS ile ortalamalar arasındaki farkın güven aralığının hesaplanması

<https://drive.google.com/open?id=0B0ieO-Fy3z-5M0o1WV92NFpBWmM>

adresinden indirebileceğiniz diyabet.sav veri setinde yaş değişkeni için kadınlar ve erkekler arasındaki ortalama farkın $\%95$ güven aralığını hesaplayalım:

Analyze > Compare Means > Independent-Samples t test [“Test variables” kutusuna “Age” değişkenini, “Grouping variable” kutusuna “sex” değişkenini koyalım. “Define Groups” butonunu tıklayıp “Group 1” için 1, “Group 2” için 2 yazalım > Continue > OK. Aşağıdaki çıktıyı elde ederiz:

Sex of the patient	N	Mean	Std. Deviation	Std. Error Mean
Age Male	235	56,20	12,662	,826
Female	194	52,31	11,975	,860

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
		Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower
Age	Equal variances assumed	,219	,640	3,238	427	,001	3,881	1,199	1,525	6,237
	Equal variances not assumed			3,255	419,167	,001	3,881	1,192	1,538	6,225

Erkekler ve kadınların yaş farkının güven aralığı [1,53-6,24] arasındadır.

Orantılar için

Kategorik verilerle çalıştığımızda orantılar arasındaki farkın standart hatası

SEM(fark) = $\sqrt{[(p_1q_1/n_1) + (p_2q_2/n_2)]}$ formülünden hesaplanır. İncelediğimiz araştırmadaki Tablo 1'e göre 1. Grupta 20 kadın, 2. Grupta 18 kadın bulunuyor. Bu gruplar arasında cinsiyetler arasındaki farkın güven aralığına bakalım:

$$SEM(fark) = \sqrt{[(20/25) \times (5/25)/25] + [(18/25) \times (7/25)/25]} = 0,12$$

$$\%95 \text{ GA} = \text{Ortalama 1} - \text{Ortalama 2} \pm t_{0,05} \times SEM$$

$$\%95 \text{ GA} = (0,8 - 0,72) \pm 2,009 \times 0,12$$

$$\text{İki yüzde arasındaki farkın güven aralığı: } [-0,16 - 0,32]$$

Yani kadınlar ve erkekler arasındaki fark %16 erkekler lehine olabileceği gibi %32 kadınlar lehine de olabilir.

Odds oranları için güven aralığı

Odds oranı daha sonraki konularda işlenecektir. Ancak, burada güven aralığı açısından yorumlayabilmek için kısaca değinilecektir.

Odds oranı bir olayın olma olasılığının olmama olasılığına oranı olarak tanımlanabilir. Yani sigara içenlerde kanser olma olasılığı / sigara içmeyenlerde kanser olma olasılığı. Bu oran, risk faktörlerinin hesaplanmasında sıkça kullanılan önemli bir parametredir. Akciğer kanseri örneğinde odds oranı 10 olsa, sigara içenler içmeyenlere göre 10 kat daha fazla kanser oluyorlar gibi bir yorum yapabiliriz. Araştırmalarda odds oranı bir değer olarak verilmekle birlikte beraberinde güven aralığının da belirtilmesi daha uygun olur. Odds

oranlarının güven aralığının 1'i içermesi durumunda risk açısından bir anlamlılık olmadığı yorumu yapılabilir.

Relative Deviate formülünün farkların güven aralığına uygulanması

Daha önceden gördüğümüz gibi relative deviate (rölatif sapma) $RS = (\text{Ortalama 1} - \text{Ortalama 2}) / \text{farkların standart hatası}$ formülünden hesaplanıyordu.

Yukarıdaki örneklere uygularsak:

- Tablo 1'de Grup 1 ve Grup 2 arasındaki yaş farkı için: $(51,0 - 43,76) / 4,30 = 1,68$
Sonuç 1,96'dan küçük olduğundan aradaki fark istatistiksel olarak anlamsızdır.
- Tablo 1'de Grup 1 ve Grup 2 arasındaki cinsiyet farkı için: $(0,8 - 0,72) / 0,12 = 0,66$
Sonuç 1,96'dan küçük olduğundan aradaki fark istatistiksel olarak anlamsızdır.
- Diyabet.sav veri setinde erkekler ve kadınlar arasındaki yaş farkı için: $(56,2 - 52,31) / 1,199 = 3,24$
Sonuç 1,96'dan büyük olduğundan aradaki fark istatistiksel olarak anlamlıdır.

Bölüm 15: Araştırma Tasarımı – I

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun araştırma tasarımları hakkında genel bilgi sahibi olması, araştırma tasarımlarını sınıflandırabilmesi ve açıklayabilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Araştırma tasarımlarının önemini tartışabilmek
- Cevaplanabilir klinik soru formatını açıklayabilmek
- Deneysel araştırmaları açıklayabilmek
- Gözlemsel araştırmaları açıklayabilmek
- “Nedensellik” terimini ve önemini açıklayabilmek
- Araştırmalarda kontrol grubunun önemini açıklayabilmek
- Araştırmalardaki hata çeşitlerini (bias) açıklayabilmek

Daha araştırma planlama aşamasında araştırmamızın adını koymamız ve özelliklerini belirlememiz çok önemlidir. Tasarım aşamasının iyi düşünülmediği araştırmalarda araştırmamızın eksikleri veri toplama veya raporlama aşamasında ortaya çıkabilir. Bu durumda geriye dönüp eksikleri giderme imkanımız da olmadığından emek ve kaynaklarımızı boşa harcamış olma riskimiz vardır.

Örneğin, kırsal bölgedeki bir okulda sağlık taraması yaptığımızı düşünelim. Eğer araştırmamızı iyi planlamazsak verileri topladıktan sonra fark edeceğiz ki, sonuçlarımızın şehirdeki öğrencilerden farklı olduğunu göstermek için şehirdeki bir okulu da kontrol grubu olarak almamız gerekirdi.

Araştırma tasarımı yaparken mutlaka yapmamız gereken iki şey vardır: birisi **klinik olarak cevaplanabilir bir soru** sormak (Problem nedir? Neden araştırma yapıyoruz?) ve ikincisi de bununla bağlantılı olarak neyi ölçmek istediğimizi (sonuç ölçütlerimiz, outcome measures) belirlemektir.

Klinik olarak cevaplanabilir soruda 4 özelliğin olması gerekir. Hatırlamamız için İngilizce PICO akronimi faydalı olabilir: **P**roblem (hasta, hastalık ya da problem), **I**ntervention (uyguladığımız girişim), **C**omparison (karşılaştırma) ve **O**utcome (incelemek istediğimiz sonuç / çıktı). Bu akronime bir de **S** harfi (setting, araştırmanın yapıldığı ortam) ekleyenler de vardır.

Örneğin: “Şişmanlığın tedavisinde kalori diyeti ile glisemik endekse dayalı bir diyet arasında vücut kitle endeksini (VKİ) düşürme açısından fark var mıdır?” (P: şişmanlık, I: kalori diyeti, C: glisemik diyet, O: VKİ).

PICO formatı her araştırma tasarımına uygulanamayabilir. Tanımlayıcı epidemiyolojik bir araştırma yapıyorsak sorumuz “Toplumumuzdaki erişkinlerde diyabet yaygınlığı ne kadardır?” şeklinde olabilir.

PICO formatından yararlınsak da yararlanmasak da sonuç ölçütlerimizi mutlaka belirlemeliyiz. Tercih edilen, bir araştırmada *sadece bir ana sonuç ölçütü* olmasıdır. Yukarıdaki şişmanlık örneğinde VKİ iyi bir örnektir. Bununla birlikte araştırmalarda genelde çok sayıda ek değişken de (örn. Kan basıncında düşme, yaşam kalitesinde artma, depresyon skalasında düşme vs.) incelenmeye çalışılır. Sonuç ölçütlerinin çok olması araştırma hatalarını artıracaktır.

Deneysel / gözlemsel araştırmalar

Deneysel araştırmalarda araştırmacı deneklere bir şekilde **müdahalede** bulunur. Bu müdahale, **hayvan deneylerinde** bir ilaç verip etkisini araştırmak, **toplum deneylerinde** iyot verip guatr yaygınlığındaki değişimi incelemek veya **insan deneylerinde** diyabetli gruplara bir eğitim verip kan şekeri kontrolündeki değişimi araştırmak şeklinde olabilir.

Deneysel araştırmalarda incelenen değişkene (**sonuç ölçütü**, bağımlı değişken) etki eden faktörleri kontrol etmek kısmen mümkün olduğundan en güvenilir bilgiyi bu araştırma tasarımından elde ederiz. Bununla birlikte, bu araştırma tasarımı daha masraflıdır ve etik sorunlar diğer araştırma tasarımlarından daha fazladır.

Gözlemsel araştırmalarda ise araştırmacı bir müdahalede bulunmayıp sadece gözlem yapar. Bu araştırma tasarımında elde ettiğimiz bilgilerin değeri daha düşüktür. Bununla birlikte, özellikle epidemiyolojik araştırmalarda başka bir tasarım söz konusu olmayabilir ve sonuçları klinik açıdan çok yararlı olabilir. John Snow’un Londra’da kolera salgınlarını gözlemleyerek etkenin kaynağını tahmin etmesi klasik bir gözlemsel araştırma örneğidir.

Araştırma çeşitleri

Araştırma tiplerini sınıflandırırken hangi açıdan baktığımız önemlidir. Bakış açısına göre Tablo 1’deki şekilde sınıflandırmalar yapılabilir:

Tablo 1: Çeşitli yönlerden araştırma tasarımlarının sınıflandırılması.

Analiz	• Tanımlayıcı • Analitik
Müdahale varlığı	• Gözlemsel • Deneysel
Zaman açısından	• İleriye doğru (prospektif) • Geriye doğru (retrospektif) • Anlık tanımlayıcı (kesitsel)
Kontrol grubu	• Kontrollü • Kontrolsüz
Körleme	• Açık etiketli • Tek kör • Çift kör • Üçlü kör
Örnekleme	• Rastgele (randomize) • Randomize değil

Dolayısıyla bir araştırma tanımlayıcı, analitik ve ileriye doğru olabilir. Diğer taraftan, deneysel araştırmalar doğası gereği zamanda ileriye doğru olduğundan “deneysel ve prospektif” ifadesi gereksizdir. “Körleme” ise deneysel araştırma tasarımları için söz konusu olan bir unsurdur. Kesitsel bir araştırma ise, **kesitsel tanımlayıcı** olabileceği gibi **kesitsel analitik** de olabilir.

Zaman içerisindeki akışı ve tipik kullanımlarıyla birlikte çeşitli araştırma tasarımları Tablo 2’de özetlenmiştir.

Araştırma çeşitlerine birer örnek Tablo 3’te verilmiştir.

Tablo 2: Sık kullanılan araştırma çeşitleri (kanıt değerine göre azalan sırayla)

Çeşidi	Zaman	Şekli	Geçmiş	Şimdi	Gelecek	Gelecek	Kullanımı
Deneysel	Prospektif	Deneysel		Müdahale	Takip	Sonuçların gözlenmesi	- Tedavi etkisini araştırma - Önleyici girişimleri araştırma - Laboratuvar deneyleri
Kohort	Prospektif	Gözlemsel		Kohortun tanımlanması ve risk faktörlerinin değerlendirilmesi	Takip	Sonuçların gözlenmesi	- Prognoz ve hastalığın doğal seyri - Etyoloji
Vaka kontrol	Retrospektif	Gözlemsel	Risk faktörlerinin değerlendirilmesi	Vakalar ve kontrollerin belirlenmesi			Etyoloji (özellikle nadir hastalıkları için)
Tekrarlayan kesitsel	Zamanda bir an	Gözlemsel		Veri toplama	Veri toplama	Veri toplama	Zaman içerisindeki değişim (trend)
Kesitsel	Zamanda bir an	Gözlemsel		Veri toplama			- Prevalans - Referans aralıkları - Bir grubun sağlık durumu

Tabloda belirtilenlerin dışında metodolojik araştırmalar (geçerlilik/güvenilirlik araştırmaları gibi), kalitatif (nitel) araştırmalar ve ekolojik araştırmalar gibi farklı araştırma yöntemleri de vardır.

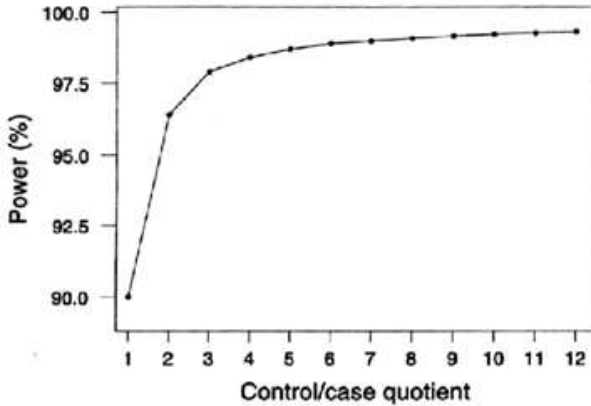
Tablo 3: Araştırma çeşitlerine örnekler.

Araştırma çeşidi	Örnek
Kesitsel tanımlayıcı	Yazar: Geyik MF ve ark. Başlık: Tıp Fakültesi Öğretim Üyelerine Uygulanan Eğiticilerin Eğitimi Etkinliğinin Araştırılması Amaç: Bu çalışmada eğiticilerin eğitimi kursunu alan öğretim üyelerinin eğitime bakış açılarını ortaya koymak hedeflenmiştir. http://konuralptipdergi.duzce.edu.tr/Dokumanlar/sayi5-3/06_KTD-2013-5-3.pdf
Kesitsel analitik	Yazar: Bölge KN ve ark. Başlık: Lise son sınıf öğrencilerinde hedef odaklı olmak sınav kaygısını nasıl etkiliyor? Amaç: Çalışmamızın amacı lise son sınıf öğrencilerinde hedefe yönelimin sınav kaygısı üzerine etkisini araştırmaktır. http://www.tahud.org.tr/dokuman/download?fl=493
Tekrarlayan kesitsel	Yazar: Aktürk Z ve ark. Başlık: Türkiye’de 2008-2011 Yılları Arasında Aile Hekimlerinin TAHUD, Özlük Hakları ve Sevk Zinciri Hakkındaki Görüşleri: Zamanla Neler Değişti? Amaç: 2008 yılından başlayarak her yıl aile hekimliği uzmanlarının görüşlerinin bir anket aracılığıyla sorulması amaçlandı. http://dergipark.ulakbim.gov.tr/tjfmpe/article/download/5000055908/5000053117
Kohort	Yazar: Arnlov ve ark. Başlık: Relations of Insulin Sensitivity to Longitudinal Blood Pressure Tracking Amaç: We investigated the relations of insulin sensitivity to the development of hypertension and BP tracking in a large community-based sample of middle-aged adults stratifying our analyses by age, baseline BMI, and BP level. http://circ.ahajournals.org/content/112/12/1719.long
Vaka kontrol	Yazar: Akin E, Berkem M. Başlık: İntihar Girişiminde Bulunan Ergenlerde Olumsuz Yaşam Olayları: Vaka Kontrol Çalışması Amaç: Bu çalışmada intihar girişiminde bulunan 12-18 yaş aralığındaki ergenlerde olumsuz yaşam olaylarının incelenmesi ve kontrol grubuyla karşılaştırılması amaçlanmıştır. http://www.firattipdergisi.com/text.php3?id=814
Deneysel	Yazar: Özger H ve ark. Başlık: Çeşitli kemik replasman materyallerinin sıçanlarda kemik rejenerasyonu üzerine etkisi Amaç: Biz bu deneysel çalışmamızda kemik replasmanında en çok kullanım alanı bulan transplantasyon ve replasman yöntemlerini aynı şartlar altında karşılaştırmalı olarak uygulamayı ve bunların kemik rejenerasyonu üzerine etkilerini gözlemlenmeyi rasyonu amaçladık. http://www.aott.org.tr/article/download/5000008583/5000008382
Geçerlik/ Güvenirlik	Yazar: Taştan K ve ark. Başlık: Development and Validation of a Personality Type Inventory Based on Enneagram

	Amaç: The purpose of this study was to develop an inventory measuring personality types based on the Enneagram principles and test its validity and reliability among a Turkish speaking community.
Duyarlılık/ Özgüllük	Yazar: Kunduracioğlu A ve ark. Başlık: Yayma Negatif Tüberküloz Olgularının Tanısında MTD Gen-Probe® Testi, BACTEC 960™ Sistemi ve Löwenstein-Jensen Kültür Yöntemlerinin Performansının Karşılaştırılması Amaç: Bu çalışmada, klinik ve radyolojik olarak TB kuşkusu olan, ancak ARB yayması negatif hastaların pulmoner örneklerinde MTD Gen-Probe® testinin, akciğer TB tanısındaki duyarlılık ve özgüllüğünün araştırması, sonuç raporlama sürelerinin BACTEC 960™ (MGIT) sistemi ve LJ kültür sonuç süreleriyle karşılaştırılması amaçlanmıştır. http://www.mikrobiyolbul.org/fulltext.aspx?issue_id=189&ref_ind_id=21273
Nitel	Yazar: Güldal D ve ark. Başlık: Aile hekimliği çekirdek yeterlilikleri hastaların beklentileri ile uyumakta mıdır? Amaç: Bu niteliksel çalışmada birinci basamak sağlık hizmetleri konusunda hasta beklentilerinin neler olduğu, bu beklentilerin aile hekiminin özellikleri ve tanımlanmış temel yeterlilikleri ile ne kadar örtüştüğünün araştırılması amaçlanmıştır. http://www.turkailehekderg.org/makaleler/arastirma/aile-hekimligi-cekirdek-yeterlilikleri-hastalarin-beklentileri-ile-uyusmakta-midir/
Ekolojik	Yazar: Genowska A ve ark. Başlık: Environmental and socio-economic determinants of infant mortality in Poland: an ecological study. Amaç: The aim of the present ecological study was to assess the relationship between occupational environment, industrial pollution, socio-economic status and infant mortality in Poland. http://www.ncbi.nlm.nih.gov/pubmed/26195213

Kesitsel analitik araştırmalar ile vaka kontrol araştırmaları bazen karıştırılmaktadır. Şu açıklamalar ayırımı yapmayı kolaylaştırabilir:

- Kesitsel araştırmaların esas amacı bir durumun yaygınlığını saptamaktır. Vaka kontrol araştırmaları bu tür bir bilgi sağlamaz.
- Vaka kontrol araştırmalarında bir olay meydana gelmiştir ve bu olaya neden olan risk faktörü/faktörleri araştırılmaktadır. Kesitsel analitik araştırmalarda ise incelenen popülasyonda bir bağımlı değişkenin durumu (örn. Sigara içen/içmeyen) çeşitli gruplar arasında karşılaştırılmaktadır (örn. erkek/kadın)
- Vaka kontrol araştırmalarında en az vaka sayısı kadar kontrol seçilir (kontrol grubunun sayısı vakaların 2-4 katına kadar olabilir). Kesitsel araştırmalarda karşılaştırılan gruplardaki birey sayılarını araştırmacı belirlemez.



(Woodward, M. Epidemiology Study Design and Analysis. Boca Raton: Chapman and Hall; 1999, p.265)

- Vaka kontrol arařtırmalarında kontroller aynı örneklemden rastgele seçilmiş bireyler olmayıp kontrollerin ailelerinden, çevrelerinden veya hastaneden seçilebilir.

Nedensellik

Arařtırma yapmadaki amacımız genelde bir **neden – sonuç ilişkisi** bulmaktır. Neden – sonuç ilişkisini inceleyen en güvenilir arařtırmalar deneysel arařtırmalardır. Bununla birlikte, gözlemsel arařtırmalara bakarak ta nedenler hakkında çıkarımlar yapabiliriz. Ancak, her arařtırma sonunda bir nedensellik çıkarımı yapamayabiliriz. Sigara içenlerde içmeyenlere oranla akciğer kanserinin anlamlı derecede daha sık görölmesini bulduğumuzda “Sigara içilmesiyle akciğer kanseri gelişmesi arasında bir ilişki vardır” deriz. “Sigara akciğer kanseri yapar” diyebilmemiz için bazı kriterlerin dikkate alınması gerekir. Bu konuda en iyi bilinen kriterler Hill tarafından önerilenlerdir (Austin Bradford Hill, 1897-1991):

- Sebep sonuçtan önce olmalıdır.
- İlişki mantıklı olmalıdır, yani sonuçlar biyolojik açıdan açıklanabilmelidir.
- Bulgular birkaç arařtırma tarafından desteklenmelidir.
- Sebep ve sonuç arasındaki ilişki kuvvetli olmalıdır.
- Sebeple ilgili doz – cevap ilişkisi olmalıdır.
- Etken faktörü ortadan kaldırmak hastalık riskini azaltmalıdır.
- Alternatif açıklamalar (hipotezler) dikkate alınmalıdır.

- Uygun bir deneyle durum ortaya konulabilmelidir.

Kesitsel / longitudinal araştırmalar

Kesitsel araştırmalar zamanda belli bir noktayı inceler. Bu araştırmalar özellikle toplumdaki bir durumun nokta prevalansını incelemek için yararlıdır.

Nokta prevalansı = Belli bir zaman anında hastalığı olan bireyler / Aynı andaki toplam nüfus

Burada sebep – sonuç arasındaki zaman ilişkisini bilemediğimizden ilgili faktör hastalığa *yol açmıştır* diyemeyiz. Ancak “risk faktörüyle hastalık arasında bir ilişki vardır” diyebiliriz. Diğer taraftan hastalığın insidansını veya seyrini (trend) hesaplamak için de bu araştırma tasarımı uygun değildir. Ancak, ucuz ve çabuk yapılan araştırmalar olduğundan bu tasarım çokça tercih edilmektedir.

Zaman içerisinde farklı noktalarda **tekrarlayan kesitsel araştırmalar** (örn. Her yıl tıp fakültesine kayıt yaptıran öğrencilerde riskli davranış yaygınlığını incelemek) yapılabilir. Bu araştırma tasarımı belli bir durumun zaman içerisindeki seyrini (trend) incelemek için yararlıdır. Ancak, her defasında farklı bir topluluk söz konusu olduğundan sonuçların zaman içerisindeki değişimden mi, yoksa örneklemdaki farklılıklardan mı kaynaklandığını kestirmek zordur.

Longitudinal araştırmalarda bir örneklem zaman içerisinde belli bir süre takip edilir. Bunlara **prospektif** araştırmalar denir. Geçmişe ait güvenilir bilgilere ulaşabilmemiz halinde geçmişten başlayıp şimdiki zamana doğru inceleme yapan prospektif araştırmalar da tasarlanabilir. Longitudinal araştırmaları yapmak daha fazla zaman ve kaynak gerektirir. Verilerin deneklerin hafızasından veya hasta kayıtlarından elde edilmesi halinde bu araştırma tasarımında önemli hatalar da söz konusu olabilir.

Deneyisel araştırmalar prospektif olarak değerlendirilirler. Gözlemsel araştırmalar ise hem prospektif, hem de retrospektif olabilir.

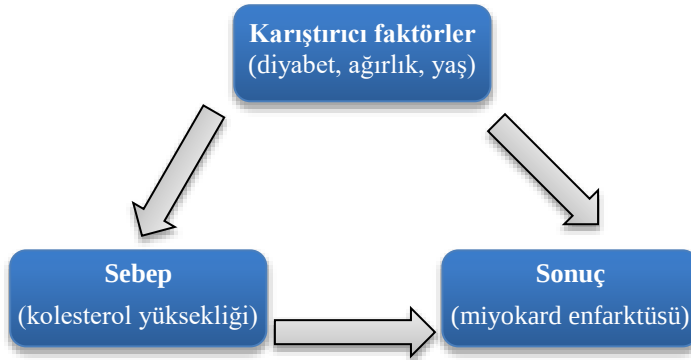
Kontrol grupları

Herhangi bir araştırma tasarımında veya sonuçların değerlendirilmesinde bir kontrol grubu kullanılması elzemdir. Örneğin, sigara ile akciğer kanseri ilişkisini araştırmak istediğimizde kanser olanlar ve olmayanlar olmak üzere iki örneklem grubumuz olmalı. Aksi halde sadece kanser olanların incelenmesini klinik açıdan yorumlamak ve nedensellik açısından değerlendirmek zor olacaktır.

Hatalar (bias)

Araştırma sonuçlarımız ve toplumdaki gerçek durum arasında fark olması halinde hata yapmış oluruz. Araştırmalarda sık karşılaşılan hata çeşitleri aşağıda sıralanmıştır:

- **Gözlemci hatası:** gözlemciden kaynaklanan hatalardır. Bir anketörün boyları ölçerken deneklerin ayakkabılarını çıkarmaması gibi.
- **Karıştırıcı faktörler** (confounding): hem risk faktörü, hem de sonuç ölçütünü etkileyen faktörlerin ayarlanamamasıdır. Örn. erkek olmak kalp krizi için bir risk faktörüdür. Erkekler aynı zamanda kadınlara göre daha fazla sigara içerler. Bu durumda sigara ile kalp krizi arasındaki ilişkiyi araştırdığımızda cinsiyeti de dikkate almamız (cinsiyet açısından düzeltme yapmamız) gerekir. Karıştırıcı faktörleri aşağıdaki şekilde şematize edebiliriz:



- **Seçime bağlı hatalar:** araştırmaya dahil edilen hastaların toplumu temsil etmemesi durumudur.
- **Ölçüme bağlı hatalar** (information bias): ölçümler sistematik bir şekilde hatalı yapılmaktadır. Laboratuvarda miyadı geçmiş bir kit kullanılması veya ağırlık ölçerken baskülümüzün kalibrasyonunun hatalı olması örnek olarak verilebilir.
- **Yayın hataları:** araştırmacılar genelde sadece anlamlı çıkan sonuçları yayınlama meylinde.

Bunların dışında **hatırlama** hataları, **sağlıklı birey etkisi** (healthy entrant effect), **değerlendirme** (assessment) ve **gruplara ayırma** (allocation) hataları olabilir.

Alıştırma 6

1. <https://drive.google.com/open?id=0B0ieO-Fy3z-5QUQ1bWJtWjdCZnM> adresinden odev6.sav isimli SPSS dosyasını kullanarak kızlarla erkeklerin yaşları arasında istatistiksel bir fark olup olmadığına bakalım. Söz konusu problemde bağımsız iki grupta nümerik verileri karşılaştıracamız (Sayfa 41'deki akış şemasına bakınız). Bu amaçla yapacağımız bağımsız gruplarda t testine SPSS'te Analyze> Compare means>Independent samples t test sekmelerinden ulaşabiliriz. Test değişkenimize yaş, grouping variable olarak da cinsiyet (0 ve 1) girelim. OK dediğimizde aşağıdaki tabloları elde etmeliyiz:

Group Statistics					
	Cinsiyet	N	Mean	Std. Deviation	Std. Error Mean
Deneklerin yaşı	Kız	1564	18,38	1,373	,035
	Erkek	1546	18,25	1,298	,033

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Deneklerin yaşı	Equal variances assumed	2.112	.148	2.680	3108	.007	.128	.048	.035	.221
	Equal variances not assumed			2.690	3101.685	.007	.128	.048	.035	.221

Görüldüğü üzere $p=0,007$ olacak şekilde kızlarla erkeklerin yaşları arasında istatistiksel olarak anlamlı bir fark vardır. Kızların ve erkeklerin yaş ortalamasının %95 güven aralığını hesaplayarak sonucu aşağıdaki tabloya yazınız ve sonucu yorumlayınız.

	Cinsiyet	N	Mean	Std. Deviation	%95 güven aralığı
Deneklerin yaşı	Kız	1564	18,38	1,37	
	Erkek	1546	18,25	1,30	

Not: Bulduğunuz güven aralığını analyze, explore sekmesinden de hesaplatarak sağlamasını yapabilirsiniz.

2. Türk Psikiyatri Dergisi (<http://www.turkpsikiyatri.com/>) 2005; 16(2):77-82 yayınlanan “Depresyonlu Hastalarda Kemik Mineral Yoğunluğunda Değişiklik Var mı?” isimli makalede Tablo 1’de hastalarla kontrollerin yaşları arasında anlamlı bir farkın olmadığı belirtilmektedir ($p>0,05$). Yaşların %95 güven aralıklarını hesaplayarak durumun sağlamasını yapınız.

TABLO 1. Hasta ve Kontrollerin Demografik ve Klinik Özellikleri.

	Hastalar (n=42)	Kontroller (n=23)	Karşılaştırma
Yaş (ya)	37.57±8.70	33.73±7.16	t=1.80; p>0.05
Cinsiyet (E/K)	21/21	11/12	-
BKİ (kg/m ²)	27.06±3.46	25.25±3.56	t=1.99; p>0.05
MADRS skoru	34.97±5.71	-	-
Toplam hastalık süresi (ay)	63.60±82.95	-	-
Epizot sayısı	2.78±2.75	-	-
Son epizotun süresi (ay)	5.94±9.03	-	-

E: Erkek, K: Kadın, BKİ: Beden Kitle İndeksi.

Çözüm için bakınız s. 471

Bölüm 16: Araştırma Tasarımı – II

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun araştırma tasarımlarında varyasyonu azaltma ve tahminlerin hassasiyetini artırma konusunda bilgi sahibi olması amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

Hedefler:

- Varyasyonu azaltmak yani tahminlerimizin hassasiyetini artırmanın yollarını sayabilmek ve açıklayabilmek
 - Replikasyon
 - Örneklem büyüklüğü
 - Özel araştırma tasarımları
 - Bloklama
 - Çapraz tasarım
 - Faktöriyel deneyler

Varyasyon

Verilerimizdeki varyasyonlar ölçüm hataları gibi bilinen faktörlere bağlı olabileceği gibi açıklanamayan rastgele varyasyona da bağlı olabilir. Daha önce bahsedildiği gibi, verilerimizdeki varyasyonu standart hataya bakarak anlarız. Bir değişkende aşırı varyasyon (sapma) olması halinde geniş bir güven aralığı ve büyük bir standart hatamız olacak ve bu değişkenle ilgili tahminlerimiz hassas olmayacaktır.

Verilerimizdeki varyasyonu azaltmak ve tahminlerimizin hassasiyetini artırmak isteriz. Araştırma tasarımlarını kullanarak bunu sağlamanın çeşitli yolları vardır.

Kopyalama (Replikasyon)

Aynı bireyde birden fazla ölçüm yaparsak ölçümlerimizin doğruluk oranını artırabiliriz. Ancak, bu durumda ölçümlerimiz birbirinden bağımsız olmayacağı için analiz sırasında dikkatli olunmalıdır. Her bireyden yapılan tekrarlayan ölçümlerin (kopyaların) ortalamasını almak bir çözüm olabilir. Replikeler ölçümlerin analizi için geliştirilen özel istatistiksel yöntemleri de kullanabiliriz.

Örneklem büyüklüğü

Araştırma tasarımının en vazgeçilmez bir ögesi de örneklem büyüklüğüdür. Yeterince büyük bir örneklem alarak standart hatayı düşürebilir ve araştırmamızın hassasiyetini artırabiliriz. Araştırmamıza başlamadan önce mutlaka örneklem ve güç (power) hesaplaması yapmalıyız ve makalemizin yöntem bölümünde de bu bilgileri sunmalıyız.

Özel araştırma tasarımları

Daha önce bahsettiğimiz temel araştırma tasarımlarını değiştirerek daha hassas ölçümler yapabiliriz. Deneklerimizi gruplara ayırırken tamamen rastgele davranmamız gerekir. Bununla birlikte, gruplar arasında yine de farklılıklar olabilir. Bu durumda varyasyonu önlemek için **bloklama** veya **çapraz tasarımlar** diye bilinen yöntemler kullanabiliriz.

Bloklama

Deneklerimizi birbirine yakın, daha homojen gruplara ayırabiliriz. Böylece bu alt gruplar (bloklar) arasında varyasyon daha az olacaktır. Örneğin, yaş gruplarına göre 5'er veya 10'ar yıllık aralıklara ayırabiliriz. Karşılaştırmayı yaparken tüm denekleri karşılaştırmak yerine bu alt blokları karşılaştırdığımız takdirde varyasyon sorununu büyük oranda azaltabiliriz.

Paralel ve çapraz çalışmalar

Çoğu araştırma paralel bir tasarıma sahiptir. Çalışma ve kontrol grupları birbirine paraleldir. İki gruptaki bireyler birbirleriyle karşılaştırılır.

Daha önce gördüğümüz gibi aynı bireyde yapılan ölçümlerin varyasyonu bireyler arası ölçümlerdeki varyasyona göre daha düşüktür. Bu bilgiyi kullanarak bireylerin kendi kontrolleri olmasını sağlayabiliriz. Böylece birey içi karşılaştırmalar yapacağımızdan ölçümlerimiz daha hassas olacaktır.

Çapraz tasarım birey içi karşılaştırmalara bir örnektir. Önce her iki tedavi grubuna ilgili girişim uygulanır. Ardından, ilacın etkisinin geçmesi için bir süre beklenir (washout = yıkama dönemi). Sonra, gruplara verilen tedaviler değiştirilerek (çaprazlama) deneye devam edilir. Süre sonunda hem gruplar arasında, hem de grup içi karşılaştırma yapılabilir.

Şüphesiz, çapraz tasarımlar verilen tedavinin hastalığı tam geçirmeyip sadece geçici olarak azalttığı durumlarda (hipertansiyon, diyabet gibi) uygulanabilir.

Faktöriyel tasarımlar

Genelde araştırmalarımızda bir faktörü inceleriz (egzersiz yapmanın kilo vermeye olan etkisini araştırmak gibi). Oysa bazen birden fazla faktörü (egzersiz ve diyetin kilo vermeye olan etkisini araştırmak gibi) araştırmak isteyebiliriz. Bu durumda çok sayıda masraflı araştırma yapmak yerine faktöriyel bir tasarım düşünebiliriz. Örnek olarak "US Physicians Health study" (<http://phs.bwh.harvard.edu/phs20.htm>) verilebilir. Bu araştırmada aspirin ve beta karotenin kalp hastalığını önlemedeki etkisini araştırmak için 2x2 faktöriyel bir tasarım uygulanmıştır.

Bu araştırmadaki tedavi kombinasyonları aşağıdaki tabloda verilmiştir:

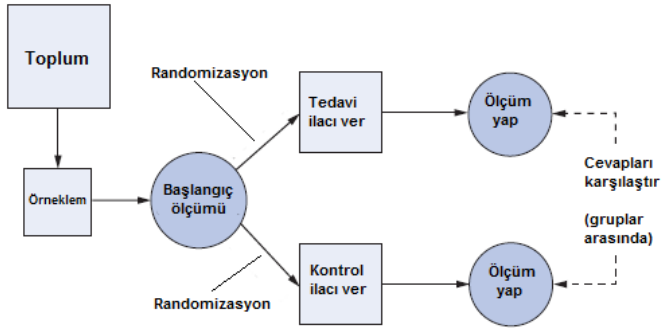
		Beta karoten	
		Hayır	Evet
Aspirin	Hayır	<i>Hiçbiri</i>	<i>Beta karoten</i>
	Evet	<i>Aspirin</i>	<i>Aspirin + beta karoten</i>

Bu araştırmada beta karotenin etkisini araştırmak için sol sütundaki denekleri sağdakilerle, aspirinin etkisini araştırmak için ise üst satırdaki hastaları alttakilerle karşılaştırırız.

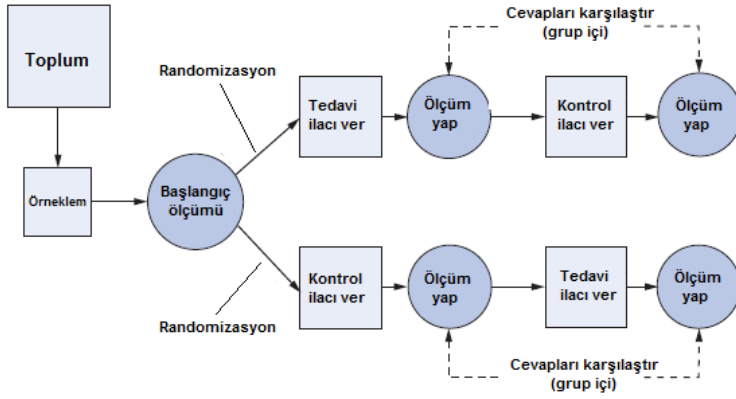
Ayrıca bu tasarımda iki faktör arasında **etkileşim** (interaction) olup olmadığını da araştırabiliriz. Etkileşim saptamamız halinde yukarıdaki örnekte aspirin ve beta karotenin birlikte verilmesinin ayrı ayrı etkilerinden daha fazla bir yarar sağlayacağı anlamına gelir. Bu yönüyle de bu araştırma tasarımı iki farklı araştırma yapılmasına göre daha üstün bilgiler sağlar. Ayrıca kaynakları daha etkin kullanmış oluruz ve daha küçük bir örnekleme daha hassas yorumlar yapabiliriz.

SPSS ile faktöriyel araştırma tasarımlarının analizi için General Linear Model seçeneğinden yararlanabiliriz (Analyze>General Linear Model>Univariate). (Sayfa 262’de “Bölüm 28: Nümerik Veri İki Kategorik Bağımsız Değişken” başlığına ve sayfa 240’e bakabilirsiniz).

(a) Paralel



(b) Çapraz



Bölüm 17: Klinik Araştırmalar

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun klinik araştırmaların planlanması ve raporlanması hakkında bilgi sahibi olmaları amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Klinik araştırmaların (deneysel araştırmalar) nasıl yapıldığını açıklayabilmek
- *in vitro* ve *in vivo* terimlerini açıklayabilmek
- Faz 0, I, II, III, IV çalışmaları açıklayabilmek
- Klinik çalışmalarda kontrol grubunun önemini savunmak
- Pozitif ve negatif kontrol grubunu açıklayabilmek
- Sonuç ölçütlerini açıklayabilmek
- Randomizasyonu açıklayabilmek ve randomizasyon çeşitlerini sayabilmek
 - Tabakalı
 - Blok
 - Sistematik
 - Küme
- Körmeyi tanımlayabilmek ve körleme çeşitlerini açıklayabilmek
 - Tek kör
 - Çift kör
- Etik onay ve bilgilendirilmiş rızanın önemini savunmak
- Araştırmaya dahil etme ve çıkarma kriterlerini açıklayabilmek
- CONSORT statement’i açıklayabilmek
- Intention to treat analizini açıklayabilmek
- On – treatment analizini açıklayabilmek

Buraya kadar çeşitli araştırma tasarımlarını özet olarak gördük. Klinik araştırmalar, kohort çalışmaları ve vaka kontrol çalışmaları daha ayrıntılı olarak incelenecektir.

Yeni bir tedavinin insanlardaki etkisini incelemek için yapılan araştırmalara **klinik araştırmalar** denir. Klinik araştırmalardan önce laboratuvarında canlılar dışında yapılan deneylere *in vitro* (Latince: camda) araştırmalar denir. *in vitro* araştırmalar sonrasında hayvanlarda veya insanlarda yapılan deneylere ise *in vivo* (Latince: canlıda) araştırmalar denir.

Klinik araştırmalar şu şekilde sınıflandırılır:

- **Faz 0** - 10-15 kişilik grupta düşük dozda insanlarda farmakokinetik ve farmakodinamiği araştıran çalışmalar;
- **Faz I** – 20-80 kişilik grupta güvenilirlik, tolerabilite, farmakokinetik ve farmakodinamiği araştıran çalışmalar;
- **Faz II** araştırmalar – 20-300 kişilik grupta doz, etkinlik ve toksisiteyi araştıran çalışmalar
- **Faz III** – 300-3000 kişilik çok merkezli gruplarda ilacın tüm yönleriyle araştırılması. Ruhsatlandırma işlemi bu aşamada yapılır. 2008 yılı itibariyle faz II araştırmalarının ancak %18'i faz III'e ilerleyebilmektedir (https://en.wikipedia.org/wiki/Phases_of_clinical_research).
- **Faz IV** çalışmalar ise ilacın ruhsatı alınıp kullanıma konulmasından sonra yapılan, ilaca yeni bir pazar bulma, ilaç etkileşimlerini, emziren annelerde ve gebelerdeki kullanımlarını inceleyen klinik araştırmalardır.

Faz araştırmalarıyla ilgili bir özet, aşağıdaki tabloda verilmiştir:

Aşama	Amacı	Doz	Katılımcı sayısı
Klinik öncesi	Etkinlik, toksisite ve farmakokinetik bilgiler elde etmek için insan dışı deneyler	Sınırsız	-- (in vitro ve in vivo)
Faz 0	Farmakodinamik ve farmakokinetik inceleme; oral biyoyararlanım ve yarı ömrün incelenmesi	Çok küçük, tedavi eşliğinin altında	10
Faz I	Gönüllülerde yapılan deneylerle doz aralığı araştırılır; etkinlik araştırması için güvenli olup olmadığına karar verilir.	Genelde tedavi eşliğinin altında ama giderek artırılır	20-100
Faz II	Tedavi edici özelliğinden emin olunmayan ilacın ilacın etkinlik ve güvenliğini test etmek için yapılan araştırma	Tedavi dozu	100-300
Faz III	Tedavi edici özelliği kabul edilen ilacın etkinlik ve güvenliğini test etmek için yapılan araştırma	Tedavi dozu	1000-2000
Faz IV	İlaç piyasaya çıktıktan sonraki (Postmarketing) uzun süreli etkilerin araştırılması	Tedavi dozu	Hizmet alan herkes
Faz V	Toplanan veriler üzerinden translasyonel araştırma	--	Kullanım raporları

Karşılaştırmalar

Klinik araştırmalarda karşılaştırma yapmak için mutlaka bir kontrol grubu olmalıdır. Kontrol grubu olmadan bulgularımızı yorumlamamız zordur. Kontrol grubuna başka bir standart tedavi verilebileceği gibi (**pozitif kontrol**) etik sorun oluşturmaması halinde plasebo da verilebilir (**negatif kontrol**).

Araştırmanın çıktıları (sonlanım noktaları, endpoints, outcomes)

Bireylerin araştırılan ilaçtan yararlandıklarını gösteren ana ölçütün ne olduğu önceden belirlenmelidir (Cevaplanabilir klinik soru oluşturma ve PICO formatını hatırlayınız.) (örn. yeni bir tansiyon ilacı için ana çıktı veya sonuç ölçütü sistolik veya diyastolik kan basıncı olabilir). Buna **birincil sonuç ölçütü** (veya ana sonuç ölçütü) denir ve genelde tedavinin etkinliğiyle ilgilidir. Ayrıca ilacın yan etkileri, hasta memnuniyeti gibi konuları inceleyen **ikincil sonuç ölçütleri** de olabilir. Araştırma devam ederken ara ölçümler yapılarak tedavi ve kontrol grupları arasında önemli bir farkın olup olmadığına bakılabilir. Tedavi kolunda anlamlı bir üstünlük olması halinde araştırmaya devam etmek etik olmayabileceğinden araştırmanın durdurulması gerekebilir.

Gruplara ayırma

Bireyin araştırmaya katılması durumunda tedavi veya kontrol gruplarından birine ayrılması gerekecektir. Bu ayırma (allocation) işlemi bir bilgisayar veya rastgele sayılar tablosu aracılığıyla şansa dayalı (rastgele, random) yapılır. Bu işlem **randomizasyon** (randomization) olarak bilinir. Rastgele seçimden de öte, vakaların gruplara atanması ve girişim çeşidinin seçilmesi sırasında kapalı zarf gibi bir yöntem kullanılarak hekimin (veya deneyi yapanın) etkisi tamamen ortadan kaldırılmalıdır (**concealed allocation**). Böylece çalışma ve kontrol gruplarındaki bireylerin demografik özellikler açısından birbirine yakın olmaları sağlanabilir ve **ayırma hataları** (allocation bias) ve **karıştırıcılık** (confounding) olarak bilinen sorunlar önlenir.

Doğru bir randomizasyon yapıldığı halde çalışma ve kontrol grupları arasında demografik özellikler açısından fark olması halinde bu farkın rastlantısal olduğunu kabul etmeliyiz. Bu açıdan bakıldığında randomizasyon yapıldığı halde grupları aradaki fark açısından incelemenin gereksiz olduğu ortaya çıkmaktadır.

Tedavi ve kontrol gruplarına ayrılan ve randomizasyon yapılan araştırmalar **randomize kontrollü araştırmalar** (Randomized Controlled Trials - RCT) olarak bilinir ve bu tür araştırmalar ideal kabul edilir.

Tabakalı randomizasyon (deneklerimiz geldikleri coğrafi bölgelere göre tabakalara ayrılabilir ve her bölge ayrı ayrı randomize edilir) ve **bloklama** (deneklerimiz yaş, boy gibi özelliklerine göre belli aralıklarda bloklara ayrılır ve bloklarda ayrı ayrı randomizasyon yapılır) gibi yöntemler kullanılarak

randomizasyon daha da hassas hale getirilebilir. Kapıdan giriş sırası, haftanın günü veya doğum tarihi gibi özelliklere göre yapılan **sistematik gruplandırmalarda** uygulamayı yapan hekimin taraf tutma riski olabilir. Bunun yerine kapalı bir yöntemle gruplara atama (concealed allocation) tercih edilmelidir. Sistematiik gruplandırmadan mümkün olduğunca kaçınılmalıdır.

Bazen randomizasyonu **küme** düzeyinde yapabiliriz (örn. bir şehrin belli mahallelerini seçmek veya muayenehaneler içerisinde bazı muayenehaneleri seçmek gibi). Bu durumda kümeler içerisindeki bireylerin birbirine benzeme olasılığı olduğundan örneklem sayısının hesaplanmasında dikkatli olunmalıdır.

Körleme

Hastanın veya klinisyenin çalışma ve kontrol grubunun hangisi olduğunu bilmesi halinde **gözlem hataları** ortaya çıkabilir. Bu durum özellikle sonuç ölçtümüzün subjektif olması halinde söz konusudur. Örn. hasta memnuniyetini araştırıyorsak plasebo kullandığını bilen hastanın “Yeterince yararlanmadım” seçeneğini işaretleme ihtimali yüksek olacaktır. Bu nedenle mümkün olduğunda hasta, klinisyen ve ölçümü yapan kişinin gruplar hakkında bilgi sahibi olmasını önlemeliyiz. Hem hekimin, hem de hastanın grupları bilmemesi halinde **çift kör** (double blind) araştırmadan, hekimin veya hastanın durumu bilmesi halinde ise **tek kör** (single blind) araştırmadan bahsederiz. Araştırmanın yürütülmesi ve analizi sırasında da taraf tutmalar söz konusu olabileceğinden araştırmanın konuyla ilgisi olmayan birisi tarafından yürütülmesi ve analiz edilmesi halinde **üçlü kör** (triple blind) ifadesini de kullanabiliriz.

Etik

Klinik araştırmalarda etik onayı çok önemlidir. Hakemli, indekslenen tıbbi dergiler makalemizi yayınlamadan önce etik onayını mutlaka soracaklardır. Üniversiteler veya araştırma hastaneleri bünyesinde oluşturulmuş bölgesel etik kurullardan etik onay alınmalıdır. Etik onayda Helsinki Deklarasyonu (<http://www.wma.net/en/20activities/10ethics/10helsinki/index.html>; <http://www.wma.net/en/30publications/10policies/b3/17c.pdf>) dikkate alınır. Ayrıca her katılımcıdan (veya ebeveynden) **bilgilendirilmiş olur** (informed consent) alınmalıdır.

Araştırma protokolü

Araştırmaya başlamadan önce araştırmayı tüm yönleriyle anlatan ve açıklayan bir protokol hazırlanmalıdır. Burada araştırmanın amacı, yöntemi, araştırmaya alınma ölçütleri (**dahil edilme kriterleri** ve **çıkartılma kriterleri** – inclusion and exclusion criteria), tedavi planı, veri toplama ve analiz yöntemleri, araştırma ekibi ve problem çıkması halinde neler yapılacağı gibi konular belirtilmelidir.

Gerçek bir tedavi etkisini ortaya çıkarabilmek için araştırmanın örneklem sayısının yeterli olması çok önemlidir. Örneklem sayısı ve nasıl hesaplandığı da protokolda belirtilmelidir.

Hata olasılığını azaltmak için analizlerin **intention-to-treat** (tedavi niyeti) esasına göre yapılması önerilmektedir. Bu yöntemde tedaviyi uygulasa da uygulamasa da tüm katılımcıların sonuçları analize dahil edilmektedir. Araştırmadan ayrılan bireyler hakkında da veri sağlanmaya çalışılmalıdır. Sadece araştırmayı tam olarak tamamlayan bireylerin dahil edildiği **on-treatment** analiz yöntemi tedavi karşılaştırmalarında hatalara yol açabileceğinden tavsiye edilmemektedir.

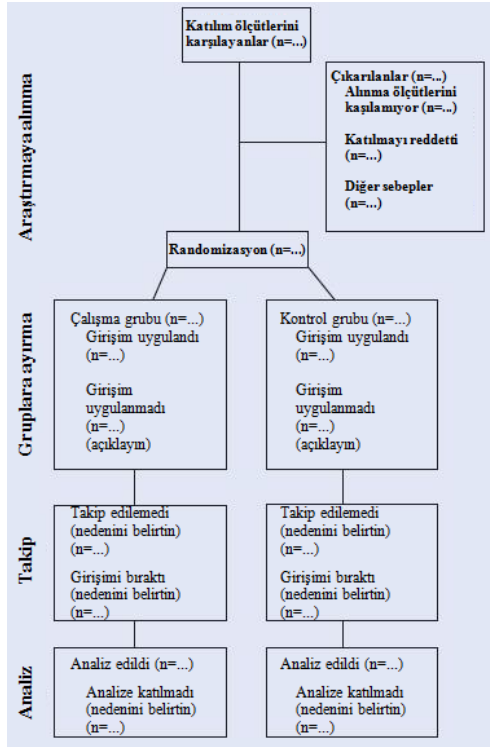
Tablo: CONSORT bildirisi (Consolidation of Standards for Reporting Trials www.consort-statement.org) kontrol listesi⁵.

MAKALEDEKİ BÖLÜM ve Başlık	Madde	Açıklama
BAŞLIK ve ÖZET	1	Gruplara ayrılma nasıl yapıldı (örn. rastgele, randomize...)
GİRİŞ		
Arka plan	2	Bilimsel arka plan ve gerekçenin açıklanması
YÖNTEM		
Katılımcılar	3	Katılımcıların alınma/dışlanma kriterleri ve araştırmanın yapıldığı ortam
Girişim	4	Her grup için yapılan girişimin ayrıntılı olarak açıklanması ve ne zaman ve nerede yapıldığı
Amaçlar	5	Araştırmanın hedefleri ve hipotezleri
Çıktılar	6	Birincil ve ikincil sonuç ölçütleri ve ölçümlerin doğruluğunu sağlamak için neler yapıldığı (birden çok ölçüm, veri toplayıcılarının eğitimi vs.)
Örneklem büyüklüğü	7	Örneklem büyüklüğü nasıl hesaplandı ve varsa ara analiz ve araştırmayı durdurma kuralları
Randomizasyon – ayırma sırası	8	Gruplara ayırma yönteminin ve sırasının açıklanması
Randomizasyon –	9	Gruplar kesinleşene kadar ayırma işleminin gizli tutulması

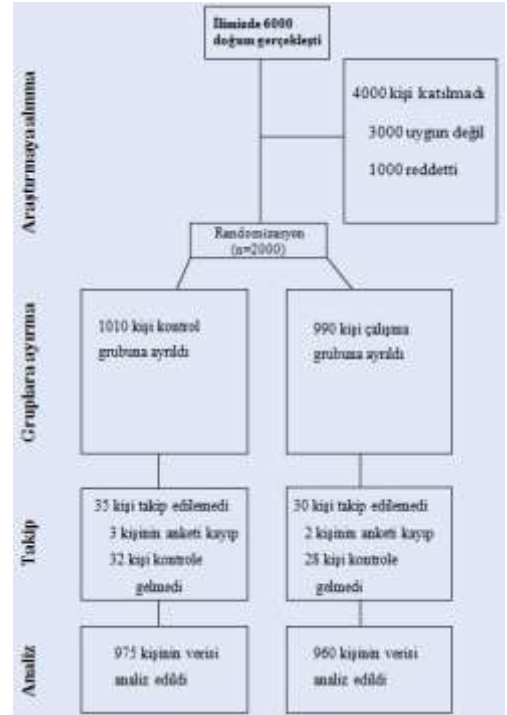
⁵ Didem Sunay, Tijen Şengezer, Meltem Oral, Zekeriya Aktürk, Kenneth F. Schulz, Douglas G. Altman, David Moher, for the CONSORT Group. CONSORT 2010 Raporu: Randomize Paralel Grup Çalışmalarının Raporlanmasında Güncellenmiş Kılavuzlar. Eurasian Journal of Family Medicine. 2013;2(1):1-10 <http://ejfm.trakya.edu.tr/archive/show/volume2/number1/263102>

ayırma gizliliği		
Randomizasyon – uygulama	10	Gruplara ayırma sırasını kim belirledi, katılımcıları çalışmaya kim aldı ve kimin hangi gruba gireceğini kim belirledi?
Körleme	11	Katılımcılar, girişimi uygulayanlar ve sonuçları değerlendirenler kimin hangi grupta olduğunu biliyor muydu? Körlemeden nasıl emin olundu?
İstatistik yöntemler	12	Grupları ana sonuç ölçütü açısından karşılaştıran istatistik yöntemler. Alt grup analizleri ve ağırlıklandırılmış analizler gibi ek analiz yöntemleri
BULGULAR		
Katılımcı akışı	13	Her aşama için katılımcı akışı (Bakınız Şekil). Protokolden sapmaların gerekçeleriyle birlikte açıklanması.
Araştırmaya alınma	14	Araştırmaya alınma ve takip tarihleri.
Başlangıç verileri	15	Her grup için başlangıçtaki demografik ve klinik özellikler.
Analiz edilenler	16	Gruplarda analize katılan birey sayısı ve “intention-to-treat” uygulanıp uygulanmadığı. Sonuçları olabildiğince sayılarla verin (%50 yerine 10/20)
Çıktı ve tahminler	17	Gruplarda birincil ve ikincil sonuç ölçütleri için sonuçların özetlenmesi ve etki büyüklüğü (effect size) ve hassasiyetin (%95 GA) belirtilmesi
Ek analizler	18	Alt grup analizleri, ağırlıklandırılmış analizler gibi yapılan diğer analizlerin belirtilmesi
İstenmeyen etkiler	19	Her grup için bütün önemli istenmeyen etkiler ve yan etkiler belirtilmelidir.
TARTIŞMA	20	Araştırma hipotezlerini, olası hata kaynaklarını, hassasiyeti vs. dikkate alarak bulguların yorumlanması.
Yaygınlaştırılabilirlik	21	Araştırma bulgularının başka ortamlarda uygulanabilirliği (yaygınlaştırılabilirliği, external validity)
Genel kanıt	22	Güncel kanıtlar ışığında sonuçların genel olarak yorumlanması.

Şekil: Randomize kontrollü bir araştırma için CONSORT bildirisi taslağı.



Şekil: Randomize kontrollü bir araştırma için CONSORT bildirisi örneği.



Bölüm 18: Önceki Konuların Özeti

Amaç: Bu oturum sonunda katılımcıların istatistik kursunun ilk 16 konusunda öğrenilenleri tekrar etmeleri ve pekiştirmeleri amaçlanmıştır.

Yönerge: aşağıdaki soruları önce kaynak kullanmadan, sonra da notlarınızdan yararlanarak cevaplandırınız. Kaynak kullanmadan %80'ini, kaynak kullanarak %100'ünü cevaplandırabilmelisiniz.

1.	- Bir araştırmada erkeklerle kadınlar arasındaki antihipertansif tedaviye uyum durumu karşılaştırılmak isteniyor. - Hangi istatistik hipotez test/testler yapılabilir?
2.	- Bir araştırmada migrenli hastalarda ilaç verilmeden önce ve sonra baş ağrısının şiddeti (yedili görsel analog skala kullanılarak) araştırılmak isteniyor. - Hangi istatistik hipotez test/testler yapılabilir?
3.	- Bir araştırmada çeşitli sınıflardaki tıp öğrencilerinde girdikleri son sınavda geçer not alıp almadıkları karşılaştırılmak isteniyor. - Hangi istatistik hipotez test/testler yapılabilir?
4.	- Bir araştırmacı mahallesindeki erişkinlerin sistolik kan basınçlarını ölçmek ve toplum ortalamasıyla karşılaştırmak istiyor. - Hangi istatistik hipotez test/testleri yapılabilir?
5.	- Üç ayrı antihipertansif ilaç kullanan hasta grubunda kandaki pentraksin düzeylerini karşılaştırmak istiyorsunuz. - Hangi istatistik hipotez test/testleri yapılabilir?
6.	Veri çeşitlerini sınıflandırınız
7.	- Aşağıdaki terimleri örnek vererek açıklayınız: - Yüzde - Oran - Orantı - Hız - Puan
8.	- Değişken 1: Hasta memnuniyeti: 1-hiç memnun değil, 2-az memnun, 3-orta halli,

	4-memnun, 5-çok memnun • ne tür bir değişkendir?
9.	• Değişken 2: Öğrenim durumu: 1-ilkokul, 2-ortaokul, 3-lise, 4-üniversite, 5-doktora • ne tür bir değişkendir?
10.	• 8 ve 9. sorulardaki değişkenlerin ortalama ve standart sapmaları hesaplanabilir mi? • Neden?
11.	• SPSS hangi yazılım veya standartlardan veri alabilir (import) ve hangi formatlarda veri ihraç edebilir (export) ?
12.	• <u>Veri toplama aşamasında</u> değişkenlerimize SPSS'te kullanacağımız isimlerin verilmesi neden önemlidir? • SPSS'te değişkenler için kategorilerin (values) tanımlanması neden önemlidir? Bu açıdan veri toplama aşamasında nasıl bir yöntem izlenmelidir?
13.	• Değişkenimizde birden fazla seçeneğin olması durumunda SPSS'te hangi yöntemler izlenebilir? • Hatırlatma: Az sayıda seçenek vardır ve bunlardan çoğunun seçilmesi söz konusu olabilir Çok sayıda seçenek vardır ama bunlardan sadece birkaçının seçilmesi söz konusudur
14.	• Aşağıdaki anket sorusu hakkında ne söyleyebilirsiniz? • Kaç yaşındasınız? a) 20'den az, b) 20-40, c) 41-60 d) 60'dan fazla

15	Araştırmadaki bireylere numara verilmesi ve SPSS’te de bu numaralar için bir değişken tanımlanması neden önemlidir?										
16	- Veri analizine başlamadan hata ayıklaması yapmak neden önemlidir? - SPSS’te hata ayıklamasını hangi yöntemlerle yapabiliriz?										
17	SPSS’te uç değerler ve aşırı değerleri nasıl saptarsınız?										
18	SPSS ile diyabet.sav veri setinde “duration” değişkeninin histogram grafiğini çizin ve yorumlayınız.										
19	- Aşağıdaki ne tür bir grafiktir. SPSS ile diyabet.sav veri setinde bu grafiği yapınız? Marital status - Single - Married - Divorced - Widow Pies show counts <table border="1"> <thead> <tr> <th>Marital status</th> <th>Percentage</th> </tr> </thead> <tbody> <tr> <td>Single</td> <td>5.71%</td> </tr> <tr> <td>Married</td> <td>79.76%</td> </tr> <tr> <td>Divorced</td> <td>4.76%</td> </tr> <tr> <td>Widow</td> <td>9.76%</td> </tr> </tbody> </table>	Marital status	Percentage	Single	5.71%	Married	79.76%	Divorced	4.76%	Widow	9.76%
Marital status	Percentage										
Single	5.71%										
Married	79.76%										
Divorced	4.76%										
Widow	9.76%										
20	Merkezi eğilim ölçütleri nelerdir?										
21	- Geometrik ortalama nedir? - Antilog nasıl hesaplanır?										
22	Ağırlıklı ortalamayı açıklayınız?										
23	Yaygınlık ölçütleri nelerdir?										
24	Referans aralığı nedir? Nasıl hesaplanır?										
25	Standart sapmanın formülünü belirtiniz										

26	• Bir istatistiksel hipotez testinde yanılma ihtimalimiz %5'tir. • Araştırmamızda üç adet hipotez testi yaptığımızı kabul edelim. • Bu araştırmada hipotez testlerinden en az birinin hatalı olma olasılığı nedir?
27	• Normal dağılımı (Gauss dağılımı) tanımlayan parametreler hangileridir?
28	• Normal dağılımında verilerin – %68'i – %95'i – %99'u • ortalama $\pm$ kaç standart sapma aralığındadır?
29	• Normal Sapma (Standardized Normal Deviate) formülünü belirtiniz (z formülü)
30	• Bir durum için sadece başarı ve başarısızlık şeklinde iki sonuç olabileceğini varsayalım. Bu veri hangi teorik dağılıma uyar?
31	• Sağa doğru eğimli (pozitif) verileri normal dağılıma yaklaştırmak için hangi veri dönüştürme yöntemleri kullanılabilir?
32	• Ortalamanın standart hatası (SEM) nasıl hesaplanır? • SEM bize ne hakkında bilgi verir?
33	• Bir kategorik değişken için ortalamanın standart hatası nasıl hesaplanır?
34	• Ortalamanın güven aralığı nasıl hesaplanır?
35	• Kategorik veriler için ortalamanın güven aralığı nasıl hesaplanır?
36	• Nümerik veriler için ortalamalar arasındaki farkın güven aralığı nasıl hesaplanır?
37	• Kategorik veriler için ortalamalar arasındaki farkın güven aralığı nasıl hesaplanır?
38	• Ortalamalar arasındaki farkın güven aralığı nasıl yorumlanır?
39	• Odds oranı nedir? • Odds oranı için güven aralığı nasıl yorumlanır?
40	• Klinik olarak cevaplanabilir bir soru sormak için kullanılan PICO akronimini örnek vererek açıklayınız?
41	• Araştırmalarımızda nedenselliği irdelemek için Hill tarafından önerilen kriterler nelerdir?

42	• Araştırmalarda ne tür hatalar olabilir?
43	• Confounding (karıştırmacılık) nedir? Örnek veriniz?
44	• Temel araştırma tasarımlarını sayınız.
45	• Varyasyonu azaltmak için ne tür araştırma tasarımları yapabiliriz?
46	• Çapraz araştırma tasarımını açıklayınız.
47	• Faktöriyel araştırma tasarımını açıklayınız.
48	• Intention to treat analizini açıklayınız? Bu yöntemin aksi nedir?
49	• CONSORT bildirisi nedir?
50	• Concealed allocation (ayırma gizliliği) nedir? • Ayırma gizliliği neden önemlidir?

Aıştırma 7

Aşağıdaki bağlantıdaki “Re-engineering systems for the treatment of depression in primary care: cluster randomised controlled trial” başlıklı makaleyi indiriniz ve sorulara cevap veriniz:

<http://www.bmj.com/cgi/reprint/bmj.38219.481250.55v1.pdf>

<http://www.bmj.com/content/329/7466/602.pdf%2Bhtml>

Sorular:

1. Araştırmanın başlığında "cluster randomized controlled trial" olduğu belirtilmiş. Buradaki kümeler nelerdir? Küme örneklem olmasının örneklem sayısı açısından ne gibi olumsuzluğu olur?
2. Araştırmanın “inclusion” (dahil edilme) ve “exclusion” (çıkarılma) kriterleri nelerdir?
3. Araştırmanın hasta akışı CONSORT şemasına uygun olarak verilmiş midir? İnceleyiniz.
4. Araştırmanın analizinde "intention to treat" yöntemi kullanılmış mıdır? Bu yöntemin araştırmaya nasıl etkisi olduğunu yorumlayabilir misiniz?
5. Araştırmada örneklem hesabıyla ilgili bilgi verilmiştir. İlgili bölümü bulup örneklem hesabında daha önceki araştırmaların sonuçlarının ve klinisyen açısından önemlilik derecesinin veriliş verilmediğini kontrol ediniz. Bulduklarınızı yazınız.
6. Araştırmada nasıl bir körleme kullanılmıştır?
7. Araştırmanın ana sonuç ölçütü (primary endpoint) nedir? (örneklem hesabında da bu sonuç ölçütünün kullanıldığına dikkat ediniz)
8. Araştırmanın sunumunda CONSORT önerilerinden dikkate alınmayanlar var mıdır? (ya da araştırmada sizin saptayabildiğiniz eksiklikler var mıdır?)

Çözüm için bakınız s. 473

Bölüm 19: Kohort Araştırmaları

Hamit ACEMOĞLU

Amaç: Bu konu sonunda okuyucunun kohort türü araştırma tasarımı hakkında bilgi sahibi olmaları amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Kohort çalışmaların yapılış yöntemini açıklayabilmek
- Tarihsel (historical) kohortu açıklayabilmek
- “Sağlıklı birey etkisi” ni açıklayabilmek
- Takipte kaybolan birey olmamasının önemini savunmak
- Rölatif riski hesaplayabilmek
- Rölatif riskin hastalığın toplumdaki prevalansı ile ilişkisini açıklayabilmek
- Kohort çalışmaların avantaj ve dezavantajlarını sayabilmek
- Klinik kohortu tanımlayabilmek

Kohort araştırmalarda bireyler zaman içerisinde ileriye doğru takip edilir. Esas amacı belli **risk faktörlerinin gelecekte** hastalığa yol açıp açmayacağını araştırmasıdır. Kohort araştırmalar genelde prospektif olmakla birlikte, verilere ulaşılabilme durumunda geçmişten başlayıp bugünkü duruma gelen kohort araştırmalar da olabilir (**tarihsel kohort** veya **retrospektif kohort**).

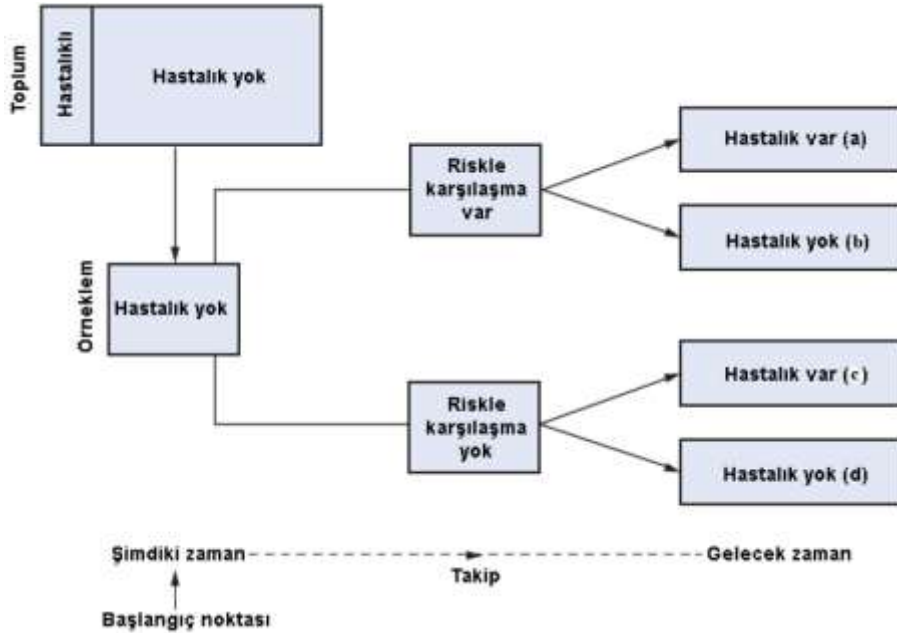
Sabit (fixed) kohortlarda zaman içerisinde araştırmayı terk eden bireylerin yerine yenisi alınmaz. **Dinamik** kohortlarda ise bireyler araştırmadan ayrılabilir ve araştırmaya alınma ölçütlerini karşılayan diğer bireyler dahil edilir.

Kohortun seçilmesi

Seçilen kohort toplumu temsil edici olmalıdır. Kohortumuz bir kasaba, bir fabrikada çalışanlar, bir meslek grubu (örn. doktorlar) veya takip ettiğimiz bir klinik grup (örn. polikliniğimize gelen hipertansif hastalar) olabilir.

Nedensellik bağlantısını kurabilmemiz için risk faktörünün hastalıktan önce gelmesi gerektiğinden kohort çalışmasına alınma aşamasında **bireylerin sağlıklı olması gerekir**. Bu durum, araştırmanın ilk zamanlarında hastalığın prevalansının toplumdakinden düşük olmasına yol açar. Buna **sağlıklı birey etkisi** (healthy entrant effect) denir ve dikkat edilmesi gereken bir hatadır. Diğer taraftan araştırmaya katılmaya gönüllü olan insanlar diğerlerine göre sağlığına daha fazla önem verir.

Kohort çalışmalar uzun sürdüğünden zaman içerisinde bireylerin araştırmayı terk etme (**drop out**) ihtimali vardır. Araştırma süresi sonunda kayıp birey sayısının fazla olması araştırmanın değerini düşürecektir. Bu durum kohort



Şekil: Bir kohort araştırma tasarımı (parantez içerisindeki frekansların açıklaması için aşağıdaki tabloya bakınız).

Kohort çalışmalarında dikkat edilmesi gereken bir konu da incelenen riskin zaman içerisinde aynı olmamasıdır. Sigara kullanımıyla kalp hastalığı arasındaki risk ilişkisini araştırdığımızı düşünelim. Bireyin başlangıçtaki sigara kullanım durumu değişebilir. Daha fazla içmeye başlayabileceği gibi bırakabilir de. Bu nedenle kohort araştırmalarında denekler sık aralıklarla yeniden değerlendirmeli ve risk durumu incelenmelidir.

Kohort araştırmaların analizi

Tablo: Kohort araştırma sonuçlarının özetlenmesi.

		İkinci gözlem		Toplam
		Hastalık var	Hastalık yok	
Birinci Seçim	Maruziyet var	a	b	a+b
	Maruziyet yok	c	d	c+d

	Toplam	a+c	b+d	a+b+c+d=n

“ $[a/(a + b)] - [c/(c + d)]$ ” formülü kullanılarak hastalığın ne kadarının risk faktörüne atfedilebileceği hesaplanabilir (**atfedilen risk**).

* DeneySEL çalışmalar, Kohort araştırmaları ve vaka kontrol çalışmalarından elde edilebilecek bazı göstergelerin örnekleriyle hesaplanmasını konu sonundaki tabloda bulabilirsiniz.

Kohort araştırmalarda bireyler zaman içerisinde takip edildiklerinden hastalık gelişme riskini hesaplamak mümkündür.

Risk hesabı (insidans) = çalışma boyunca hastalık gelişenler / kohorttaki toplam birey sayısı = **$a+c/n$**

Aynı şekilde riskle karşılaşanlar ve karşılaşmayanlardaki hastalık gelişme olasılığı da hesaplanabilir.

Riskle karşılaşanlarda hastalığın gelişme olasılığı: **risk₊** = $a/a+b$

Riskle karşılaşmayanlarda hastalığın gelişme olasılığı: **risk₋** = $c/c+d$

Son olarak ta nispi risk (relative risk; RR) hesaplanır:

$$RR = \text{risk}_+ / \text{risk}_- = [a/(a+b)] / [c/(c+d)]$$

Rölatif risk, riskle karşılaşanlarda hastalık riskinin arttığını veya azaldığını gösterir. RR = 1 olması durumunda risk faktörüyle karşılaşan ve karşılaşmayanlarda hastalık gelişme durumunun eşit olduğu anlaşılır.

Sigara içmeyi bir risk faktörü olarak alıp bir kohort çalışması planlayalım. Sonuç ölçtümüz kalp krizi geçirmek olsun. RR = 2 olması durumunda sigara içenlerin içmeyenlere göre kalp krizi geçirmek açısından 2 kat daha fazla risk altında oldukları anlaşılır. RR 1’den küçük olması durumunda ise incelenen risk faktörünün hastalıktan koruyucu bir etki yaptığını anlamalıyız.

Rölatif risk her zaman hastalığın görülme sıklığıyla birlikte değerlendirilmelidir. Hastalığın çok nadir olması durumunda rölatif risk dikkatle değerlendirilmelidir. Örn:

Bir ilaç firmasının geliştirdiği grip aşısının etkisini araştırmak için bir kohort çalışması planladığını düşünelim (risk faktörü = grip aşısı yapılmak). Sonuç ölçütleri sağlıklı erişkinlerde gripten ölüm olsun. Araştırma sonunda aşıya maruz kalanlarda RR = 0,8 bulunsun. Grip ve pnömoniden ölüm oranlarının 100 binde 18 civarında (http://www.cdc.gov/nchs/data/nvsr/nvsr56/nvsr56_16.pdf) olduğunu

biliyoruz. Bu durumda 0.8'lik rölatif riskin gerçekte ne kadar fayda sağlayacağı ve kar/zarar oranları tartışılabilir.

Rölatif risk için de güven aralıkları hesaplanmalı ve güven aralıklarının 1'i içerip içermediğine bakılmalıdır. **1'i içermesi** riskin eşit olduğu, yani arada fark olmadığı anlamına gelir.

Kohort araştırmaların olumlu yönleri

- Olayların zamanla ilişkisi değerlendirilebilir.
- Birden fazla sonuç ölçütünü bir arada değerlendirebiliriz.
- Hastalığın insidansını hesaplayabiliriz.
- Çok sayıda risk faktörü ile ilgili ayrıntılı bilgi toplayabiliriz.
- Nadir risk faktörlerini de araştırabiliriz.
- Riske maruziyet zaman içerisinde tekrar ölçülebilir.
- Vaka kontrol çalışmalarına göre **hatırlama hataları** ve **seçime bağlı hatalar** daha azdır.

Kohort araştırmaların olumsuz yönleri

- Genelde uzun sürdüğünden masraflı ve zahmetlidir.
- Sonuç ölçütünün nadir olması durumunda çok büyük bir kohort gerekir.
- Takip uzadıkça araştırmadan ayrılanlar da artacak, bu da hatalara yol açacaktır.
- Zamanın uzun olması nedeniyle ölçümlerde tutarlılığı sağlamak zor olacaktır.
- Bireyler başlangıçtaki riskli davranışlarını değiştirebilirler.
- Hastalığın sıklığı ve etiyolojisinin de zaman içinde değişmesi söz konusu olabilir.

Klinik kohortlar

Bazen kliniğimizde takip ettiğimiz belli gruptan hastaları bir risk faktörü için ileriye dönük izlemek isteriz. Rutin sağlık hizmeti sunumu sırasında kaydettiğimiz verileri böylece bir kohort araştırmasına dönüştürebiliriz. Klinik kohortlarda amaç, hastalığa sahip bireylerin gidişatını takip etmek veya farklı tedavi yaklaşımlarını karşılaştırmak olabilir.

Randomize kontrollü çalışmaların aksine, klinik kohortlarda çalışmaya alınma kriterleri uygulamayız ve takip ettiğimiz tüm hastaları dahil ederiz. Bu nedenle klinik kohortlar gerçek durumu daha iyi yansıtabilir. Ancak, diğer taraftan tedavi grupları konusunda bir randomizasyon söz konusu olmadığından karıştırıcı faktörlerden etkilenme olasılığı klinik kohortlarda daha fazladır. Örn:

Klinik kohortumuz kliniğimizde takip ettiğimiz diyabet hastalarından oluşabilir. Bu hastalarımızın bazılarını oral antidiyabetik ilaçlar, bazılarını da insülin verdiğimizizi düşünelim (risk faktörü = hastanın kullandığı ilaç türü). Sonuç ölçütümüz diyabetik nöropati gelişme durumu olsun. Beş yıl sonra her iki gruptaki komplikasyon gelişme oranı hesaplanıp birbirine bölünerek rölatif riski bulabiliriz. Direkt klinikteki uygulamamızı yansıtacağı için yararlı bir çalışma olmakla birlikte, insülin ve oral antidiyabetik gruplarını rastgele seçmediğimiz ve klinik tercihlerimize göre yaptığımızdan karıştırmacılık söz konusu olabilir ve genelleme yapmamızı zorlaştırır. Örn. insülin tedavisi verdiğimiz grup daha yaşlı bireylerden oluşabilir. Yaşlılarda komplikasyonların görülme olasılığı daha fazla olacağından iki gruptaki hastalarımızın yaşa göre rastgele dağılmamış olması karıştırmacılığa yol açacaktır.

Örnek 1: risk azalması			Örnek 2: risk artışı		
	Risk pozitif (R)	Kontrol (K)	Toplam	(R)	(K)
Hastalık var (H)	RH = 15	KH = 100	115	RH = 75	KH = 100
Hastalık yok (S)	RS = 135	KS = 15	285	RS = 75	KS = 150
Toplam birey (T)	RT = RH + RS = 150	KT = KH + KS = 250	400	RT = 150	KT = 250
Olay yüzdesi (OY)	ROY = RH / RT = 0,1 (%10)	KOY = KH / KT = 0.4 (%40)		ROY = 0.5 (%50)	KOY = 0.4 (%40)
Eşitlik	Değişken		Kısaltma	Örnek 1	Örnek 2
RO – KOY	< 0: toplam risk azalması (<i>absolute risk reduction</i>)		TRAz (ARR)	(–) 0,3; (–%30)	
	> 0: toplam risk artışı (<i>absolute risk increase</i>)		TRArt (ARI)		0,1; (% 10)
(ROY – KOY) / KOY	< 0: rölatif risk azalması (<i>relative risk reduction</i>)		RRAz (RRR)	(–) 0,75; (–%75)	
	> 0: rölatif risk artışı (<i>relative risk increase</i>)		RRArt (RRI)		0,25; (%25)
1 / (ROY – KOY)	< 0: tedavi için gereken kişi (<i>number needed to treat</i>)		TGK (NNT)	(–) 3,33	
	> 0: zarar için gereken kişi (<i>number needed to harm</i>)		ZGK (NNH)		10
ROY / KOY	Rölatif risk (<i>relative risk</i>)		RR (RR)	0,25	1,25
(RH / RS) / (KH / KS)	Odds oranı (<i>odds ratio</i>)		OR (OR)	0,167	1,5
RH / (RH + KH) – RS / (RS + KS)	Atfedilen risk (<i>attributable risk</i>)		AR (AR)	(–) 0,34; (–%34)	0,095; (%9,5)
(RR – 1) / RR	Atfedilen risk yüzdesi (<i>attributable risk percent</i>)		ARY (ARP)		%20
1 – RR (veya 1 – OR)	Koruyucu kısım (<i>preventive fraction</i>)		KK (PF)	0,75, veya %75	

Tablo: Deneysel çalışmalar, Kohort araştırmaları ve vaka kontrol çalışmalarından elde edilebilecek bazı göstergelerin örnekleriyle hesaplanması.⁶

⁶ https://en.wikipedia.org/wiki/Relative_risk_reduction

Bölüm 20: Vaka Kontrol Çalışmaları

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun vaka kontrol çalışmaları hakkında bilgi sahibi olmaları amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Vaka kontrol çalışma tasarımını açıklayabilmek
- Vaka seçiminde “incident” ve “prevalent” terimlerini açıklayabilmek
- Kontrol grubunun nasıl seçileceğini açıklayabilmek
- Vaka kontrol çalışmalarında eşleştirmeyi açıklayabilmek
- Odds oranının nasıl hesaplanacağını açıklayabilmek
- Odds oranının nasıl yorumlanacağını açıklayabilmek
- Vaka kontrol çalışmalarının olumlu ve olumsuz yönlerini sayabilmek

Vaka kontrol çalışmaları belli özelliklere sahip hastalıklı bireyleri (**vakalar**) aynı özellikler açısından hasta olmayan bireylerle (**kontroller**) karşılaştırır. Böylece incelenen özelliklerin vaka veya kontrol grubunda daha fazla olup olmadığı araştırılır. Bu tür retrospektif araştırmalar hastalığın insidansı veya prevalansı hakkında bilgi vermezler. Hastalığın riskini hangi faktörlerin artırdığı veya azalttığı hakkında bilgi verirler.

Vakaların seçilmesi

Vakalar tanı aldıkları anda araştırmaya dahil edilebilecekleri gibi (**incident cases**) daha önceden tanı konmuş vakalar da araştırmaya dahil edilebilir (**prevalent cases**). Bu durumun belirtilmesinde yarar vardır. Zira, özellikle popüler hastalıklarda tanı aldıktan sonra hastanın hastalığı ve risk faktörleri üzerinde düşünme ve tutum değiştirme ihtimali vardır.

Örn. akciğer kanseri tanısı konan hastalarda sigaranın zararlı etkilerini bilme durumunu bir risk faktörü olarak incelemek istediğimizi düşünelim. Hasta kanser olduğunu öğrendikten sonra büyük ihtimalle sigara ve zararlı etkileri hakkında daha fazla şey duyacak veya araştırıp öğrenecektir. Bu durumda yeni tanı konmuş hasta bize daha geçerli ve güvenilir bilgi verecektir.

Araştırmamıza mümkün olduğunca fazla vaka almalıyız. Böylece yorumlarımızı genellememiz mümkün olur. Gerekirse hastane kayıtları incelenerek araştırma dönemindeki vakaların tümüne ulaşmaya çalışılmalıdır.

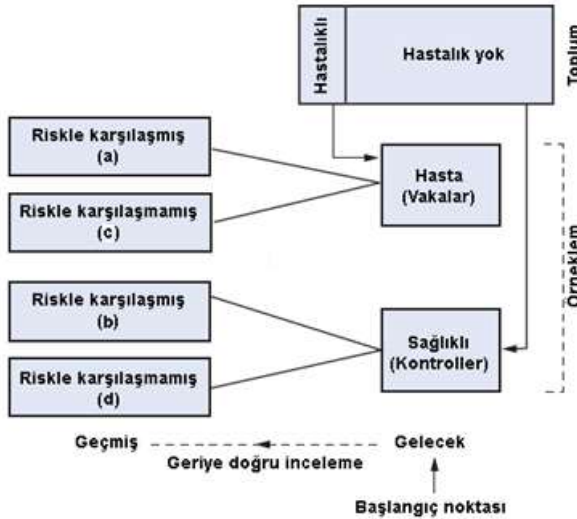
Aralık 2008’de Edirne’de miyokart enfarktüsünden ölen bireylerde sigara içme, diyabet varlığı ve hiperlipidemi gibi risk faktörlerini

araştırdığımızı düşünelim. Sadece fakülte hastanesi kayıtlarını inceleyip enfarktüstten ölenlerin dosyalarını çıkarmak ve risk faktörlerine bakmak toplumu yansıtmaya açısından yeterli olmayacaktır. Ölümler genelde hastanelerde olacağından araştırmaya tüm hastaneleri dahil etmeliyiz. Hatta bununla da yetinmeyip belediyenin ölüm kayıtlarından Aralık 2008’de ölenlerin listesini almalı ve ölüm nedenlerini bulmaya çalışmalıyız.

Kontrol grubunun seçilmesi

Kontrol grubundaki bireyler muayeneden geçirilmeli ve söz konusu hastalığa sahip olmadıkları kesinleştirilmelidir. Bazen bir vaka için birden fazla kontrol alınabilir. Vaka başına kontrol sayısını dördün üzerine çıkarmanın maliyet etkin olmadığı bildirilmektedir.

Kontroller olabildiğince hastalarla aynı ortamdaki seçilmelidir. Bireylere ulaşma kolaylığı nedeniyle kontroller genelde hastanelerden seçilmektedir. Halbuki, hastanedeki bireylerin söz konusu risk faktörüne maruz kalmış olma olasılığı toplumdakinden daha fazladır. Bu nedenle bu yöntem pek önerilmemektedir. Diğer taraftan toplumdaki bireyler kontrol grubu olma konusunda istekli olmayabilir. Kontrollerin vakalarla aynı mahalleden seçilmesi sosyal yapılarının aynı olmasını sağlayabilir.



Eşleştirme

Bazı vaka kontrol araştırmalarında birbirine olabildiğince benzeyen vakalar ve kontroller seçilerek eşleştirme yapılır. Genel olarak cinsiyete ve yaşa göre eşleştirme önerilmektedir. Bununla birlikte, incelenmek istenen risk faktörü açısından veya hastalığın nedenini oluşturabilecek faktörler açısından eşleştirme

yapmamak önemlidir. Eşleştirilen faktör ile hastalık arasındaki ilişkiyi incelemek mümkün değildir.

Eşleştirilmiş vaka kontrol çalışmalarında vakalarla kontroller birbirinden bağımlı olacağından analiz sırasında dikkatli olunmalıdır.

Eşleştirilmemiş vaka kontrol araştırmalarının analizi

Vaka kontrol araştırmalarında gözlenen frekansların tablosu aşağıda verilmiştir. Vakalar toplumdan seçilmeyip hastalık durumlarına göre seçildiklerinden hastalığın kesin riskini hesaplamak mümkün değildir. Bu durumda **odds oranı** (<http://www.merriam-webster.com/dictionary/odd>: differing markedly from the usual or ordinary or accepted – normalden, alışılmıştan veya beklenenden farklı olan) hesaplanır.

Odds Oranı = Vakalarda (hasta olanlarda) riskle karşılaşma odds'u / Kontrol grubunda (hasta olmayanlarda) riskle karşılaşma odds'u

Vakalarda riskle karşılaşma odds'u şöyle hesaplanır:

Vakalarda riskle karşılaşma ihtimali / Vakalarda riskle karşılaşmama ihtimali

Kontrol grubunda riskle karşılaşma odds'u ise şöyle hesaplanır:

Kontrol grubunda riskle karşılaşma ihtimali / Kontrol grubunda riskle karşılaşmama ihtimali

Tablodaki frekanslar üzerinden yapacak olursak:

$$Odds_{vaka} = [a/(a+c) / c/(a+c)] = a/c$$

$$Odds_{kontrol} = [b/(b+d) / d/(b+d)] = b/d$$

Bu durumda

$$Odds Oranı (OR) = (a/c) / (b/d) \text{ yani } ad/bc$$

olur.

Tablo: Vaka kontrol araştırmasında gözlenen frekanslar.

		Hastalık		Toplam
		Vaka	Kontrol	
Risk faktörü	Var	a	b	a+b
	Yok	c	d	c+d
	Toplam	a+c	b+d	$n = a+b+c+d$

Bir hastalığın nadir olması durumunda odds oranı ile rölatif risk hesabı birbirine yakın değerler verecektir. Örn:

Hastalık prevalansının 100/2000 = %5 olduğu bir çalışma için:

		Hastalık		
		Vaka	Kontrol	Toplam
Risk faktörü	Var	25	95	120
	Yok	75	1805	1880
Toplam		100	1900	2000

Relatif Risk (RR = nisbi risk) = $(25/120) / (75/1880) = 5,2$

Odds Oranı (OR) = $(25/75) / (95/1805) = 6,3$

Aynı hesabı hastalık prevalansının 10/2000 yani %0,5 olduğunu kabul ederek yaparsak:

		Hastalık		
		Vaka	Kontrol	Toplam
Risk	Var	2	118	120
	Yok	8	1872	1880
Toplam		10	1990	2000

RR = $(2/120) / (8/1880) = 3,92$ **OR** = $(2/8) / (118/1872) = 3,97$

Odds oranının 1 olması durumunda riskle karşılaşan ve karşılaşmayan gruplarda hastalığın gelişmesi açısından bir fark olmadığı anlaşılır. Odds oranının yüksek olması riskin fazlalığını, 1'den küçük olması ise incelenen risk faktörünün hastalığa karşı koruyucu olduğunu gösterir. Odds oranı için de güven aralıkları hesaplanabilir.

Vaka kontrol çalışmalarının olumlu yönleri

- Çabuk ve kolay yapılabilir ve ucuzdur
- Özellikle nadir hastalıklar için uygundur
- Çok sayıda risk faktörü araştırılabilir
- Vakaların takipte kaybolması (drop out, loss at follow up) söz konusu değildir.

Vaka kontrol çalışmalarının olumsuz yönleri

- Hatırlama hataları olabilir. Bir akciğer kanseri hastası ne kadar sigara içtiğini hatırlayabilir ama kontrol grubundaki bireyin sigara içme durumunu hatırlama olasılığı daha azdır.
- Hastalığın başlangıcı risk faktöründen öncesine dayanıyorsa nedensellik bağı kurulamaz.
- Risk faktörleriyle karşılaşmanın nadir olduğu durumlarda vaka kontrol çalışmaları uygun değildir.

Aıştırma 8

Araştırma 1: Bir vaka kontrol çalışmasının sonuçları.

		Hastalık		
		Var	Yok	Toplam
Risk	Var	100	8000	8100
	Yok	10	15000	15010
	Toplam	110	23000	23110

Araştırma 2: Bir kohort çalışmasının sonuçları

		Hastalık		
		Var	Yok	Toplam
Risk	Var	600	250	850
	Yok	400	1750	2150
	Toplam	1000	2000	3000

1. Araştırma 1 için odds oranını hesaplayınız
2. Araştırma 2 için rölatif (nisbi) riski hesaplayınız.
3. Araştırma türleri belirtilmemiş olsaydı bu araştırma türlerini tablolardan tahmin etmek için nasıl bir mantık yürütürdünüz?

Çözüm için bakınız s. 474

Bölüm 21: Hipotez Testleri

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun hipotez testlerinin oluşturulması ve yorumlanması hakkında bilgi sahibi olması amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Sıfır hipotezini oluşturabilmek
- Alternatif hipotezi oluşturabilmek
- “İki yönlü hipotez” ve “tek yönlü hipotez” kavramlarını açıklayabilmek
- Hipotez testlerinde test istatistiklerinin yerini açıklayabilmek
- Hipotez testlerinde p değerinin yerini açıklayabilmek
- p değerinin nasıl yorumlanacağını açıklayabilmek
- Parametrik / Nonparametrik testleri açıklayabilmek
- Hipotez testleri ile güven aralığı ilişkisini açıklayabilmek
- Equivalence ve non-inferiority (eşdeğerlik) terimlerini açıklayabilmek

Hipotez testlerinde genel yaklaşım

Hipotez testleri, araştırmamızın başında sorduğumuz sorunun ve kurduğumuz hipotezin kabulü veya reddi için ne kadar kanıt olduğunu anlamamıza yardım eder. Hipotez testlerinde kural olarak aşağıdaki 5 aşama izlenir:

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması
2. Verilerin toplanması
3. İlgili sıfır hipotezi için *test istatistiğinin* hesaplanması
4. Test istatistiğinden elde edilen değer bilinen bir olasılık dağılımı ile karşılaştırılması
5. P değerinin ve sonuçların yorumlanması

H_0 ve H_1 hipotezi

Sıfır hipotezi, “arada fark yoktur” şeklinde kurulur. Örn.

Toplumda kadınlarla erkeklerin sigara içme durumlarını karşılaştırmak istersek kuracağımız H_0 hipotezi şöyle olur:

“Toplumda kadınlarla erkeklerin sigara içme sıklıkları arasında fark yoktur”

Alternatif hipotez ise H_0 ’ın doğru olmadığını iddia eder. Yukarıdaki örnekte H_1 şöyle olur:

“Toplumda kadınlarla erkeklerin sigara içme sıklıkları farklıdır”

Dikkat edilirse yukarıdaki hipotezlerde aradaki farkın yönü belirtilmemiştir. Yani “erkeklerin kadınlardan daha fazla sigara içtikleri” veya “kadınların erkeklerden daha fazla sigara içtikleri” gibi bir iddia yoktur.

İki yönlü ve tek yönlü hipotez

Karşılaştırılan grupların herhangi birisinin fazla olabileceği duruma **iki yönlü hipotez** (two-tailed) denir. Araştırmanın başında bir grubun kesinlikle ötekinden küçük (ya da büyük) olacağını çoğunlukla bilemediğimizden genelde iki yönlü hipotez kurulur. **Tek yönlü hipotez** ise gruplardan birinin ötekinden kesinlikle daha düşük (veya yüksek) olamayacağını bilindiği nadir durumlarda kurulur. Örn:

AIDS hastalığında A ilacı kullanılması durumunda hastaların tamamı (%100) bir süre sonra ölmektedir. Yeni bir B ilacı A ilacıyla karşılaştırılmak isteniyor. Ölümü önlemede B ilacının A ilacından daha kötü olma olasılığı olmadığından hipotezimizi tek yönlü kurabiliriz. Bu durumda H_0 şöyle olur:

“AIDS’ten ölümleri önlemede B ilacı A ilacından daha üstün değildir.”

Aynı şey **eşdeğerlik (non inferiority)** araştırmaları için de geçerlidir. Piyasada klasik olarak kullanılan bir orijinal amoksisillin-klavulonik asit ürünü olduğunu düşünelim. Farklı bir ilaç firması aynı etken maddede bir ürünü piyasaya sürmek istiyor. Bu durumda yeni ilacın orijinali kadar etkili olduğunu ispat etmesi yeterli olacaktır. Zaten aynı etken maddesi olduğundan orijinal ilaçtan daha etkili olduğu gibi bir iddiası yoktur. H_0 hipotezimiz şöyle olur:

“Otitis medianın tedavisinde yeni çıkarılan B ilacı orijinal A ilacından daha az etkili değildir.”

Bekleme salonunda hasta memnuniyetini artırmak için yapılan iyileştirmelerin etkinliği araştırılmak isteniyor:

“Bekleme salonundaki iyileştirmeler hasta memnuniyetini artırmamıştır.”

İki yönlü hipoteze göre tek yönlü hipotezi test etmenin daha kolay olacağı ve araştırma için daha az vaka gerekeceği açıktır.

Hipotez kurulduktan sonra uygun bir istatistik yöntem seçilir ve istatistik test uygulanır. Uygulanacak formülden elde edilecek değer H_0 hipotezini reddetmeye yöneliktir ve mutlak değeri (artı mı eksi mi olduğuna bakılmaksızın) ne kadar yüksek olursa o kadar kuvvetlidir.

Olasılık dağılımları başlığı altında bahsedildiği gibi, istatistik testler **teorik olasılık dağılımlarına** göre yorumlanırlar. Test istatistiğinden elde ettiğimiz değeri teorik dağılımımızın olasılık yoğunluk fonksiyonunun (çan eğrisi) iki veya bir ucu ile karşılaştırarak **p değerimizi** elde ederiz. Bilgisayar programları bu değeri otomatik olarak verir. **p değeri, sıfır hipotezinin doğru olması durumunda sonuçlarımızı elde etme olasılığımızı verir.**

SPSS ile bir örnek yapalım (diyabet.sav veri seti). Veri setimizde kadınlarla erkeklerin boylarını karşılaştırmak istiyoruz. Erkeklerin VEYA kadınların boylarının daha uzun olabileceğini varsayıyoruz. İki yönlü hipotez kurmalıyız (SPSS'te tek yönlü p değeri sadece özel durumlar için verilmektedir.):

H_0 : boy uzunluğu açısından kadınlarla erkekler arasında fark yoktur.

H_1 : boy uzunluğu açısından kadınlarla erkekler arasında fark vardır.

“boy” nümerik bir değişken. “erkekler” ve “kadınlar” olmak üzere birbirinden bağımsız iki grup var. Konu 1'deki akış şemalarından da görülebileceği gibi “bağımsız örneklerde t testi” yapmalıyız.

Analyze > Compare Means > Independent-Samples t test [“Test variables” kutusuna “height” değişkenini, “Grouping variable” kutusuna “sex” değişkenini koyalım. “Define Groups” butonunu tıklayıp “Group 1” için 1, “Group 2” için 2 yazalım > Continue > OK. Aşağıdaki çıktıyı elde ederiz:

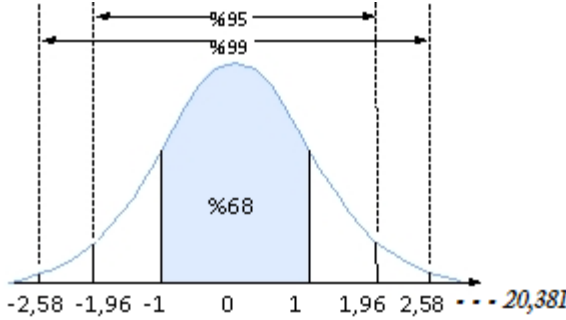
Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Height	Equal variances assumed	.116	.733	20,381	404	.000	13,1159	.8435	11,8508	14,3810
	Equal variances not assumed			20,342	394,130	.000	13,1159	.8448	11,8483	14,3835

İki yönlü p değeri 0,001'den küçük

p değerinin yorumlanması

Sıfır hipotezini kabul veya reddetmemiz için ne kadar kanıtı ihtiyacımız olduğunu önceden belirlemeliyiz. Tıptaki çalışmalar için genelde 0,05'ten küçük bir p değeri anlamlılık için yeterli sayılmaktadır. Halbuki yanılma payının çok daha önemli olduğu astronomi gibi bilim dallarında p değerleri için çok daha küçük sınırlar kullanılmaktadır. Görüldüğü gibi bu sınırın belirlenmesi nispeten **subjektiftir**. Hata yaptığımızda ciddi sonuçlar oluşabilecekse p değerini %5 yerine %1 veya binde bir alabiliriz. Buna testimizin **anlamlılık düzeyi** deriz ve makalemizde de “p anlamlılık düzeyi %5 alınmıştır” gibi ifade ederiz.

Verilerimiz normal dağıldığından elde ettiğimiz test istatistiği sonucunu (20,381) normal dağılım bilgilerimizi kullanarak yorumlayabiliriz. Bu amaçla normal dağılım olasılık yoğunluk fonksiyonu eğrisine bir göz atmak bize fikir verebilir:



Standart normal dağılımda verilerin %1'i 2,58 standart sapma sınırındaydı. Bu durumda bizim bulduğumuz 20,381 değerinin %0,1 ($p < 0,0001$) sınırından da ötede olacağına şaşırılmamalıyız.

Hesapla bulduğumuz değeri (burada t değeri) ilgili teorik dağılım için geliştirilmiş tablolardaki değerlerle karşılaştırarak da p değerini elde edebiliriz.

Makalede sadece p anlamlılık düzeyini (mesela $< 0,05$) vermek yerine tam olarak p değerini vermek okuyucu için daha bilgilendirici olacaktır.

Parametrik ve parametrik olmayan testler

Parametrik hipotez testlerini kullanabilmemiz için bazı varsayımların karşılanması gerekmektedir. İlgili hipotez testi başlığının altında belirtilecek olan bu varsayımlardan en önemlisi dağılımımızın bir teorik dağılıma benzemesidir. Eğer ilgili varsayımları karşılayamıyorsak parametrik olmayan (**nonparametrik**) hipotez testlerini kullanmamız gerekecektir. Nonparametrik testlere dağılım şartı olmayan (distribution free) veya sıralama testleri (rank methods) de denir. Bu testler verilerin gerçek değerleri yerine veri setindeki büyüklük sıralarını dikkate alarak hesap yapar. Bu nedenle de olasılık dağılımı şartları yoktur.

Bununla birlikte, normal dağılım şartıyla ilgili farklı öneriler de vardır. Bruce Weaver'ın sunumu'ndan¹ özetlediğimiz aşağıdaki hususlar bu konuda fikir verebilir:

1. Toplamı tahmin etmeye yönelik testlerde normallik daha önemlidir

¹ Weaver B. Silly or Pointless Things People Do When Analyzing Data.

https://www.nosm.ca/uploadedFiles/Research/Northern_Health_Research_Conference/Weaver,%20Bruce_Silly%20or%20Pointless%20Things.pdf

2. Hipotez testleri örneklem aralarında fark olup olmadığına bakar; dolayısıyla toplum önemli değildir.
3. Örneklem arttıkça normal dağılım önemsizleşir (central limit teorem nedeniyle)
4. Örneklem arttıkça normalite testlerinin normalden sapmayı yakalama şansı artar
5. Gruplar arasındaki n eşit olursa normallik daha az önemlidir. Gruplar arasındaki birey sayısı farkı 4 kattan az ise t testi yapılabilir.
6. Her şeye rağmen istatistik kitaplarının çoğunda normallik önemli bir varsayım olarak belirtilmektedir.
7. Her bir grupta 30 kişi kuralında da daha esnek olunmalı. Belki 15 denilebilir.

Nonparametrik testlerin diğer özellikleri de **örneklem sayısının düşük olduğu** durumlarda ve değişkenin **kategorik** özelliklerde olduğu durumlarda kullanılabilirlerdir.

Bununla birlikte, nonparametrik testler verinin gerçek değerini göz ardı ettiğinden parametrik testlere oranla daha güçsüzdür. Ayrıca bu verilerden *farkın önemlilik düzeyini* (effect of interest) de hesaplayamayız. Bu nedenle, eğer mümkünse veri dönüşümü yaparak verilerimizin teorik dağılımlara benzemesini sağlamaya çalışırız.

Hangi istatistik testi kullanacağımıza karar vermeden önce değişkenimizin türünü ve karşılaştırmak istediğimiz grupları göz önüne almalı, sonra ilgili hipotez testini yapabilmek için gerekli varsayımları karşılayıp karşılamadığımıza bakmalıyız.

Hipotez testi mi güven aralığı mı?

Güven aralıkları ve hipotez testleri yakından ilişkilidir. Hipotez testindeki amaç bir karar vermek ve tam bir p değeri sunmaktır. Güven aralığı ise ortalamalar arasındaki fark gibi dağılım genişliğiyle ilgili bilgiler sunarak sonuçları klinik olarak yorumlayabilmemizi sağlar. Ayrıca p değerini bilmesek bile güven aralığına bakarak farkın anlamlı olup olmadığı hakkında yorum yapabiliriz. Örneğin nümerik bir değişkeni iki grup arasında karşılaştırdığımızda ortalamalar arasındaki farkın %95 güven aralığına baktığımızı varsayalım:

Erkeklerle kadınların sistolik kan basınçları arasındaki farkın ortalaması 8,5 mmHg olarak bulunmuştur. Aradaki fark istatistiksel olarak anlamlı çıkmıştır ($p=0,003$). Aradaki farkın %95 güven aralığı ise [2,3 – 11,2] aralığındadır.

%95 güven aralığımızın sıfırı içermemesi durumunda iki ortalama arasındaki fark için p değerinin 0,05'ten küçük olduğunu, yani anlamlı olduğunu söyleyebiliriz ama tam p değerini veremeyiz.

Günümüzde araştırma makalelerinde p değeriyle birlikte genelde güven aralığı da sunulmaktadır.

Bölüm 22: Hipotez Testlerinde Hatalar

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun hipotez testlerindeki karar vermeye etki eden faktörler ve verilen kararda olabilecek hatalar hakkında bilgi sahibi olması amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Effect of interest’i (etki genişliği, farkların önemlilik derecesi) açıklayabilmek
- Tip 1 hatayı açıklayabilmek
- Tip 2 hatayı açıklayabilmek
- Araştırmanın gücünü açıklayabilmek
- Araştırmanın gücüne etki eden faktörleri sayabilmek
 - Örneklem büyüklüğü
 - Gözlemlerin değişkenliği (variabilitesi)
 - Effect of interest
 - Önemlilik düzeyi (yanılma payı)
- Bağımsız örneklemelerde t testi ve ANOVA için güç hesabı yapabilmek

Karar verme

Temel istatistik uygulamalarını **tanımlayıcı istatistikler** ve **hipotez testleri** (önemlilik testleri) diye ayırabiliriz. Bundan önceki konularda tanımlayıcı istatistikleri (grafiklerle gösterme, ortalama, standart sapma, güven aralığı vs.) gördük. Hipotez testleri ise bir iddia veya varsayım hakkında karar vermemize yardımcı olur.

İki farklı sağlık kuruluşunda tedavi gören hipertansif hastaların sistolik kan basınçlarıyla ilgilendiğimizi düşünelim. Öncelikle iki grup açısından karşılaştıracığımız bu nümerik değişkenin gruplar arasında ne kadar farklı olmasının bizim için (klinik açıdan) önemli olduğuna karar vermeliyiz. İki grup arasında yakalamak istediğimiz, önemli olduğunu düşündüğümüz bu farka **comparison of interest, true difference of means, effect of interest** veya **treatment effect** (saptamak istediğimiz en küçük fark, etki genişliği, farkların önemlilik derecesi) denir. Sıfır hipotezimiz “iki sağlık kuruluşunda tedavi gören hipertansif hastaların sistolik kan basınçları arasında fark olmadığı” şeklindedir. İki yönlü H_1 hipotezimiz ise A VEYA B hastanesindeki hastaların kan basıncı yüksek olacak şekilde arada bir fark olduğu şeklindedir. İşte “treatment effect” arada fark vardır veya yoktur derken kastettiğimiz FARKIN ne kadar olduğudur. Bu nedenle araştırmaların yöntem bölümünde güç hesabı ile birlikte etki genişliğinin de ne kadar alındığı belirtilmelidir.

Yanlış karar verme

Kurduğumuz hipotezle ilgili karar verirken en doğru kararı vermek isteriz. Halbuki, toplumun tamamını değil de sadece bir örneklemini incelediğimizden kararımızda hata ihtimali her zaman olacaktır. Bu hataların neler olabileceğini anlamak için aşağıdaki tabloya bakabiliriz.

Tablo: Hipotez testlerinin sonuçları.

	H₀ reddedildi	H₀ kabul edildi
H₀ doğru	Tip I hata	Doğru karar
H₀ yanlış	Doğru karar	Tip II hata

Tip I hata: sıfır hipotezi doğru olduğu halde reddedilmesidir. Sonuçta incelediğimiz gruplar arasında fark olmadığı halde “fark vardır” şeklinde karar veririz. Tip I hataya yol açma olasılığı α (**alfa**) ile gösterilir. Bu aynı zamanda testimizin anlamlılık düzeyidir. p değerimizin α ’dan daha düşük olması halinde sıfır hipotezini reddederiz.

Araştırmamızın proje aşamasında α değerimizin kaç olacağına karar vermeliyiz. Bu değer geleneksel olarak tıbbi çalışmalarda 0,05 olarak alınmakla birlikte, incelediğimiz durumun hassasiyetine ve önemine göre daha küçük veya daha büyük de olabilir.

Tip II hata: yanlış olduğu halde sıfır hipotezini kabul etmemizdir. Sonuçta incelediğimiz gruplar arasında bir fark olduğu halde “fark yoktur” sonucunu çıkarabiliriz. Tip II hataya yol açma olasılığı β (**beta**) ile gösterilir. $1 - \beta$ ise testimizin **gücünü** (power) verir.

Bu durumda araştırmanın gücünü şu şekilde tanımlayabiliriz: “*araştırmanın gücü, doğru olduğu durumda sıfır hipotezini kabul edebilme olasılığımızdır*”. Örnek verecek olursak:

H₀: A hastanesi ve B hastanesinde takip edilen hipertansif hastalar arasında sistolik kan basınçları açısından anlamlı bir fark yoktur.

Gerçekte (ki, “gerçeği” kimin bildiği felsefi bir tartışma konusu olabilir) de A hastanesindeki hastaların kan basınçlarının B hastanesindekilerden (kabul ettiğimiz “effect of interest” açısından) farklı olmadığını varsayalım. Bu durumda H₀’ın kabul edilmesi gerekir.

*Biz de istatistik analizimiz sonunda doğru bir karar verdik ve H₀’ı kabul ettik. İşte bu doğru kararı verebilmemiz yaptığımız testin **gücüdür**.*

Yukarıdaki açıklamalar doğrultusunda hipotez testimizin olası sonuçlarını aşağıdaki tablodaki gibi gösterebiliriz:

Hipotez testi sonucunda verilen karar			
		H ₀ reddedildi	H ₀ kabul edildi
Gerçek	H ₀ doğru	Tip I hata (%5)	Doğru karar (Araştırmanın gücü; Power) (%80)
	H ₀ yanlış	Doğru karar (%95)	Tip II hata (%20)

Tabii ki araştırmamızın gücünün %100 olmasını isteriz. %80’lik bir güç ise genelde kabul edilebilirdir. Ancak, düşük de olsa tip II hata yapma olasılığımız her zaman vardır. Kaldı ki, tip II hata yapmayacağız diye işi sıkı tutarsak bu sefer tip I hata yapma olasılığımız artacaktır.

Diğer taraftan, araştırmanın gücüne etki edecek faktörleri biliyoruz ve araştırmamızı planlarken bunları dikkate alabiliriz.

Araştırmanın gücü ve ilişkili faktörler

Araştırmamızın planlama aşamasında gücünü hesaplamamız çok önemlidir. Böylece -eğer varsa- gruplar arasında saptamayı hedeflediğimiz farkın miktarını (effect of interest) ve araştırmamızın *yeterli* bir güce sahip olup olmadığını önceden belirlemiş oluruz (*yeterli güç*’ten genelde %80 anlaşılır). Sadece %40’lık bir güce sahip bir araştırmayı yürütmek etik açıdan kabul edilemez olduğu gibi ekonomik açıdan da bir israftır.

Araştırmanın gücüne etki eden başlıca faktörler şunlardır:

1. **Örneklem Büyüklüğü:** örneklem arttıkça araştırmanın gücü de artar. Dolayısıyla büyük bir örneklem – eğer varsa – klinik olarak daha küçük bir farkı saptayabilir.
2. **Gözlemlerin Variabilitesi:** gözlemlerin standart sapması düştükçe araştırmanın gücü de artar.
3. **Etki Genişliği¹** (Effect of interest): gruplarımız arasındaki küçük bir farkı yakalamak istediğimizde araştırmanın gücü düşük olur. Büyük farkları saptama konusunda güç daha fazladır.

¹ Bazı formüllerde etki genişliği, bazılarında ise hata marjı kullanılmaktadır. İkisini açıklamak için şöyle bir örnek verebiliriz:

A ilacını kullanan 50 ve B ilacını kullanan 50 kişide kan şekeri ölçtüğümüzü varsayalım. Ortalama kan şekeri sonuçları şöyle olsun: A ilacı kullananlarda: 100 mg/dl,

4. **Anlamlılık Düzeyi:** anlamlılık düzeyini yüksek tuttuğumuzda (yani tip I hata payını artırdığımızda) araştırmamızın gücü de artar. Bu durumda tip II hata da azalmış olur.
5. **Hipotezin Yönü:** Hipotezin tek yönlü veya iki yönlü olarak kurulması örneklem sayısını etkileyecektir. İki yönlü hipotez kurulması durumunda daha fazla örneklem gerekecektir.
6. **Hipotez Testi:** Yapılacak hipotez testine göre örneklem hesabı değişecektir. Araştırmanın gücünü hesaplariken öncelikle hipotez testine karar vermek gerekecektir.
7. **Araştırma Tasarımı:** Araştırma tasarımına göre bazı araştırmalarda veri kayıpları beklenebilir. Örneğin, araştırmadan ayrılanların %10 civarında olması beklenirse, bu durumda örneklem sayısının üzerine %10 ilave edilmelidir. Diğer taraftan, örneklemin alt kümelerden oluşması halinde kümelerin içindeki bireyler birbirine benzeyeceğinden alınan örneklem sayısı küme katsayısı dikkate alınarak artırılmalıdır.

Bu noktada güven aralığı konusunu da tekrar hatırlatmakta fayda var. Güven aralığının geniş olması araştırmanın örneklem sayısının az olmasından kaynaklanıyordu ve dolayısıyla araştırmamızın hassasiyetinin düşük olduğunu gösteriyordu. Bu da araştırmanın gücünün düşük olması anlamına gelir.

Birden çok hipotezi test etme

Araştırmamızın tek bir ana sonuç ölçütünün olması ve olabildiğince az hipotez testi yapılması önerilmekle birlikte, sıklıkla bir veri setinde birden fazla hipotez testi uygularız. Olasılık kurallarından hatırlayacağımız gibi birden çok test yaptığımızda araştırmamızın hata ihtimali önemli ölçüde artacaktır. Örn:

Bir hipotez testi uyguladığımızda %5 yanılma yapımızın olduğunu varsayalım. Aynı araştırmada iki test uyguladığımızda biri VEYA

B ilacını kullananlarda: 105 mg/dl. Her iki ortalamaya ait standart sapmalar (SS) da 10 civarında olsun.

Etki genişliği (effect size): $(\text{Ortalama A} - \text{Ortalama B}) / \text{SS} = 5 / 10 = 0,5$ olarak hesaplanır. Etki genişliği Cohen tarafından kabaca şöyle sınıflandırılmıştır: küçük: 0,2, orta: 0,5, büyük: 0,8.

Hata payı (margin of error): Hata payı genelde %95 güven aralığı ile hesaplanır. Örneğimizdeki iki ortalama arasındaki farkın %95 güven aralığı 1,03 ile 8,97 arasında (7,94 büyüklüğünde) çıkar.

Hem etki genişliği, hem de hata payının hesaplanmasında kullanılan parametreler aynıdır (ortalamalar, standart sapmalar ve örneklem sayıları). Bazı örneklem miktarı hesaplamalarında etki genişliği, bazılarında ise hata payı tercih edilmektedir.

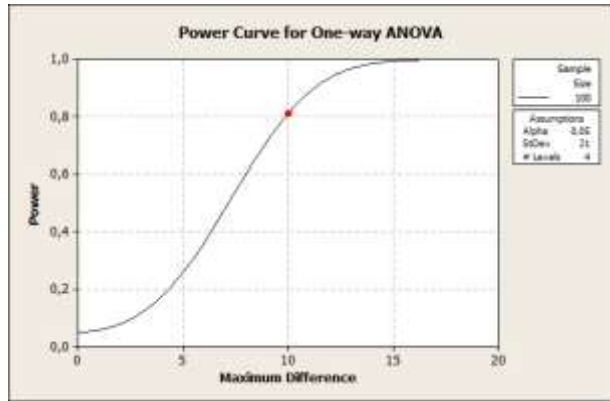
ötekinin yanlış olma olasılığı %5+%5 = %10 olacaktır. Böylece 10 ayrı test yapsak araştırmamızın genel hata ihtimali %50'ye çıkacaktır.

Çoklu hipotez testi uygulamanın risklerinden kurtulmak için yapacağımız karşılaştırmaları iyi seçmeli ve en aza indirmeliyiz. Bunun dışında bazı istatistiksel yaklaşımlar da yararlı olabilir. Örneğin **Bonferroni** yaklaşımında her bir hipotez testinden elde ettiğimiz p değerini araştırmamızda toplam uyguladığımız test sayısı ile çarparak düzeltilmiş p değerleri elde edilir. Örn: Araştırmamızda 3 adet hipotez testi uyguladığımızı düşünelim:

Hipotez	Hesaplanan p değeri	Bonferroni düzeltmesinden sonra p
H_0 -1: A ve B tedavisi alanlar arasında kan basıncı açısından fark yoktur	0,02	$0,02 \times 3 = 0,06$
H_0 -2: Kadınlarla erkekler arasında kan basıncı açısından fark yoktur	0,01	$0,01 \times 3 = 0,03$
H_0 -3: Zenginlerle fakirler arasında kan basıncı açısından fark yoktur	0,008	$0,008 \times 3 = 0,024$

Bonferroni düzeltmesi, alfa (tip I) hatasına karşı koruma sağlarken beta hatasını dikkate almaması ve gereksiz kullanılması açısından eleştirilmektedir.

Yanda MINITAB® 15 ile (<http://www.minitab.com/>) yapılan bir power hesabı örneği verilmiştir. Dikkat edilirse saptanmak istenen fark (maximum difference) azaldıkça araştırmanın gücü de (power) azalmaktadır.



Bu örneklem hesabı ANOVA için yapılmıştır. Dört farklı grupta sistolik kan basınçlarının karşılaştırılacağı varsayılmıştır. Sistolik kan basıncı için standart sapma 21 mmHg, her bir gruptaki vaka sayısı 100, tip I hata ise 0,05 olarak belirlenmiştir. Bu durumda gruplar arasındaki 10 mmHg'lik bir farkı saptamak için %81 güç oluşmaktadır.

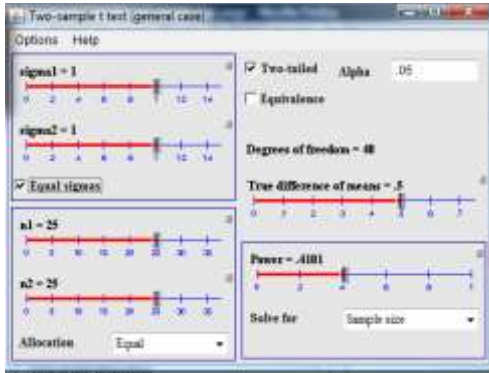
İnternette ücretsiz kullanabileceğiniz örneklem ve güç hesap araçları da mevcuttur. Bunlardan birisi Russ Lenth'in Java modülüdür: <http://www.stat.uiowa.edu/~rlenth/Power/>. Bu kez nümerik veri iki grup karşılaştırması yapalım:

Bir grupta kadınlarla erkeklerin kan basınçlarını karşılaştırmak istiyoruz. Bağımsız örneklemelerde t testi (independent samples t-test veya two sample t test) yapmamız gerekiyor ("İstatistiksel testler ve kullanım yerleri" konusuna bakınız. Nümerik veri – iki bağımsız grup).

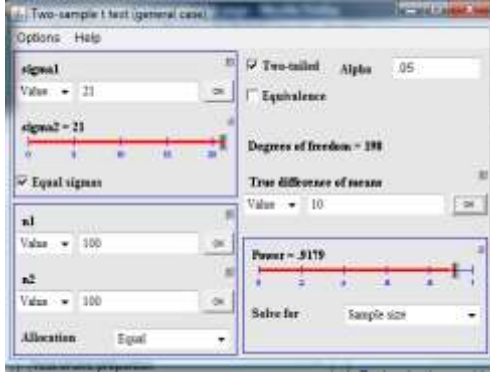
Russ Lenth'in sitesinden "Two samples t-test'i" seçim



Run selection'ı tıkladığımızda aşağıdaki pencere açılacaktır (bilgisayarımızda Java kurulu olmalı):

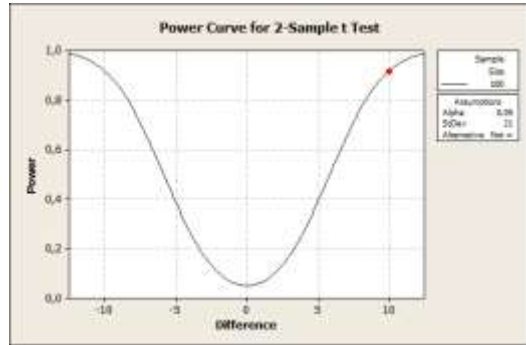


Standart sapmayı (sigma1 ve sigma2) önceki örnekte olduğu gibi 21 mmHg alalım. Varyansları eşit kabul edelim (Equal sigmas). Örneklem sayısını her iki grup için 100 alalım ($n1=100$, $n2=100$). İki uçlu alfa 0.05 kalsın. Ortalamalar arasında saptamak istediğimiz fark (True difference



of means) 10 mmHg olsun.
Power %91,79 olarak çıkacaktır:

Aynı alıştırmayı MiniTab® ile
yapsak şu sonucu elde ederiz:



Alıştırma 9

Hipotez testlerinde yaptığımız hataları yargıda yapılan hatalara benzetelim. Bir şüpheli hakimın karşısına çıkarıldığında iki ihtimal vardır: ya beraat eder, ya da ceza alır. Her iki durumda da hukuk sisteminin yanılma olasılığı vardır.

Avukatın hipotezi (H_0): “Müvekkilim atfedilen suçu işlememiştir”		Hakimin kararı	
		H_0 reddedildi: Suçlu	H_0 kabul edildi: Suçsuz
Gerçek (Zanlının kendisinin bildiği gerçek)	H_0 doğru (kabul): Suçsuz	Tip 1 hata Tolerans payı %5	Doğru karar Hukukun gücü (%80)
	H_0 yanlış (ret): Suçlu	Doğru karar (%95)	Tip 2 hata Tolerans payı %20

Dikkat ederseniz yargının gücü suçluya suçlu demesinde değildir; esas güç, suçsuzla suçsuz diyebilmektir!

Soru 1: İstatistiksel anlamlılık düzeyi p olarak gösterilen alfa hatası genelde kaç olarak alınır ve neyi ifade eder?

Soru 2: Tip 1 ve Tip 2 hata arasında nasıl bir ilişki vardır?

Soru 3: Bir araştırmanın gücüne neler etki eder?

Soru 4: Örneklem sayısını gerektiğinden çok yüksek tutarak yanıltıcı bir sonuç elde edebilir miyiz?

Çözüm için bakınız s. 474

Bölüm 23: Nümerik Veri Tek Grup (Tek Örneklemde t Testi)

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun tek gruptan elde edilen nümerik verilerin toplum ortalamasıyla karşılaştırılmasında kullanılan tek örneklemde *t*-testi (one-sample *t*-test) ve işaret testi (sign test) hakkında bilgi sahibi olması amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Tek örneklemde *t*-testi varsayımlarını açıklayabilmek
- SPSS ile tek örneklemde *t*-testi yapabilmek
- İşaret testinin kullanım yerini tartışabilmek
- İşaret testinin nasıl yapılacağını açıklayabilmek

Mahallemizdeki bireylerin boylarını ölçtük [mahalle=tek grup; boy=nümerik veri]. Sonuçlarımızı il geneli ile karşılaştırmak istiyoruz. Daha önceki araştırmalardan ilimizdeki bireylerin boy ortalamasının 161,5 cm olduğunu biliyoruz.

*Bu karşılaştırmayı **tek örneklemde t-testi** (one-sample *t*-test) ile yapabiliriz. Tek örneklemde *t*-testini yapabilmemiz için verilerimizin bazı varsayımları karşılaması gerekir. Bu varsayımları karşılamamız halinde alternatifimiz **nonparametrik** bir test olan **işaret testi** olabilir.*

Tek örneklemde t-testi varsayımları

Tek örneklemde *t*-testinin varsayımları;

- Değişkenimizin toplumda normal dağılması ve
- Yeterli bir örneklemeye sahip olmamızdır.

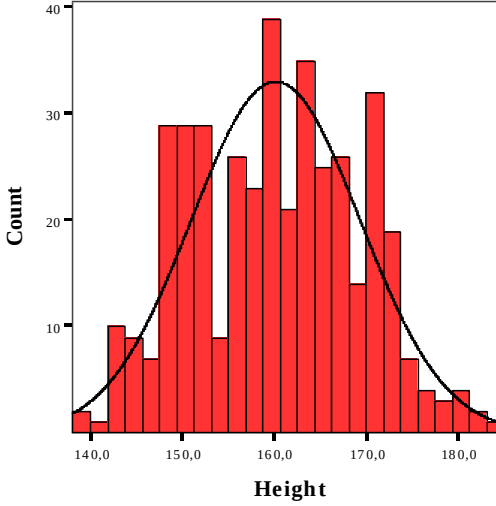
Toplunun varyansını (σ) bilmememiz durumunda verilerimizin dağılımı teorik dağılımlardan *t* dağılımına uyacaktır. Eğer toplumun varyansını bilirsek veya örneklem sayımız çok yüksekse bu durumda normal dağılım bilgilerine dayanan *t* testini kullanabiliriz. Ancak, her iki durumda da sonuçlar aynı çıkacaktır.

Örnek uygulama

Diyabet.sav verisetini kullanarak bir uygulama yapalım:

Öncelikle tek örneklemde *t*-testi varsayımlarını karşılayıp karşılamadığımızı test etmemiz gerekecek:

- Boy değişkeni toplumda normal dağılır. Örneklemimizde de boy değişkeninin histogram grafiğine baktığımızda normal dağılıma yakın olduğunu görüyoruz:



Skewness değerine de bakılacak olursa normal dağıldığı görülecektir:

	N	Skewness	
	Statistic	Statistic	Std. Error
Height	406	-,006	,121
Valid N (listwise)	406		

- Örneklem sayımız da 406 olup yeterlidir.

Hipotez testleri için 5 basamaklı genel yaklaşımımızı uygulayacak olursak:

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

H_0 : İlimizdeki bireylerin boy ortalaması ile mahallemizdeki bireylerin boy ortalaması arasında fark yoktur.

H_1 : İlimizdeki bireylerin boy ortalaması ile mahallemizdeki bireylerin boy ortalaması farklıdır.

2. Verilerin toplanması:

Verilerimizi topladık ve SPSS'e girdik. Diyabet.sav verisetinde "Weight" değişkenini oluşturduk.

3. İlgili sıfır hipotezi için *test istatistiğinin* hesaplanması:

Tek örneklemede *t* testi için kullanacağımız formül:

$$t = \frac{(\bar{x} - \mu_1)}{s/\sqrt{n}}$$

$\bar{X}$ = örneklemimizin aritmetik ortalaması.	160,0
μ_1 = toplum ortalaması	161,5
<i>s</i> = örneklemimizin standart sapması	9,2
<i>n</i> = örnek sayısı	406

Buradan $t = -3,285$ olarak hesaplanır.

4. Test istatistiğinden elde edilen değerin bilinen bir olasılık dağılımı ile karşılaştırılması:

Sonucumuzu *t* tablosu ile (Appendix 2) karşılaştıracak olursak $t = 3,285$ ve serbestlik derecesi = 405 için *p* değerinin 0,01'den küçük olduğunu görürüz.

<i>df</i>	Two-tailed <i>P</i> -value			
	0.10	0.05	0.01	0.001
1	6.314	12.706	63.656	636.58
2	2.920	4.303	9.925	31.600
<i>SD = 405</i> $t = 3,285$				
100	1.660	1.984	2.626	3.390
200	1.653	1.972	2.601	3.340
5000	1.645	1.960	2.577	3.293

5. *P* değerinin ve sonuçların yorumlanması

P değeri 0,05'ten küçük olduğu için H_0 hipotezini reddederiz.

Dolayısıyla mahallemizin insanları ile il genelindeki bireyler arasında boy açısından istatistiksel olarak anlamlı bir fark vardır.

Mahallemizdeki bireyler daha kısadır.

Toplum ortalamasının %95 güven aralığını da hesaplayacak olursak:

$$\%95 \text{ GA} = \bar{X} \pm 1,96 \times (s/\sqrt{n}) = 160 \pm 1,96 \times 0,45$$

$$\%95 \text{ GA: } [159,11 - 160,88]$$

Bulduğumuz güven aralığı toplum ortalamasını (161,5) içermediğinden aradaki farkın anlamlı olduğunu söyleyebiliriz.

Örneğin SPSS ile yapılması

SPSS ile yapacak olursak:

Analyze>Compare Means >One-Sample T Test>[“height” değişkenini
“Test Variable(s)” alanına geçirelim]> “Test Value” kısmına 161,5
yazalım>ok

	N	Mean	Std. Deviation	Std. Error Mean
Height	406	160,015	9,2030	,4567

	Test Value = 161.5					
	t	Df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Height	-3,252	405	,001	-1,4852	-2,383	-,587

%95 güven aralığına da baktığımızda iki ortalama arasındaki farkın -2,3 ile -0,58 arasında olabileceğini görüyoruz. Sonuçta fark %95 olasılıkla hep eksidir. Aradaki farkın sıfır veya artı olamayacağından istatistiksel olarak mahallemizdeki bireylerin il genelindeki bireylerden daha kısa olduğunu söyleyebiliriz.

Varsayımlar karşılanamıyorsa

Tek örneklemede t-testi varsayımlarının karşılanamadığı durumlarda ya verilerimizin varsayımlarını karşılamasını sağlamalıyız (normal dağılıma uydurmak için veri dönüştürme veya örneklem sayısını artırma) ya da bu parametrik testin nonparametrik versiyonundan yani işaret testinden yararlanmamız gerekir. Bu noktada tek örneklemede t testinin normallikten küçük sapmalara dayanıklı olduğunu belirtmeliyiz.

İşaret testi

İşaret testi dağılımımızın ortancasını esas alarak hesaplanır.

Eğer dağılımımızın toplum dağılımından farklı değilse o zaman verilerimizin yarısı toplum ortalamasının üstünde, yarısı da altında olmalıdır.

İşaret testi, değişkenin gerçek değerlerini dikkate almayıp sadece toplum ortalamasından büyük mü küçük mü olduklarına baktığından zayıf ve basit bir testtir.

Yukarıdaki örneğe benzer bir çalışma yaptığımızı ama daha küçük bir örneklemden veri topladığımızı düşünelim:

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması

H_0 : İlimizdeki bireylerin boy ortalaması ile örneklemimizdeki bireylerin boy ortalaması arasında fark yoktur.

H_1 : İlimizdeki bireylerin boy ortalaması ile örneklemimizdeki bireylerin boy ortalaması farklıdır.

2. Verilerin toplanması

Birey no	Boy (cm)	Birey no	Boy (cm)
1	151,0	15	176,0
2	170,0	16	169,0
3	170,0	17	167,0
4	164,0	18	160,0
5	165,0	19	173,0
6	160,0	20	172,0
7	162,0	21	173,0
8	164,0	22	158,0
9	161,0	23	173,0
10	.	24	165,0
11	169,0	25	162,0
12	170,0	26	165,0
13	157,0	27	174,0
14	172,0	28	158,0

3. İlgili sıfır hipotezi için *test istatistiğinin* hesaplanması

Örnekleminizde 28 birey var. Boy değişkenimizi küçükten büyüğe doğru sıralayalım:

Birey no	Boy (cm)	Birey no	Boy (cm)
10	.	26	165,0
1	151,0	17	167,0
13	157,0	11	169,0
22	158,0	16	169,0
28	158,0	2	170,0
6	160,0	3	170,0
18	160,0	12	170,0
9	161,0	14	172,0
7	162,0	20	172,0
25	162,0	19	173,0
4	164,0	21	173,0
8	164,0	23	173,0
5	165,0	27	174,0
24	165,0	15	176,0

1 bireyin boyla ilgili verisinin girilmediğini görüyoruz.

Toplum ortalamasına (örneğimizde 161,5 cm) eşit olan bireyleri hesaba katmayacağız.

161,5'e eşit olan birey olmadığını görüyoruz. Geriye kalan 27 bireye (**n'**) baktığımızda sıralamaya göre

7 bireyin boyunun toplum ortalamasından küçük,

20 bireyin boyunun ise toplum ortalamasından büyük olduğunu görüyoruz.

İşaret testinde bu iki değerden küçük olanı (**r**) dikkate alınır (örneğimizde 7).

$n' \leq 10$ olması durumunda işaret testi değeri $r = n'$ olarak alınır.

$n' > 10$ olması durumunda z hesabı yapılır

$$z = \frac{\left| r - \frac{n'}{2} \right| - \frac{1}{2}}{\frac{\sqrt{n'}}{2}}$$

Bizim örneğimizde $n' = 27$ değeri 10'dan büyük olduğuna göre z formülünü uygulamalıyız:

$$z = \left[\left| 7 - (27/2) \right| - (1/2) \right] / (\sqrt{27}/2) = 2,30.$$

4. Test istatistiğinden elde edilen değer bilinen bir olasılık dağılımı ile karşılaştırılması

r değerini sayfa 463'ten, z değerini ise sayfa 458'den bakarak p 'yi bulabiliriz.

z	2-yönlü P -değeri
0.0	1.000
0.1	0.920
0.2	0.841
2.2	0.028
2.3	0.021
2.4	0.016
2.5	0.012

$z = 2,3$ için iki yönlü p değerinin **0,021** (çan eğrisinin her bir tarafının olasılığı için $p = 0,021/2 = 0,0105$) olduğunu görüyoruz.

5. P değerinin ve sonuçların yorumlanması

P değeri 0,05'ten küçük olduğundan sıfır hipotezini reddetmemiz gerekir. Dolayısıyla örneklemimizle toplum ortalaması arasında boy açısından istatistiksel olarak anlamlı bir fark vardır.

Alıştırma 10

Bu alıştırmada daha önceki konuda işlediğimiz araştırmanın gücü (power) ve buna etki eden faktörleri tekrar gözden geçirmek istiyoruz.

Bu durumda araştırmanın gücünü şu şekilde tanımlayabiliriz: "*araştırmanın gücü, doğru olduğu durumda sıfır hipotezini kabul edebilme olasılığımızdır*".

Aşağıdaki bağlantıda verilen “Quality of life of people with diabetes attending primary care health centres in Riyadh: bad control - good quality?” başlıklı çalışmayı inceleyiniz:

<http://www.ingentaconnect.com/content/rmp/qpc/2007/00000015/00000005/art00009;jsessionid=1bxrllhxrip9s.victoria>

Makalenin yöntem kısmı aşağıya çıkarılmıştır:

Riyadh is the capital city of KSA with around 4 million inhabitants.¹⁷ There are 78 PHC centres in Riyadh city providing free health services to all.¹⁸ Primary care physicians (mostly without a qualification in family medicine) run the chronic disease clinics in these centres. In this health system, all families have to be registered to a PHC centre. Thus, patients with diabetes are also expected to be followed up there. However, some patients prefer to be followed up by their specialists in the hospital. Although there are no data on the proportion of patients who do so, it is usually about 20% (authors' personal observations).

Three of the PHC centres in Riyadh were selected randomly, and nurses providing the diabetes care were invited to join the study. The selected centres were from different geographical areas, namely from south, east, and central Riyadh. The total number of registered persons for each centre was 38 165, 30 961 and 27 337 persons respectively. Out of these registered persons, the number of patients with diabetes for the south, east, and central area were 697, 731 and 685 persons respectively. In KSA, 59.9% of the population is aged 19 years or less.¹⁹ The centres are divided into male and female sections and two to five practitioners serve in each section, together with a nurse experienced in diabetes care.

Data about the demographic characteristics of the patients, their medications and level of control of the disease were collected. Dartmouth Primary Care Cooperative Information Project/World Organization of Family Doctors (COOP/WONCA) charts were used as a tool to collect information on the main outcome (perceived HRQOL of the patients). These charts are developed by the WONCA Classification Committee in collaboration with the WONCA Research Committee.²⁰ The COOP/WONCA charts evaluate perceived HRQOL in six dimensions using a five-point Likert scale (1 = excellent, 5 = poor): physical fitness, feelings, daily activities, social activities, change in health and overall health. Average HRQOL scores were calculated by calculating the average scores of the six dimensions. Other outcomes measured were fasting blood glucose levels, haemoglobin A_{1c} (HbA_{1c}) levels, presence of any co-morbidity, and body mass index.

Adult diabetic patient registries in the selected health centres were reviewed. A list of all people with diabetes with active files (regular follow-up of at least three per year) was prepared. Twenty percent of the patients were sampled randomly from this list. Patients were interviewed in the health centres. The aim of the study was explained to all participants followed by an oral consent. Care was given for data privacy, and personal information of the participants was not disclosed at any stage. Written permission was taken from the Ministry of Health, where ethical perspectives are also taken into account. Data could be obtained and analysed for 440 patients.

Power and sample size calculation was based on the HRQOL outcome (physical fitness dimension). Previous studies have shown mean physical fitness scores between 3.0 and 3.7 using the COOP/WONCA charts.^{21–23} Taking the population standard deviation as 0.8 and the effect of interest as 7%, 338 patients were needed to reach a power of 90%. Student t test, Mann–Whitney U test, Chi square analysis, Spearman correlation and logistic regression analysis were used for the hypothesis testing. A P value of less than 0.05 was regarded as statistically significant.

Aşağıdaki soruları cevaplayınız:

1. Araştırmanın ana sonuç ölçütü nedir?
2. Araştırmanın hipotezi nedir? Siz olsanız nasıl bir hipotez kurardınız (H_0 ve H_1 'i kurunuz)
3. Araştırmanın örnekleme kaç kişidir?
4. Ana sonuç ölçütüyle ilgili gözlemlerin variabilitesi nedir?
5. Effect of interest (önemlilik derecesi) kaç olarak alınmıştır?
6. İstatistiksel anlamlılık düzeyi kaç olarak alınmıştır?
7. Araştırmanın gücü kaçtır?
8. Yukarıdaki verileri Russ Lenth'in web sayfasında (<http://www.stat.uiowa.edu/~rlenth/Power/>) yerine koyarak hesaplama yapınız (test olarak "two sample t-test" seçiniz). Araştırmanın gücü kaç çıkıyor?
9. Literatür tarayınız. Yönteminde örneklem hesabı belirtilmiş olan bir makale bulunuz. 3-7 arası soruları kendi bulduğunuz makaleye göre cevaplandırınız (kontrol edebilmemiz için makaleyi bizimle paylaşınız).

Çözüm için bakınız s. 475

Bölüm 24: Nümerik Veri İki Bağımlı Grup (Bağımlı Örneklemelerde t Testi, Wilcoxon Testi)

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun iki bağımlı gruptan elde edilen nümerik verilerin ortalamalarının karşılaştırılmasında kullanılan bağımlı örneklemelerde *t*-testi (paired *t*-test) ve bunun nonparametrik alternatifi olan Wilcoxon işaretli sıra testi (Wilcoxon signed rank test) hakkında bilgi sahibi olması ve bu testleri yapabilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Bağımlı örneklemelerde *t*-testi varsayımlarını açıklayabilmek
- SPSS ile bağımlı örneklemelerde *t*-testi yapabilmek
- Wilcoxon işaretli sıra testinin kullanım yerini tartışabilmek
- Wilcoxon işaretli sıra testinin mantığını açıklayabilmek
- Wilcoxon işaretli sıra testini SPSS ile yapabilmek

Diyabetik hastalarımıza bir beslenme ve egzersiz programı uyguladık. Program sonunda katılımcıların kilo verip vermediklerini araştırmak istiyoruz (H_0 : Bireylerin uygulanan beslenme ve egzersiz programı öncesi ve sonrasındaki vücut ağırlıkları arasında fark yoktur). Ölçmek istediğimiz değişken nümeriktir. Uygulama öncesi ve sonrasında elde edilen iki ağırlık ölçümü ise birbirinden bağımlıdır.

Hatırlayalım:

- Aynı bireylerde herhangi bir girişim, tedavi vs. öncesi ve sonrasında yapılan ölçümler birbirine bağımlıdır (bireyin değeri başlangıçta yüksekse sonraki ölçüm de nispeten yüksek çıkacaktır).
- Bireyler farklı olsa da iki grup arasında bire bir eşleşme yapılmışsa (örn. eşleştirilmiş vaka kontrol çalışmalarında) yine eşleşen bireylerin ölçümleri birbirine yakın çıkacağından bağımlı olarak değerlendirilir. Yaş, cinsiyet, boy, kilo ve meslek açısından eşleştirilmiş iki birey düşünelim. Bu bireylerden birinin tansiyonu 140/90 mmHg ise ötekinin de buna yakın bir değer olacağını tahmin edebiliriz.

Bizim örneğimizde bağımlı iki nümerik ölçüm söz konusudur. Kitabın başında da verilen istatistik testlerin uygulama yerleri akış şemalarına baktığımızda seçeceğimiz testin bağımlı örneklemelerde *t*-testi veya Wilcoxon işaretli sıra testi olacağını görürüz.

Bağımlı örneklemelerde t-testi

Varsayımları

- Değişken, toplumda normal dağılmalı
- Örneklem sayısı yeterli olmalı.

Testin mantığı

Eğer tedavi öncesi ve sonrasında bireylerin değerleri arasında fark yoksa, her bireyin önceki ve sonraki ölçümleri arasındaki farkların ortalamasını aldığımızda sıfır olmalıdır. Bu durumda testimiz, daha önce gördüğümüz tek örneklemde t-testine indirgenebilir. Değişkenimiz, farkların ortalaması, toplum ortalaması ise sıfırdır.

Ek açıklama

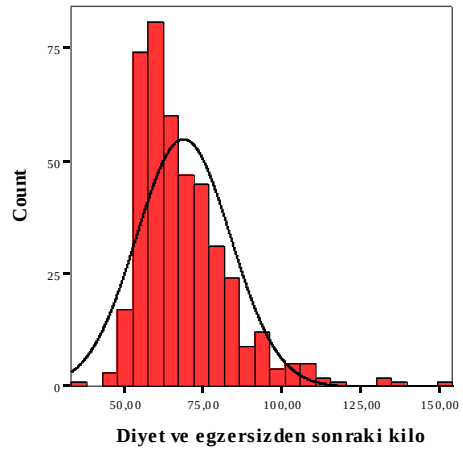
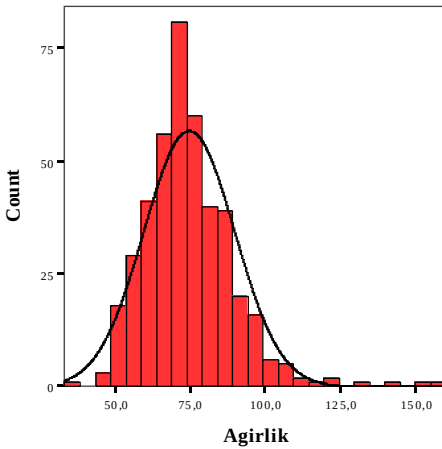
Veriler bağımlı (eşleştirilmiş) olduğundan her iki ölçümde de örneklem sayısı eşit olmalıdır. n adet bireydeki iki ölçüm arasında n adet fark olacaktır. Bunların aritmetik ortalaması $\bar{x}$ ve tahmin edilen standart sapması s_d 'dir.

Örnek uygulama

Diyabet.sav verisetini kullanarak bir uygulama yapalım:

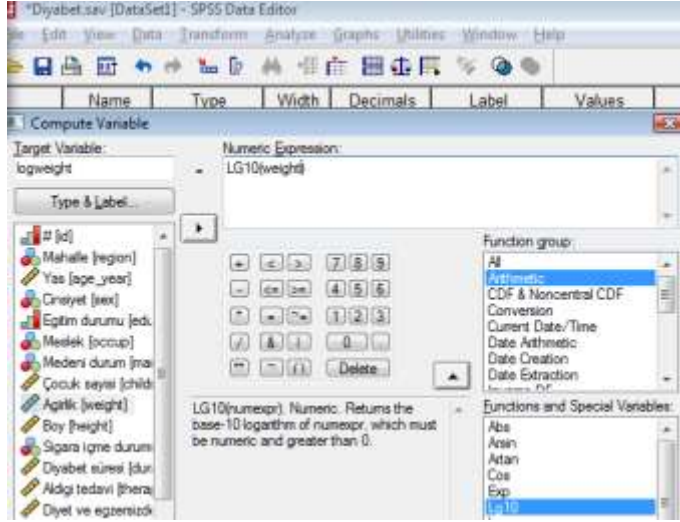
Öncelikle verilerimizin bağımlı örneklemelerde t-testi varsayımlarını (şartlarını) karşılayıp karşılamadığına bakmamız gerekecek:

- Ağırlık (weight ve weight2) değişkenlerimizin histogram grafiklerine baktığımızda normal dağılımdan hafif bir sapma olduğunu görüyoruz:



Bu durumda sağa eğimli verilerimizi normal dağılıma yaklaştırmak için logaritmik dönüşüm yapmayı deneyebiliriz. Weight ve weight2 değişkenlerimize logaritmik dönüşüm uyguladıktan sonra

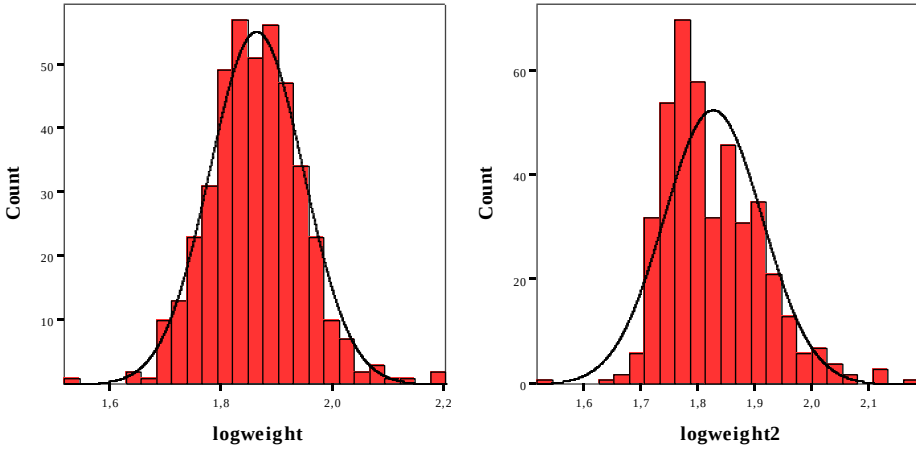
Transform>Compute variable>[“Target Variable” alanına “logweight”, “Numeric expression” alanına “LG10(weight)” girelim>ok.



Aynı işlemi weight2 için tekrarlayalım:

Transform>Compute variable>[“Target Variable” alanına “logweight2”, “Numeric expression” alanına “LG10(weight2)” girelim>ok.

histogram grafikleri aşağıdaki gibi olmaktadır:



Görüldüğü gibi verilerimizin dağılımı artık normal dağılıma daha yakındır.

- Örneklem sayısı da yeterli olduğuna göre testimizi uygulayabiliriz.

Hipotez testleri için 5 basamaklı genel yaklaşımımızı uygulayacak olursak:

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

H_0 : Bireylerin uygulanan beslenme ve egzersiz programı öncesi ve sonrasındaki vücut ağırlıkları arasında fark yoktur

H_1 : Bireylerin uygulanan beslenme ve egzersiz programı öncesi ve sonrasındaki vücut ağırlıkları arasında fark vardır.

2. Verilerin toplanması:

Verilerimizi topladık ve SPSS'e girdik. Logaritmik dönüşüm uygulayarak diyabet.sav verisetinde "logweight" ve "logweight2" değişkenlerini oluşturduk.

3. İlgili sıfır hipotezi için *test istatistiğinin* hesaplanması:

Bağımlı örneklerde t testi için kullanacağımız formül tek örneklemde t testine benzemektedir:

$$t = \frac{(\bar{d} - 0)}{SE(\bar{d})} = \frac{\bar{d}}{s_d / \sqrt{n}}$$

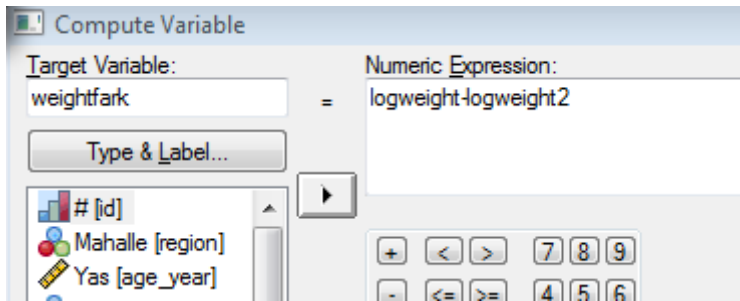
$\bar{d}$ = öncesi ve sonrasında yapılan ölçümler arasındaki farkın aritmetik ortalaması	0,0372

0 = toplum ortalaması (aradaki farkın sıfır olduğunu varsayıyoruz)	0
s_d = öncesi ve sonrasında yapılan ölçümler arasındaki farkın standart sapması	0,0614
n = örnek sayısı. (Bu değişken için mevcut veriler)	424

Buradan $t = 12,475$ olarak hesaplanır.

Not: SPSS’te (logweight-logweight2) hesabını yapıp yeni bir değişken oluşturmak ve bu değişkenin ortalamasını ($\bar{d}$) ve standart sapmasını almak için:

Transform>Compute variable>[“Target Variable” alanına “weightfark”, “Numeric expression” alanına “logweight-logweight2” girelim>ok.



Daha sonra, oluşacak yeni “weightfark” değişkeninin ortalama ve standart sapmasını alabiliriz:

Analyze>Descriptive Statistics>Descriptives>[“Variable(s)” alanına “weightfark” değişkenini geçirelim>ok.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Weightfark	424	-0,38	0,31	,04	,061
Valid N (listwise)	424				

SPSS çıktısında ortalama (0,04) ve standart sapma (0,061) değerlerinin üzerine çift tıklayarak sayıları daha ayrıntılı olarak görebiliriz.

4. Test istatistiğinden elde edilen değerin bilinen bir olasılık dağılımı ile karşılaştırılması:

Sonuçlarımız (n-1) serbestlik derecesinde t dağılımı gösterir. t tablosu ile (sayfa 459) karşılaştıracak olursak $t = 12,475$ ve serbestlik derecesi = 423 için p değerinin 0,01’den küçük olduğunu görürüz.

df	Two-tailed <i>P</i> -value			
	0.10	0.05	0.01	0.001
1	6.314	12.706	63.656	636.58
2	2.920	4.303	9.925	31.600
<i>SD</i> = 423 <i>t</i> = 12,475				
100	1.660	1.984	2.626	3.390
200	1.653	1.972	2.601	3.340
5000	1.645	1.960	2.577	3.293

5. *P* değerinin ve sonuçların yorumlanması

P değeri 0,05'ten küçük olduğu için H_0 hipotezini reddederiz.

Dolayısıyla bireylerin tedavi öncesi ve sonrasında yapılan ağırlık ölçümleri arasında istatistiksel olarak anlamlı bir fark vardır.

Ortalamalar arasındaki farkın %95 güven aralığını da hesaplayacak olursak:

$$\bar{d} \pm t_{0,05} \times (s_d / \sqrt{n})$$

($t_{0,05}$ yerine örneklemimizin serbestlik derecesinde *t* tablosunda iki yönlü *p* değerinde %5'e karşılık gelen sayıdır. Bizim tablomuzda bu kadar ayrıntı olmadığından 1,97 ile 1,96 arasında bir değer olduğunu düşünebiliriz. 1,965 alalım:

df	Two-tailed <i>P</i> -value			
	0.10	0.05	0.01	0.001
1	6.314	12.706	63.656	636.58
200	1.653	1.972	2.601	3.340
5000	1.645	1.960	2.577	3.293

$$\%95 \text{ GA} = 0,0372 \pm 1,965 \times (0,0614 / \sqrt{424})$$

$$\%95 \text{ GA: } [0,0316 - 0,0431]$$

Elde ettiğimiz değer logaritmik bir değer olduğunu unutmamalıyız. Sonucumuzu klinik olarak anlaşılır hale dönüştürmek için antilog dönüşüm uygulamalıyız:

$$\text{Antilog } (0,0316) = 10^{0,0316} = 1,08 \text{ kg. Antilog } (0,0431) = 10^{0,0431} = 1,10 \text{ kg.}$$

$$\%95 \text{ GA: } [1,08 - 1,10]$$

Bulduğumuz güven aralığı toplum ortalamasını (0) içermediğinden aradaki farkın anlamlı olduğunu söyleyebiliriz.

Örneğin SPSS ile yapılması

SPSS ile yapacak olursak:

Analyze>Compare Means >Paired-Samples T Test>[“logweight” ve “logweight2” değişkenlerini birlikte işaretleyip “Paired Variables” alanına geçirelim]>ok

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Logweight	1,86	424	,084	,004
	logweight2	1,83	424	,086	,004

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	logweight & logweight2	424	,739	,000

Paired Samples Test

		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower				Upper
Pair 1	logweight - logweight2	,037	,061	,003	,031	,043	12,486	423	,000

Varsayımlar karşılanamıyorsa

Verilerimizin bağımlı gruplarda *t*-testi için varsayımları karşılamaması halinde örneğimizde olduğu gibi veri dönüşümü uygulayabilir veya bu testin nonparametrik versiyonundan yararlanmamız gerekir.

Wilcoxon işaretli sıra testi

Mantiği

Bir önceki konuda gördüğümüz gibi, işaret testi (sign test) örneklemimizden elde ettiğimiz ortalamayı belli bir değerle (toplum ortalaması) karşılaştırmamızı sağlıyordu. İşaret testini burada da kullanabiliriz. Eşleştirilmiş örneklerdeki önce ve sonraki ölçümler arasındaki farkın sıfır olduğunu varsayarsak, hesabımız tek örneklemdeki *t*-testine benzeyecektir.

İşaret testine göre Wilcoxon işaretli sıra testi daha güçlü bir testtir. Bu test sadece farklar arasındaki işarete (artı mı eksi mi olduğuna) bakmayıp, farkın miktarını da dikkate alır.

Her iki eşleşmiş veri arasındaki fark hesaplanır. Farkı sıfır olan veriler dikkate alınmaz. Daha sonra bu farklar artı (+) veya eksi (-) olmalarına göre gruplandırılır. Ayrıca, elde edilen veriler işaretine bakılmaksızın (mutlak değerine göre) küçükten büyüğe doğru sıraya konur. En küçük farka 1, sıradakine 2 vs. numara verilir. En büyük farkın sırasına n' denir. İki veya daha fazla farkın eşit olması halinde sıralarının ortalaması alınır (örn. 0,4 – 1,2 – 1,2 – 1,2 – 3,5 için sıralama 1 – 3 – 3 – 3 – 5 şeklinde olur). Sıfır hipotezinin doğru olabilmesi için artı ve eksi farkların sıralamaların toplamı sıfır olmalıdır.

Bir örnek üzerinden Wilcoxon işaretli sıra testi

Yukarıdaki örneğe benzer bir çalışma yaptığımızı ama daha küçük bir örneklemeden veri topladığımızı düşünelim:

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması

H_0 : Bireylerin uygulanan beslenme ve egzersiz programı öncesi ve sonrasındaki vücut ağırlıkları arasında fark yoktur

H_1 : Bireylerin uygulanan beslenme ve egzersiz programı öncesi ve sonrasındaki vücut ağırlıkları arasında fark vardır.

2. Verilerin toplanması

Birey no	Tedavi öncesi ağırlık (kg)	Tedavi sonrası ağırlık (kg)
1	95,0	80,75
2	78,0	78,00
3	55,0	56
4	79,7	69,19
5	56,0	56,00
6	55,5	56
7	120,0	108
8	120,0	110
9	94,0	83
10	76,0	77
11	81,9	69,61
12	65,0	55,25
13	95,7	81,35
14	74,5	63,33
15	65,0	55,25

16	80,0	68,00
17	75,5	64,18
18	86,0	73,10
19	100,0	85,00
20	64,0	74,80
21	73,0	62,05
22	85,0	72,25
23	65,0	65
24	72,0	73
25	73,0	62,05
26	65	68
27	76,0	64,60
28	94,0	79,90
29	91,5	77,78
30	61,5	52,28
31	54	65
32	88,0	74,80

3. İlgili sıfır hipotezi için *test istatistiğinin* hesaplanması

Örnekleminizde 33 birey var. Ağırlık değişkenimiz için tedavi öncesi ve sonrası farkları hesaplayalım:

Birey no	Tedavi öncesi ağırlık (kg)	Tedavi sonrası Ağırlık (kg)	Öncesi-Sonrası
1	95,0	80,75	-14,25
2	78,0	78,00	0
3	55,0	56	1
4	79,7	69,19	-10,51
5	56,0	56,00	0
6	55,5	56	0,5
7	120,0	108	-12
8	120,0	110	-10

9	94,0	83	-11
10	76,0	77	1
11	81,9	69,61	-12,29
12	65,0	55,25	-9,75
13	95,7	81,35	-14,35
14	74,5	63,33	-11,17
15	65,0	55,25	-9,75
16	80,0	68,00	-12
17	75,5	64,18	-11,32
18	86,0	73,10	-12,9
19	100,0	85,00	-15
20	64,0	74,80	10,8
21	73,0	62,05	-10,95
22	85,0	72,25	-12,75
23	65,0	65	0
24	72,0	73	1
25	73,0	62,05	-10,95
26	65	68	3
27	76,0	64,60	-11,4
28	94,0	79,90	-14,1
29	91,5	77,78	-13,72
30	61,5	52,28	-9,22
31	54	65	11
32	88,0	74,80	-13,2

Şimdi hesapladığımız farkları mutlak değerlerine göre sıraya koyalım (sıfır olan farkları atacağız):

Öncesi-Sonrası	Sıra no
0,5	1
1	3
1	3

1	3
3	5
-9,22	6
-9,75	7,5
-9,75	7,5
-10	9
-10,51	10
10,8	11
-10,95	12,5
-10,95	12,5
-11	14,5
11	14,5
-11,17	16
-11,32	17
-11,4	18
-12	19,5
-12	19,5
-12,29	21
-12,75	22
-12,9	23
-13,2	24
-13,72	25
-14,1	26
-14,25	27
-14,35	28
-15	29 (= n')

Şimdi pozitif (T₊) ve negatif (T₋) farkların sıra değerlerini toplayalım:

T ₋ =	6+7,5+7,5+9+10+12,5+12,5+14,5+16+17+18+19,5+19,5+21+22+23+24+25+26+27+28+29	= 394,5
T ₊ =	1+3+3+3+5+11+14,5	= 40,5

$n' \leq 25$ olması durumunda T testi değeri T_+ veya T_- 'den hangisi küçükse o olur.

$n' > 25$ olması durumunda z hesabı yapılır

$$z = \frac{\left| T - \frac{n'(n'+1)}{4} \right| - \frac{1}{2}}{\left(\frac{n'(n'+1)(2n'+1)}{24} \right)}$$

Bizim örneğimizde $n' = 27$ değeri 10'dan büyük olduğuna göre z formülünü uygulamalıyız:

$$z = \left[\left| 40,5 - \frac{(29 \times 30)}{4} \right| - \frac{1}{2} \right] / \left[\frac{(29 \times 30) \times (58 + 1)}{24} \right] = -3,828.$$

4. Test istatistiğinden elde edilen değerin bilinen bir olasılık dağılımı ile karşılaştırılması

$n' \leq 25$ olması durumunda T testi değeri ayfa 459'den karşılaştırılır

$n' > 25$ olması durumunda z değeri sayfa 458'den bakılır.

z	2-yönlü P-değeri
0.0	1.000
0.1	0.920
3.3	0.001
3.4	0.001
3.5	0.000

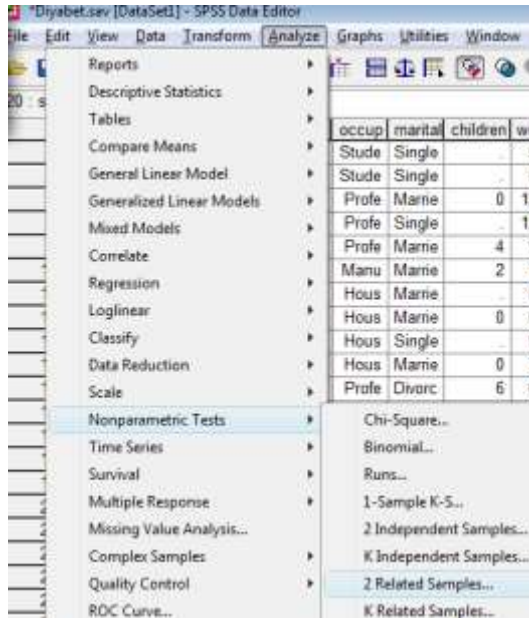
z tablosundan $z = 3,8$ için iki yönlü p değerinin **<0,05** olduğunu görüyoruz.

5. P değerinin ve sonuçların yorumlanması

P değeri 0,05'ten küçük olduğundan H_1 hipotezini kabul etmemiz gerekir. Dolayısıyla tedavi öncesi ve sonrasındaki kilo farkları arasında istatistiksel olarak anlamlı bir fark vardır.

Wilcoxon işaretli sıra testinin SPSS ile yapılması

Analyze>Nonparametric Tests >2 Related Samples>[“weight” ve “weight2” değişkenlerini işaretleyip “Test Pair(s) List” alanına geçirelim]>ok.



SPSS çıktısı aşağıdaki gibi olacaktır:

Ranks

		N	Mean Rank	Sum of Ranks
Diyet ve egzersizden sonraki kilo - Agirlik	Negative Ranks	280(a)	159,70	44715,50
	Positive Ranks	42(b)	173,51	7287,50
	Ties	102(c)		
	Total	424		

a Diyet ve egzersizden sonraki kilo < Agirlik

b Diyet ve egzersizden sonraki kilo > Agirlik

c Diyet ve egzersizden sonraki kilo = Agirlik

Test Statistics(b)

	Diyet ve egzersizden sonraki kilo – Agirlik
Z	-11,194(a)
Asymp. Sig. (2-tailed)	,000

a Based on positive ranks.

b Wilcoxon Signed Ranks Test

Bölüm 25: Nümerik Veri İki Bağımsız Grup (Student t Testi, Mann-Whitney U Testi)

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun iki bağımsız gruptan elde edilen nümerik verilerin ortalamalarının karşılaştırılmasında kullanılan bağımsız örneklemelerde *t*-testi (independent samples *t*-test veya Student *t* testi) ve bunun nonparametrik alternatifleri olan Wilcoxon sıra toplamı testi (Wilcoxon rank sum test) ve Mann-Whitney U testi hakkında bilgi sahibi olması ve bu testleri yapabilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Nümerik veri, iki bağımsız grubu tanımlayabilmeli
- Student *t* testi varsayımlarını sayabilmeli
- Boy ve ağırlık değişkenlerini kullanarak SPSS ile vücut kitle endeksini hesaplayabilmeli
- SPSS’te normal dağılımı test edebilmeli
 - Histogram
 - Kolmogorov-Smirnov
 - Skewness
- Veri dönüştürmesi uygulayabilmeli
- SPSS’te Student *t* testi yapabilmeli
- Student *t* testi SPSS çıktısını yorumlayabilmeli
- SPSS’te Mann-Whitney U testi yapabilmeli
- Mann-Whitney U testi SPSS çıktısını yorumlayabilmeli

Diyabet hastalarımızın vücut kitle indekslerinin (VKİ) erkekler ve bayanlar arasında farklılık gösterip göstermediğini araştırmak istiyoruz (H₀: Diyabet hastası erkekler ve bayanlar arasında vücut kitle indeksleri açısından fark yoktur). Ölçmek istediğimiz değişken (VKİ) nümerik sürekli bir değişkendir. Bayanlar ve erkeler olmak üzere birbirinden bağımsız iki grubumuz var.

Yukarıdaki hipotezimizi test etmek için uygulayabileceğimiz test bağımsız gruplarda *t* testidir (independent samples *t* test veya Student *t* test). Bu testi uygulayamamamız halinde nonparametrik alternatifi olan Wilcoxon sıra toplamı testi (Wilcoxon rank sum test) veya Mann-Whitney U testini yapabiliriz.

Bağımsız örneklemelerde *t* testi

Varsayımları

- Değişken, toplumda normal dağılmalı
- Değişkenin varyansları her iki grupta eşit olmalı
- Örneklem sayısı normal dağılımı ve varyansların eşitliğini test edebilecek kadar yeterli olmalı.

Mantığı

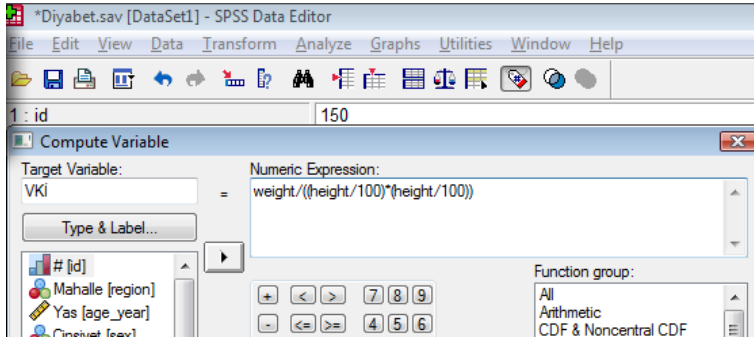
İki grupta yapılan ölçümlerin ortalamasını karşılaştırıyoruz. Sıfır hipotezimize göre iki grubun ortalamaları arasında fark olmaması gerekir ($\text{Ortalama}_1 - \text{Ortalama}_2 = 0$). Student *t* testi daha önce gördüğümüz tek örneklemde *t* testi ve bağımlı örneklemelerde *t* testi gibi ortalamalar arasındaki farka dayanmaktadır. Bu testin verileri teorik dağılımlarından *t* dağılımına benzemektedir.

Not

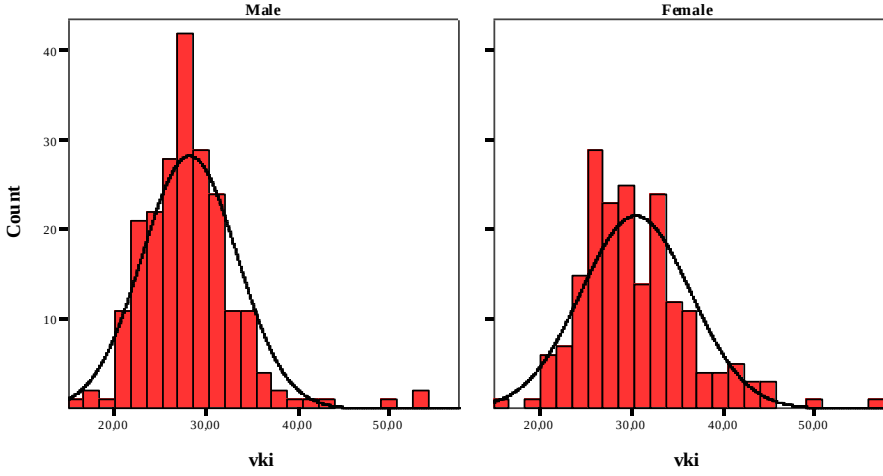
İki grubumuzun örneklemeleri n_1 ve n_2 , ortalamaları $\bar{x}_1$ ve $\bar{x}_2$, standart sapmaları ise s_1 ve s_2 'dir.

Diyabet.sav veri setimizde boy ve ağırlık değişkenleri olmakla birlikte VKİ hesaplanmamıştır. Aşağıdaki komutları kullanarak yeni bir VKİ değişkeni hesaplayabiliriz ($\text{VKİ} = [\text{ağırlık (kg)}] / [\text{boy (m)}]^2$) olduğunu hatırlayınız.

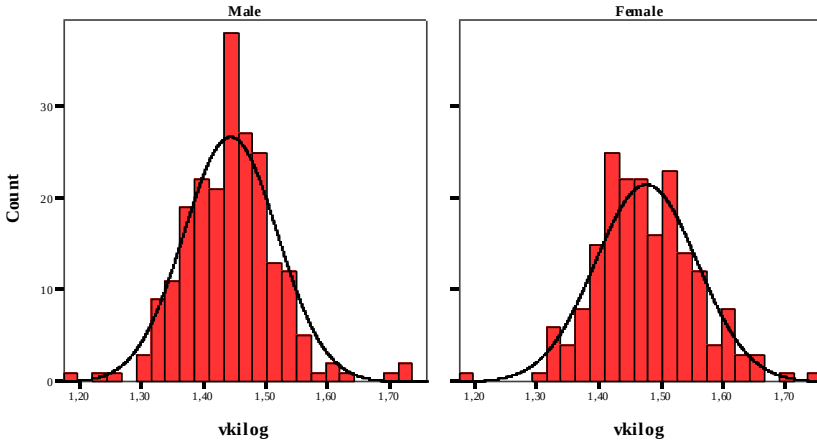
Transform>Compute variable>[“Target Variable” alanına “VKİ”, “Numeric expression” alanına “weight/((height/100)(height/100))” girelim>ok.*



Şimdi, bağımlı değişkenimizin Student *t* testi varsayımlarını karşılayıp karşılamadığına bakabiliriz. Karşılaştıracığımız gruplarda (erkek ve kadınlar) VKİ için histogram grafiği aşağıdaki gibidir:



Bu grafikten verilerin sağa eğimli olduğunu görüyoruz. Logaritmik dönüşüm yaptıktan sonra histogram aşağıdaki gibi olmaktadır:



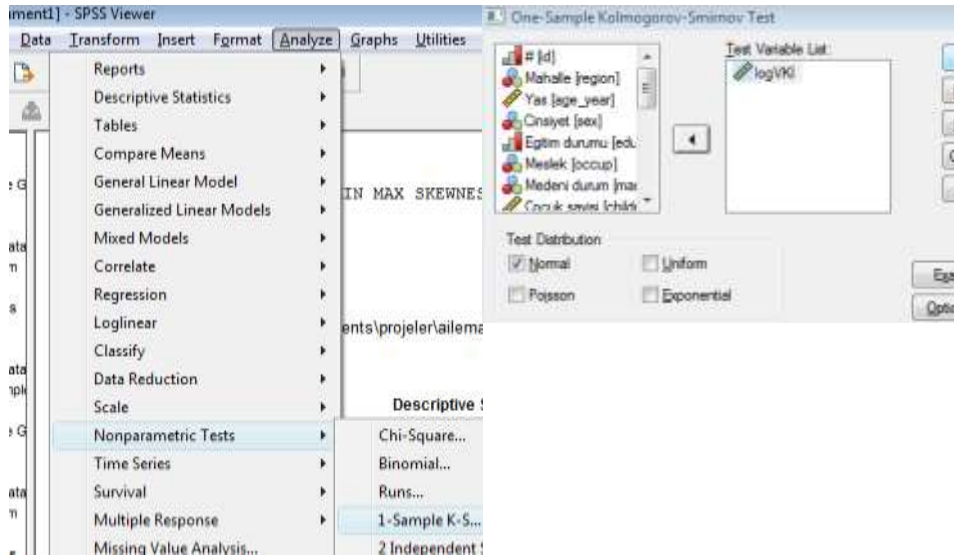
Yeni değişkenimizin (logVKİ) skewness değerine de baktığımızda normal dağılıma yaklaştığını görüyoruz:

Statistics

vkiolog			
Male	N	Valid	215
		Missing	21
		Skewness	,365
		Std. Error of Skewness	,166
Female	N	Valid	189
		Missing	5
		Skewness	,204
		Std. Error of Skewness	,177

Normal dağılımı ayrıca Kolmogorov Smirnov testi ile de ölçebiliriz:

Analyze>Nonparametric Tests>1-Sample K-S>[“logVKI” değişkenini “Test Variable List” alanına geçirelim>ok.



Aşağıdaki çıktıyı elde ederiz:

One-Sample Kolmogorov-Smirnov Test

Cinsiyet		vkilog
Male	N	215
	Normal Parameters ^{a,b}	Mean
		1,442
	Most Extreme Differences	Std. Deviation
		,0755
		Absolute
	Positive	,073
		Negative
Female	Kolmogorov-Smirnov Z	,040
		1,070
	Asymp. Sig. (2-tailed)	,202
	N	189
	Normal Parameters ^{a,b}	Mean
		1,474
	Most Extreme Differences	Std. Deviation
		,0824
		Absolute
	Positive	,046
		,043
	Negative	-,046
	Kolmogorov-Smirnov Z	,626
		Asymp. Sig. (2-tailed)
		,829

a. Test distribution is Normal.

b. Calculated from data.

Görüldüğü gibi, Kolmogorov-Smirnov testi p değeri hem erkekler, hem de kadınlar için $>0,05$ 'tir. Buradan, verilerimizin normal dağılımdan farklı olmadığı sonucunu çıkarabiliriz.

Şimdi, “logVKİ” değişkenimizi kullanarak Student t testini uygulayalım:

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

H_0 : Diyabet hastası erkekler ve bayanlar arasında vücut kitle indeksleri açısından fark yoktur

H_1 : Diyabet hastası erkekler ve bayanlar arasında vücut kitle indeksleri açısından fark vardır.

2. Verilerin toplanması:

Verilerimizi topladık ve SPSS’e girdik. “Compute” komutunu kullanarak VKİ değişkenini hesapladık. Logaritmik dönüşüm uygulayarak “logVKİ” değişkenini oluşturduk.

3. İlgili sıfır hipotezi için *test istatistiğinin* hesaplanması:

Student t testi için kullanacağımız formül de önceki t testi uygulamalarına benzemektedir:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{SE(\bar{x}_1 - \bar{x}_2)} = \frac{(\bar{x}_1 - \bar{x}_2)}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

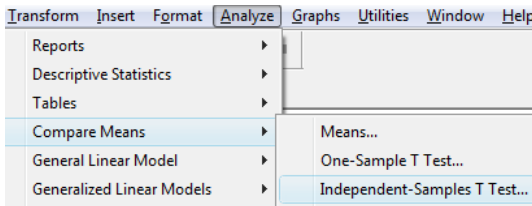
Burada s_1 ve s_2 'den “pooled standard deviation” s hesaplanması gerekmektedir:

$$s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

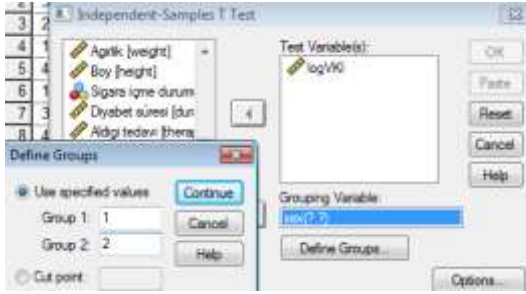
$\bar{x}_1$ (Erkek)	1,44
$\bar{x}_2$	1,47
s_1	0,076
s_2	0,082
n_1	215
n_2	189

Yukarıdaki formülü hesap makinesiyle yapmak zordur. Doğrudan SPSS ile uygulamaya geçilecektir:

Analyze>Compare Means>Independent-Samples t Test>[“logVKİ” değişkenini “Test Variable(s)” alanına, “sex” değişkenini de “Grouping Variable” alanına geçirelim.



>Define Groups>[Group 1'e 1, Group 2'ye de 2 girelim]>Continue>ok.



Aşağıdaki çıktıları elde ederiz:

Group Statistics

	Cinsiyet	N	Mean	Std. Deviation	Std. Error Mean
logVKİ	Male	215	1,44	,076	,005
	Female	189	1,47	,082	,006

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
logVKY	Equal variances assumed	3,178	,075	-4,127	402	,000	-,032	,008	-,048	-,017
	Equal variances not assumed			-4,104	384,162	,000	-,032	,008	-,048	-,017

4. Test istatistiğinden elde edilen değerin bilinen bir olasılık dağılımı ile karşılaştırılması:

İkinci tablodan görüldüğü gibi Levene testine göre varyanslar arasında fark yoktur (Sig. = 0,075; >%5). t istatistiği **-4,127** olarak hesaplanmıştır. Serbestlik derecesi (n_1+n_2-2) = 402 olarak alınmıştır. İki yönlü p değeri sıfıra yakındır (asıl değeri 0,000004462).

5. P değerinin ve sonuçların yorumlanması

P değeri 0,05'ten küçük olduğu için H_0 hipotezini reddederiz. Dolayısıyla diyabetik erkeklerin VKİ değerlerinin diyabetik bayanlarınkinden daha düşük olduğu sonucuna varabiliriz.

Farkların %95 güven aralığına da bakıldığında $VKİ_{erkek}-VKİ_{bayan}$ farkının hep eksi olduğu, sıfırı içermediği görülmektedir.

Elde ettiğimiz değerin logaritmik bir değer olduğunu unutmamalıyız. Sonucumuzu klinik olarak anlaşılır hale dönüştürmek için antilog dönüşüm uygulamalıyız:

Fark ($\log VKİ_{Erkek} - \log VKİ_{Kadın}$): -0,032. %95 GA [-0,048 – -0,017]

Antilog (0,032) = $10^{0,032} = 1,07 \text{ kg/m}^2$

Antilog (0,048) = $10^{0,048} = 1,12 \text{ kg/m}^2$.

Antilog (0,017) = $10^{0,017} = 1,04 \text{ kg/m}^2$.

Dolayısıyla, rapora şu şekilde yazabiliriz:

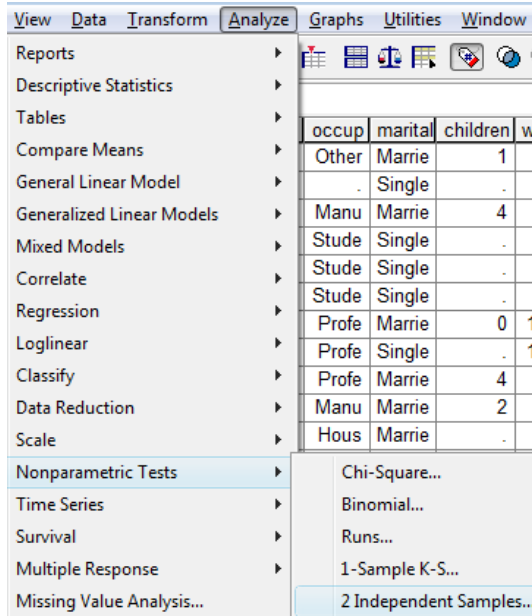
Ortalama Fark = -1,07 kg/m²; %95 GA: [-1,04 – -1,12]

Varsayımlar karşılanamıyorsa

Verilerimizin bağımsız gruplarda *t*-testi için varsayımları karşılamaması halinde örneğimizde olduğu gibi veri dönüşümü uygulayabilir veya bu testin nonparametrik versiyonundan yararlanmamız gerekir. Student *t* testinin nonparametrik alternatifleri Wilcoxon sıra toplamı testi (Wilcoxon rank sum test) ve Mann-Whitney U testidir. Mann-Whitney-U testini elle yapmak daha zordur. Bu iki testin sonuçları da hemen hemen aynıdır.

Şimdi VKİ değişkenini kullanarak Mann-Whitney U testini yapalım:

Analyze>Nonparametric Tests >2 Independent Samples>[“VKİ” değişkenini “Test Variable(s)” alanına, “sex” değişkenini de “Grouping Variable” alanına geçirelim>Define Groups>[Group 1’e 1, Group 2’ye de 2 girelim]>Continue>ok.



Aşağıdaki çıktıları elde ederiz:

Ranks

	Cinsiyet	N	Mean Rank	Sum of Ranks
VKI	Male	215	180,60	38828,50
	Female	189	227,42	42981,50
	Total	404		

Test Statistics(a)

	VKI
Mann-Whitney U	15608,500
Wilcoxon W	38828,500
Z	-4,021
Asymp. Sig. (2-tailed)	,000

a. Grouping Variable: Cinsiyet

Mann-Whitney U testi sonucunda da z testi değeri **-4,021** ve p değeri **<0,001** olduğundan H_0 hipotezini reddederiz.

Bölüm 26: Nümerik Veri İkidenden Fazla Bağımlı Grup (Tekrarlayan Ölçümlerde ANOVA, Friedman Varyans Analizi)

Hamit ACEMOĞLU

Amaç: Bu konu sonunda okuyucunun ikiden fazla bağımlı gruptan elde edilen nümerik verilerin ortalamalarının karşılaştırılmasında kullanılan tekrarlayan ölçümlerde varyans analizi (repeated measures ANOVA) ve bunun nonparametrik alternatifi olan Friedman varyans analizi hakkında bilgi sahibi olması ve bu testleri SPSS ile yapabilmesi amaçlanmıştır.

Hedefler:

- Nümerik veri, ikiden fazla bağımlı grubu tanımlayabilmeli
- Repeated measures ANOVA varsayımlarını sayabilmeli
- SPSS’te Repeated measures ANOVA testi yapabilmeli
- Repeated measures ANOVA testi SPSS çıktısını yorumlayabilmeli
- SPSS’te Friedman varyans analizi yapabilmeli
- Friedman varyans analizi SPSS çıktısını yorumlayabilmeli

Senaryo 1, Tek faktör (grup içi karşılaştırma)

Diyabetik hastalarımıza ilaç, beslenme ve egzersizden oluşan tedavi programı uyguladığımızı varsayalım. Program süresince katılımcıların Hemoglobin A1c düzeylerini araştırmak istiyoruz (H_0 : Bireylerin uygulanan tedavi programı öncesi ve sonrasındaki Hemoglobin A1c düzeyleri arasında fark yoktur). Hemoglobin A1c ölçümleri tedavi programı uygulamadan önce, programın 3. ve 6. aylarında yapılmış olsun. Ölçmek istediğimiz değişken nümeriktir. Uygulama öncesi ve sonrasında elde edilen üç ölçümler ise bağımlıdır. Hatırlayalım:

- Aynı bireylerde herhangi bir girişim, tedavi vs. öncesi ve sonrasında yapılan ölçümler birbirinden bağımlıdır
- Bireyler farklı olsa da iki grup arasında bire bir eşleşme yapılmışsa (örn. eşleştirilmiş vaka kontrol çalışmalarında) yine eşleşen bireylerin ölçümleri birbirine yakın çıkacağından bağımlı olarak değerlendirilir.

Bizim örneğimizde bağımlı üç nümerik ölçüm söz konusudur. Kitabın başında da verilen istatistik testlerin uygulama yerleri akış şemalarına baktığımızda seçeceğimiz testin 2’den fazla bağımlı örneklemelerde tekrarlayan ölçümlerde varyans analizi (repeated measures ANOVA) veya Friedman varyans analizi olacağını görürüz.

Varsayımları

Faktöriyel deneysel tasarımlar için F testi kullanıldığında aşağıdaki varsayımlar geçerlidir:

1. Sonuç ölçütü (outcome, bağımlı değişken) nümerik olmalıdır.
2. Kalıntılar (residuals) ortalaması sıfır, varyansı sabit olan bir normal dağılıma sahip olmalıdır
3. Bireyler bağımsız olmalıdır
Grup içi (within-subject) karşılaştırmanın yapıldığı bir tasarımda (tekrarlayan ölçümler gibi) ölçümler bağımsız olmadığından 3. madde birey içi karşılaştırmalar için değişir. Gruplar arası (between subjects) karşılaştırmalarda bu varsayım yine de geçerlidir.
4. Bütün gruplar arası karşılaştırmalar için grup içi kovaryans matrisleri eşit olmalıdır. Bu varsayım kovaryans matrislerinin gruplar arasında sabit olduğunu belirtir.
5. F testi kullanıldığında grup içi kovaryans matrislerinin küresel olduğu varsayılır. Bunu saptamanın bir yolu, herhangi iki ölçüm arasındaki varyans farklılıklarının tüm ölçümlerde sabit olduğunun gösterilmesidir. Ölçümler arasındaki zaman farkı arttıkça varyanslar da farklılaşacağından bu varsayım çok kez sağlanamaz. Şu testler için bu varsayıma gerek yoktur: Wilks' lambda, Pillai-Bartlett trace ve Hotelling-Lawley trace.

Tekrarlayan ölçümlerde varyans analizi, bağımsız gruplarda olduğu gibi, bağımsız değişken sayısının bir olduğu koşullarda tek yönlü varyans analizi (örn. cinsiyetin kan trigliserit düzeyine etkisi), birden çok olduğu koşullarda iki yönlü varyans analizi (örn. Cinsiyet ve mesleğin kan trigliserit düzeyine etkisi) çok yönlü varyans analizi uygulanır.

- İncelenen toplumda değişken, grupların hepsinde normal dağılmalıdır.
- Değişkenin gruplar arasındaki varyansı eşit olmalıdır.
- Örneklem sayısı bu varsayımları kontrol edecek kadar büyük olmalıdır.

Testin mantığı

Bu test grup içi ve gruplar arası varyansların oranına dayanmaktadır. Sıfır hipotezinin doğru olması halinde değişkenin grupların genelindeki varyasyonu ile gruplar arasındaki varyasyonu aynı olmalıdır.

Örnek uygulama 1

Şimdi, diyabet.sav verisetimizde “HbA1c_1, HbA1c_2 ve HbA1c_3” değişkenlerimizin zaman içinde (başlangıç, 3. ve 6. aylar) ortalamalarındaki

değişim açısından durumunu incelemek için tekrarlayan ölçümlerde varyans analizi (repeated measures ANOVA) testini uygulayalım:

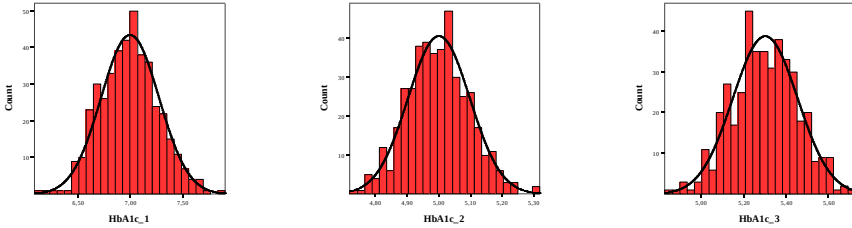
1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

H_0 : Bireylerin uygulanan tedavi programı öncesi ve sonrasındaki 3. ve 6. aylardaki vücut ağırlıkları ortalamaları arasında fark yoktur

H_1 : Bireylerin uygulanan tedavi programı öncesi ve sonrasındaki Hemoglobin A1c ölçüm ortalamalarından en az birisi diğerlerinden farklıdır.

2. Verilerin toplanması:

Veriler diyabet.sav verisetinde “HbA1c_1”, “HbA1c_2” ve “HbA1c_3” değişkenlerinin altında kaydedildi. Örneklem sayısı yeterlidir. Değişkenimizin gruplar açısından histogram grafiğine bakıldığında normal dağıldığı görülmektedir:



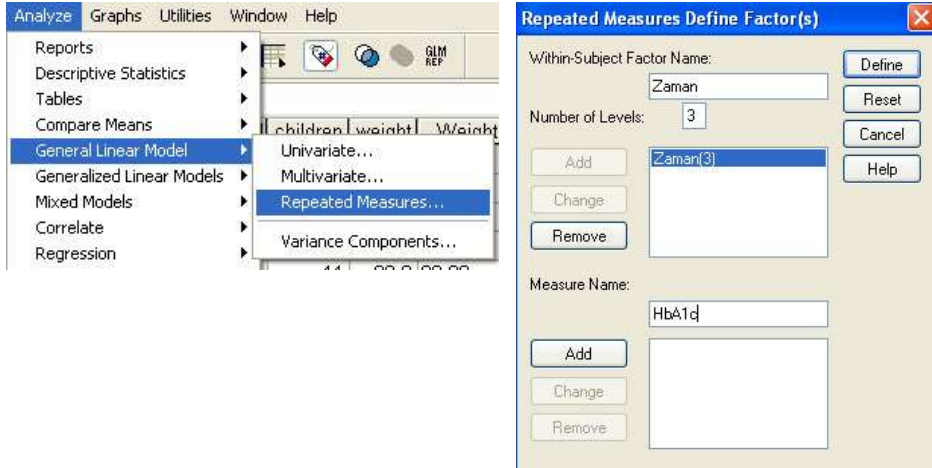
Varyansların homojenliği varsayımı ANOVA testi sırasında incelenecektir.

3. İlgili sıfır hipotezi için *test istatistiğinin* hesaplanması:

ANOVA testi grup içi ve gruplar arası varyansların birbirine oranlanmasından (F oranı) oluşur. Bu oran F dağılımına uymaktadır.

SPSS ile yapacak olursak:

Analyze > General Linear Model > Repeated Measures>[Within-Subject Factor Name kutusuna, birden çok kez ölçülmüş olan bağımlı değişkeni; zaman (ay) ve Number of Levels kutusuna da ölçüm sayısı (3) yazılmalıdır. Add butonunu klikleyerek listeye alalım. Measure name kutusuna değişkenin adını örneğimizde “HbA1c” yazıp add butonunu klikleyerek listeye alalım. Daha sonra Define butonuna basıp “Repeated Measures” diyalog kutusunu açalım.



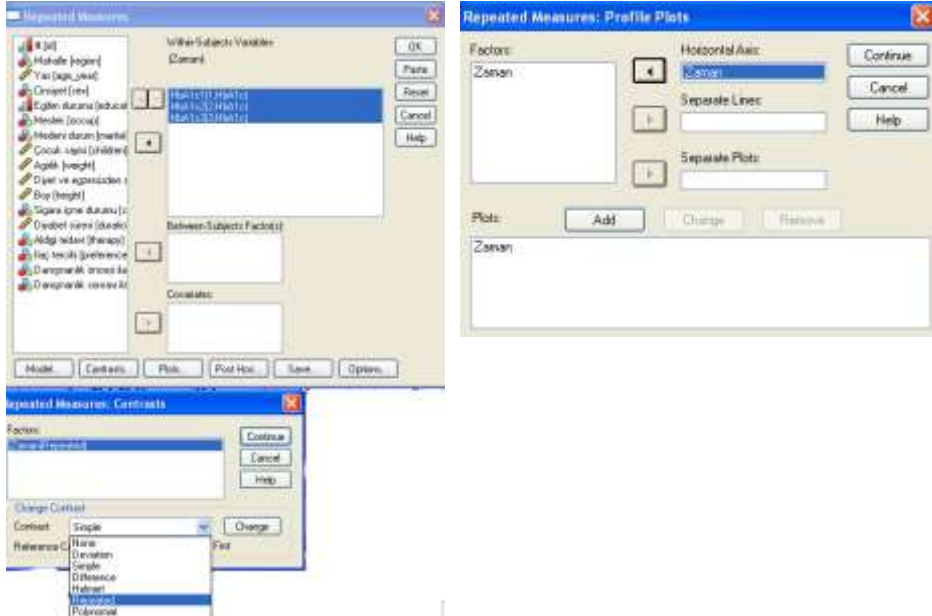
Within-Subjects variables kutusuna HbA1c_1, HbA1c_2, HbA1c_3 değişkenlerini taşıyalım.

Contrast menüsü simple seçeneğini tercih edebileceğimiz gibi bu örnekte repeated seçeneği (birinci ile ikinci ikinci ile üçüncü ölçümler karşılaştırılır) seçilebilir.

Grafik çizimi için;

>Contrast
>Repeated>Change>Continue>

>Plots> [Horizontal Axis kutucuğuna
“zaman”ı alalım]>Add>Continue



>Options> [Descriptive statistics kutucuğunu işaretleyelim]>Continue>Ok

Aşağıdaki çıktıları elde ederiz:

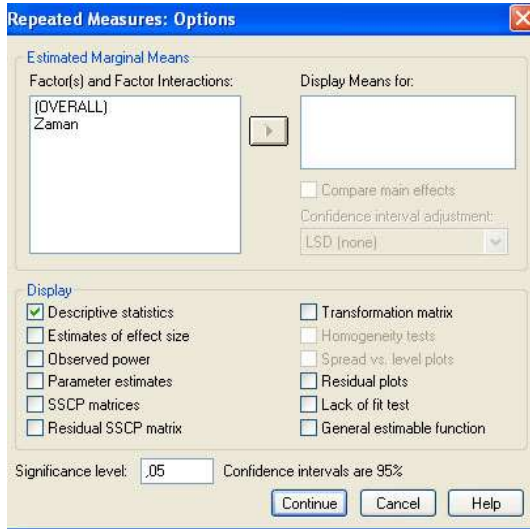
Within-Subjects Factors tablosu HbA1c değişkenine ait kaç ölçüm yapıldığını, Descriptive Statistics tablosu HbA1c'ye ait 3 ölçümün ortalama standart sapmaları verilmektedir. Bu tablo incelenerek zaman içindeki değişim hakkında bir fikir sahibi olabiliriz. İkinci ölçüm birinci ve üçüncü ölçüm ortalamasından düşüktür.

Multivariate Tests (b) tablosu üç farklı zamanda yapılan ölçümleri dört farklı istatistiksel yöntemle test etmektedir. En az bir ölçümün diğer ölçümlerden farklı olduğu görülmektedir ($p<0,001$).

Mauchly's Test of Sphericity (b) (Mauchly'nin küresellik testi) incelendiğinde $p<0,05$ olduğu için küresellik varsayımı sağlanamamıştır. Küresellik varsayımı sağlanmadığında Tests of Within-Subjects Effects tablosunda Greenhouse-Geisser veya diğer düzeltme sonuçları kullanılabilir.

Tests of Within-Subjects Contrasts tablosu repeated seçeneği seçilmiş olduğu için her ölçümü bir sonraki ölçümle karşılaştırmıştır. Buna göre birinci ölçüm ikinciden, ikinci ölçüm üçüncüden istatistiksel olarak farklıdır ($p<0,001$).

Profile Plots HbA1c ölçüm ortalamaları çizgi grafik olarak verilmiştir.



General Linear Model**Within-Subjects Factors**

Measure: HbA1c

Zaman	Dependent Variable
1	HbA1c1
2	HbA1c2
3	HbA1c3

Descriptive Statistics

	Mean	Std. Deviation	N
HbA1c1	6,99	,28	430
HbA1c2	4,99	,10	430
HbA1c3	5,30	,15	430

Multivariate Tests(b)

Effect		Value	F	Hypothesis df	Error df	Sig.
Zaman	Pillai's Trace	,979	10140,635(a)	2,000	428,0	,000
	Wilks' Lambda	,021	10140,635(a)	2,000	428,0	,000
	Hotelling's Trace	47,386	10140,635(a)	2,000	428,0	,000
	Roy's Largest Root	47,386	10140,635(a)	2,000	428,0	,000

a Exact statistic

b Design: Intercept

Within Subjects Design: Zaman

Mauchly's Test of Sphericity(b)

Measure: HbA1c

Within Subject s Effect	Mauchly's W	Approx . Chi-Square	df	Sig.	Epsilon(a)		
	Greenhous e-Geisser	Huynh-Feldt	Lower - bound	Greenhous e-Geisser	Huynh -Feldt	Lower - bound	Greenhous e-Geisser
Zaman	,704	150,31	2	,000	,772	,774	,500

		1					

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b Design: Intercept

Within Subjects Design: Zaman

Tests of Within-Subjects Effects

Measure: HbA1c

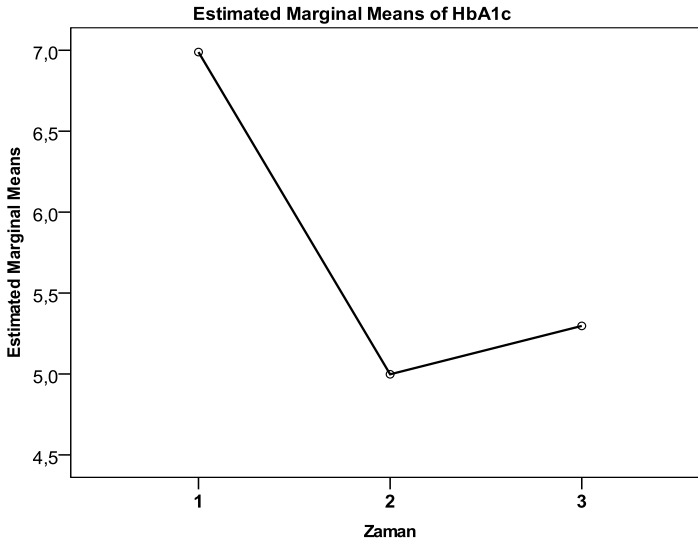
Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Zaman	Sphericity Assumed	991,215	2	495,608	13627,043	,000
	Greenhouse-Geisser	991,215	1,543	642,384	13627,043	,000
	Huynh-Feldt	991,215	1,548	640,515	13627,043	,000
	Lower-bound	991,215	1,000	991,215	13627,043	,000
Error (Zaman)	Sphericity Assumed	31,205	858	,036		
	Greenhouse-Geisser	31,205	661,959	,047		
	Huynh-Feldt	31,205	663,890	,047		
	Lower-bound	31,205	429,000	,073		

Tests of Within-Subjects Contrasts

Measure: HbA1c

Source	Zaman	Type III Sum of Squares	df	Mean Square	F	Sig.
Zaman	Level 1 vs. Level 2	1704,299	1	1704,299	20246,387	,000
	Level 2 vs. Level 3	38,423	1	38,423	1123,456	,000
Error(Zaman)	Level 1 vs. Level 2	36,112	429	,084		
	Level 2 vs. Level 3	14,672	429	,034		

Profile Plots



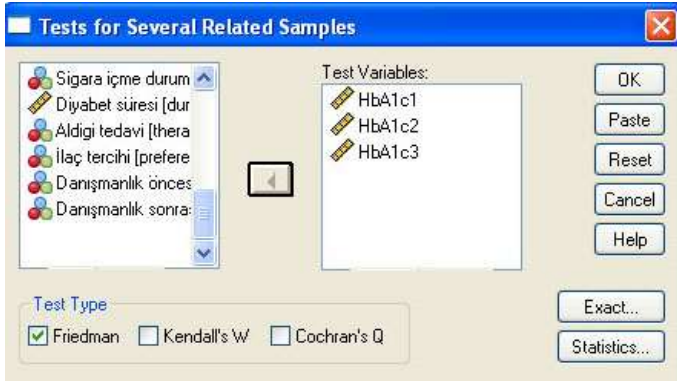
Varsayımlar karşılanamıyorsa

Verilerimizin bağımlı gruplarda tekrarlayan ölçüm için varsayımları karşılamaması halinde veri dönüşümü uygulayabilir veya bu testin nonparametrik versiyonu olan Friedman Varyans analizinden yararlanmamız gerekir.

Şimdi HbA1c değişkenini kullanarak Friedman Varyans analizini yapalım:

Analyze > Nonparametric Tests > K Related Samples > [Test Variables alanına HbA1c1, HbA1c2, HbA1c3 taşıyalım] >OK

Aşağıdaki çıktıları elde ederiz: Test Statistics(a) tablosunda $p < 0,001$ olduğundan üç farklı zamanda ölçümlerden en az birisinin diğerlerinden farklı olduğunu görürüz. Hangi grup veya grupların diğerlerinden farklı olduğunu anlamak için Wilcoxon testi ile İkişerli karşılaştırmalar yapılabilir (1. Ölçüm ile 2. ölçüm 1. Ölçüm ile 3. Ölçüm, 2. Ölçüm ile 3. Ölçüm). Ancak istatistiksel önemlilik düzeyi Bonferroni düzeltmesine göre alınmalıdır. Bonferroni düzeltmesi, önemlilik düzeyi $p < 0,05$ olarak kabul edildiği durumda, “0.05/ikili karşılaştırma sayısı” olarak belirlenir. Örneğimizde, 3 grup arasında farkın hangi grup veya gruplardan kaynaklandığını belirlerken anlamlılık düzeyi $0,05/3=0,016$ olarak kabul edilir.



NPar Tests

Friedman Test

	Mean Rank
HbA1c1	3,00
HbA1c2	1,06
HbA1c3	1,94

Ranks

Test Statistics(a)

N	430
Chi-Square	812,907
Df	2
Asymp. Sig.	,000

a. Friedman Test

Senaryo 2, İki faktör (grup içi ve gruplar arası karşılaştırma)

Erkek ve kadın diyabetik hastalarımıza ilaç, beslenme ve egzersizden oluşan tedavi programı uyguladığımızı varsayalım. Program süresince cinsiyete ve zamana göre HbA1c düzeylerini araştırmak istiyoruz (H_0 : Kadın ve erkek grupları arasında ve tedavi programı öncesi ve sonrasındaki HbA1c düzeyleri arasında fark yoktur ve cinsiyet*zaman etkileşimi(interaction) yoktur). HbA1c düzeyleri arasında yalnızca cinsiyet ya da zamanla açıklanamayan, ama bu iki faktörün değişik kombinasyonları ile açıklanan farklılıklar var ise etkileşimden bahsedebiliriz.

HbA1c ölçümleri tedavi programı uygulamadan önce, programın 3. 6. ve 12. aylarında yapılmış olsun. Ölçmek istediğimiz değişken nümeriktir.

Örnek uygulama 2

Şimdi, diyabet.sav verisetimizde “HbA1c_1, HbA1c_2, HbA1c_3 ve HbA1c_4” değişkenlerimizin cinsiyete göre (Erkek, Kadın) zaman içinde (başlangıç, 3. 6.

ve 12. aylar) ortalamalarındaki değişim açısından durumunu incelemek için tekrarlayan ölçümlerde varyans analizi (repeated measures ANOVA) testini uygulayalım:

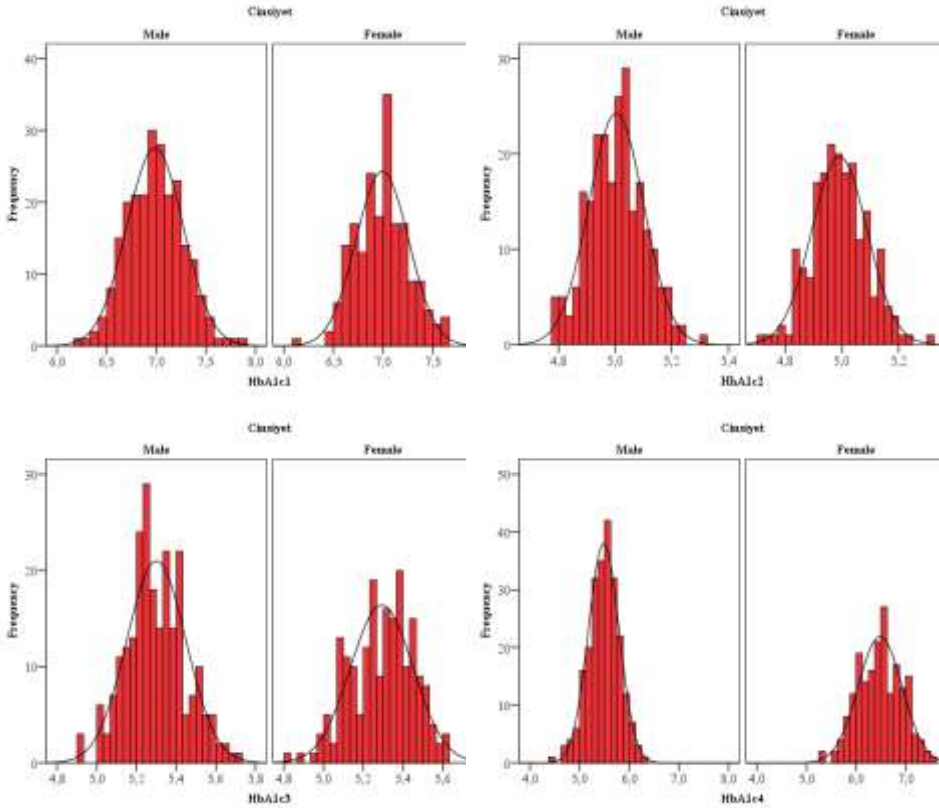
1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

H_0 : Kadın ve erkek grupları arasında ve tedavi programı öncesi ve sonrasındaki Hemogloblin A1c düzeyleri arasında fark yoktur ve/veya cinsiyet*zaman etkileşimi(interaction) yoktur

H_1 : Farklı cinsiyetteki bireylere uygulanan tedavi programı öncesi ve sonrasındaki Hemogloblin A1c ölçüm ortalamalarından en az birisi diğerlerinden farklıdır ve/veya cinsiyet*zaman etkileşimi(interaction) vardır.

2. Verilerin toplanması:

Veriler diyabet.sav verisetinde “Sex (cinsiyet)” ve “HbA1c_1”, “HbA1c_2”, “HbA1c_3” ve “HbA1c_4” değişkenlerinin altında kaydedildi. Örneklem sayısı yeterlidir. Değişkenimizin gruplar açısından histogram grafiklerine bakıldığında normal dağıldığı görülmektedir:



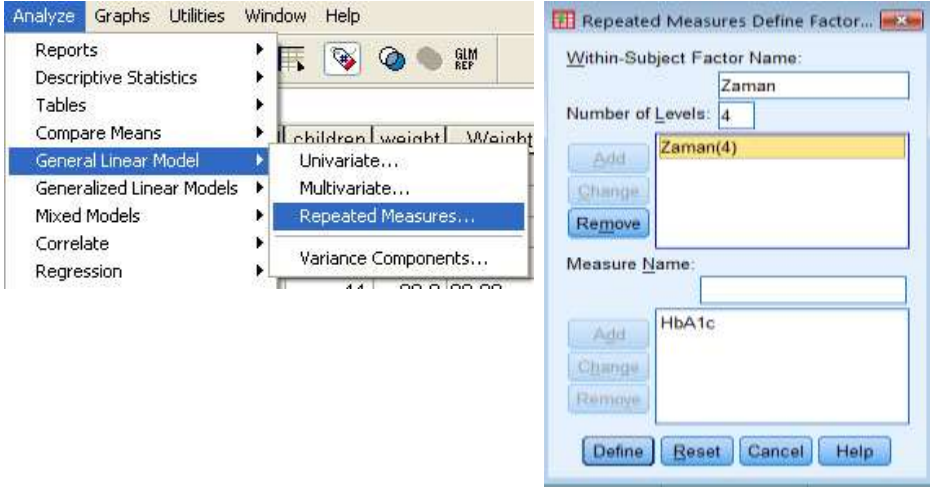
Diğer varsayımlar tekrarlayan ölçümlerde varyans analizi (repeated measures ANOVA) testi sırasında incelenecektir.

3. İlgili sıfır hipotezi için *test istatistiğinin* hesaplanması:

ANOVA testi grup içi ve gruplar arası varyansların birbirine oranlanmasından (F oranı) oluşur. Bu oran F dağılımına uymaktadır.

SPSS ile yapacak olursak:

Analyze > General Linear Model > Repeated Measures > [Within-Subject Factor Name kutusuna, birden çok kez ölçülmüş olan bağımlı değişkeni; zaman (ay) ve Number of Levels kutusuna da ölçüm sayısı (4) yazılmalıdır. Add butonunu klikleyerek listeye alalım. Measure name kutusuna değişkenin adını örneğimizde “HbA1c” yazıp add butonunu klikleyerek listeye alalım. Daha sonra Define butonuna basıp “Repeated Measures” diyalog kutusunu açalım.



Within-Subjects variables (Grup içi) kutusuna HbA1c_1, HbA1c_2, HbA1c_3, HbA1c_4 değişkenlerini taşıyalım. Between- Subjects Factor(s) (Gruplar arası) kutusuna “Sex” değişkenlerini taşıyalım.

Contrast menüsü simple seçeneğini tercih edebileceğimiz gibi bu örnekte repeated seçeneği (birinci ile ikinci ikinci ile üçüncü ölçümler karşılaştırılır) seçilebilir.

>Contrast >Repeated>Change>Continue>



Grafik çizimi için;

>Plots> [Horizontal Axis kutucuğuna “zaman” ve Separate Lines kutucuğuna “sex” alalım]>Add>Continue



>Options> [Descriptive statistics kutucuğunu işaretleyelim]>Continue>Ok

Aşağıdaki çıktıları elde ederiz: Within-Subjects Factors tablosu HbA1c değişkenine ait kaç ölçüm yapıldığını, Between-Subjects Factors tablosu cinsiyet dağılımını, Descriptive Statistics tablosu HbA1c'ye ait 4 ölçümün ortalama standart sapmalarını cinsiyete göre verilmektedir. Bu tablo incelenerek cinsiyete göre zaman içindeki değişim hakkında bir fikir sahibi olabiliriz. İkinci ölçüm birinci üçüncü ve dördüncü ölçüm ortalamasından düşüktür.

Multivariate Tests (b) tablosu üç farklı zamanda yapılan ölçümleri ve zaman*cinsiyet etkileşimini dört farklı istatistiksel yöntemle test etmektedir. En az bir ölçümün diğer ölçümlerden farklı olduğu ve zaman*cinsiyet etkileşimi olduğu görülmektedir (p=0.000).

Mauchly's Test of Sphericity (b) (Mauchly'nin küresellik testi) incelendiğinde $p < 0,05$ olduğu için küresellik varsayımı sağlanmamıştır. Küresellik varsayımı sağlanmadığında Tests of Within-Subjects Effects tablosunda Greenhouse-Geisser veya diğer düzeltme sonuçları kullanılabilir.

Tests of Within-Subjects Contrasts tablosu repeated seçeneği seçilmiş olduğu için her ölçümü bir sonraki ölçümle karşılaştırmıştır. Buna göre birinci ölçüm ikinciden, ikinci ölçüm üçüncüden istatistiksel olarak farklıdır ($p < 0,001$). Ayrıca zaman*cinsiyet etkileşiminin ilk üç ölçüm için yokken üçüncü ve dördüncü ölçümler arasında olduğu görülmektedir ($p < 0,001$).

Tests of Between-Subjects Effects tablosu erkek ve kadınların HbA1c ortalamalarının en az bir ölçümde farklı olduğunu göstermektedir ($p < 0,001$). Cinsiyet iki değişkenli olduğundan SPSS post hoc test sonuçlarını vermemektedir. Bu nedenle hangi ölçümde farklılık olduğunu bulmak için bağımsız gruplarda t testi kullanılmalıdır.

Profile Plots cinsiyete göre HbA1c ölçüm ortalamaları çizgi grafik olarak verilmiştir.

General Linear Model

Within-Subjects Factors		
Measure:HbA1c		
Zaman	Dependent Variable	
1	HbA1c1	
2	HbA1c2	
3	HbA1c3	
4	HbA1c4	

Between-Subjects Factors			
		Value Label	N
Cinsiyet	1	Male	236
	2	Female	194

Descriptive Statistics				
	Cinsiyet	Mean	Std. Deviation	N
HbA1c1	Male	6,99	,29	236
	Female	6,99	,27	194
	Total	6,99	,28	430
HbA1c2	Male	5,00	,10	236
	Female	4,99	,10	194
	Total	5,00	,10	430
HbA1c3	Male	5,30	,15	236
	Female	5,29	,16	194

	Total	5,30	,15	430
HbA1c4	Male	5,48	,31	236
	Female	6,48	,44	194
	Total	5,93	,62	430

Multivariate Tests ^b					
Effect		Value	F	Hypothesis df	Error df Sig.
Zaman	Pillai's Trace	,98	7071,2 ^a	3	426,000
	Wilks' Lambda	,02	7071,2 ^a	3	426,000
	Hotelling's Trace	49,79	7071,2 ^a	3	426,000
	Roy's Largest Root	49,79	7071,2 ^a	3	426,000
Zaman * sex	Pillai's Trace	,63	243,8 ^a	3	426,000
	Wilks' Lambda	,36	243,8 ^a	3	426,000
	Hotelling's Trace	1,72	243,8 ^a	3	426,000
	Roy's Largest Root	1,72	243,8 ^a	3	426,000
a. Exact statistic b. Design: Intercept + sex					
Within Subjects Design: Zaman					

Mauchly's Test of Sphericity^a

Measure: HbA1c

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon ^a		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Zaman	,43	358,9	5	,000	,700	,705	,333

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix. a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table. b. Design: Intercept + sex Within Subjects Design: Zaman

Tests of Within-Subjects Effects

Measure: HbA1c

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Zaman	Sphericity Assumed	998,3	3	332,8	5362,5	,000
	Greenhouse-Geisser	998,3	2,099	475,6	5362,5	,000
	Huynh-Feldt	998,3	2,115	472,1	5362,5	,000
	Lower-bound	998,3	1,000	998,3	5362,5	,000
Zaman * sex	Sphericity Assumed	80,4	3	26,8	431,8	,000
	Greenhouse-Geisser	80,4	2,099	38,3	431,8	,000

	Huynh-Feldt	80,4	2,115	38,0	431,8	,000
	Lower-bound	80,4	1,000	80,4	431,8	,000
Error(Zaman)	Sphericity Assumed	79,7	1284	,1		
	Greenhouse-Geisser	79,7	898,448	,1		
	Huynh-Feldt	79,7	905,085	,1		
	Lower-bound	79,7	428,000	,2		

Tests of Within-Subjects Contrasts

Measure:HbA1c

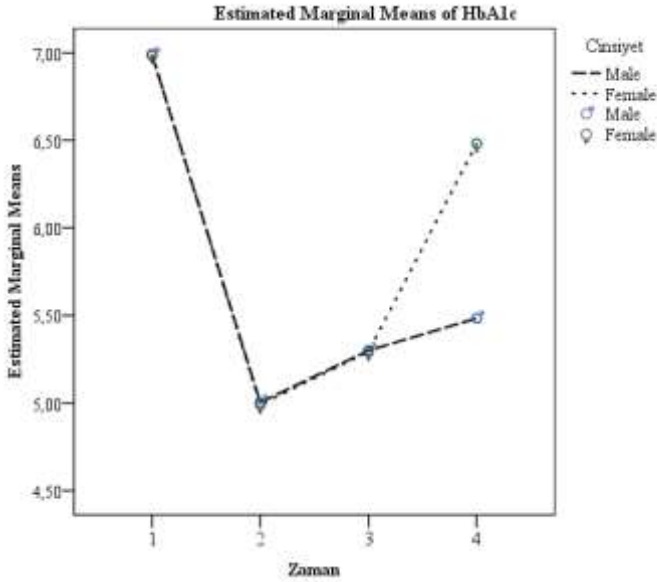
Source	Zaman	Type III Sum of Squares	df	Mean Square	F	Sig.
Zaman	Level 1 vs. Level 2	1689,5	1	1689,5	20041,4	,000
	Level 2 vs. Level 3	38,1	1	38,1	1112,3	,000
	Level 3 vs. Level 4	200,7	1	200,7	1196,6	,000
Zaman * sex	Level 1 vs. Level 2	,0	1	,0	,4	,537
	Level 2 vs. Level 3	,0	1	,0	,1	,766
	Level 3 vs. Level 4	107,6	1	107,6	641,6	,000
Error(Zaman)	Level 1 vs. Level 2	36,1	428	,1		
	Level 2 vs. Level 3	14,7	428	,0		
	Level 3 vs. Level 4	71,8	428	,2		

Tests of Between-Subjects Effects

Measure:HbA1c Transformed Variable:Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	14408,784	1	14408,9	901015,6	,000
sex	6,421	1	6,4	401,5	,000
Error	6,844	428	,01		

Profile Plots



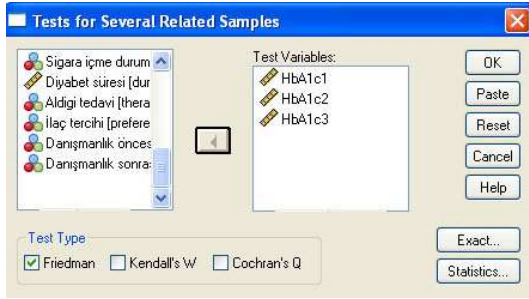
Varsayımlar karşılanamıyorsa

Verilerimizin bağımlı gruplarda tekrarlayan ölçüm için varsayımları karşılamaması halinde veri dönüşümü uygulayabilir veya bu testin nonparametrik versiyonu olan Friedman Varyans analizinden yararlanmamız gerekir.

Şimdi HbA1c değişkenini kullanarak Friedman Varyans analizini yapalım:

*Analyze > Nonparametric Tests > K Related Samples > [Test Variables alanına **HbA1c1, HbA1c2, HbA1c3** taşıyalım] >OK*

Aşağıdaki çıktıları elde ederiz: Test Statistics (a) tablosunda $p < 0,001$ olduğundan üç farklı zamanda ölçümlerden en az birisinin diğerlerinden farklı olduğunu görürüz. Hangi grup veya grupların diğerlerinden farklı olduğunu anlamak için Wilcoxon testi ile İkişerli karşılaştırmalar yapılabilir (1. Ölçüm ile 2. ölçüm 1. Ölçüm ile 3. Ölçüm, 2. Ölçüm ile 3. Ölçüm). Ancak istatistiksel önemlilik düzeyi Bonferroni düzeltmesine göre alınmalıdır. Bonferroni düzeltmesi, önemlilik düzeyi $p < 0,05$ olarak kabul edildiği durumda, “0,05/ikili karşılaştırma sayısı” olarak belirlenir. Örneğimizde, 3 grup arasında farkın hangi grup veya gruplardan kaynaklandığını belirlerken anlamlılık düzeyi $0,05/3=0,016$ olarak kabul edilir.



NPar Tests

Friedman Test

Ranks

	Mean Rank
HbA1c1	3,00
HbA1c2	1,06
HbA1c3	1,94

Test Statistics(a)

N	430
Chi-Square	812,907
Df	2
Asymp. Sig.	,000

a. Friedman Test

Bölüm 27: Nümerik Veri İkidenden Fazla Bağımsız Grup (Tek Yönlü ANOVA, Kruskal Wallis Testi)

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun ikiden fazla bağımsız gruptan elde edilen nümerik verilerin ortalamalarının karşılaştırılmasında kullanılan tek yönlü varyans analizi (ANOVA) ve bunun nonparametrik alternatifi olan Kruskal Wallis testi hakkında bilgi sahibi olması ve bu testleri SPSS ile yapabilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Nümerik veri, ikiden fazla bağımsız grubu tanımlayabilmeli
- ANOVA varsayımlarını sayabilmeli
- SPSS'te ANOVA testi yapabilmeli
- ANOVA için Post Hoc analiz yapabilmeli
- ANOVA testi SPSS çıktısını yorumlayabilmeli
- SPSS'te Kruskal Wallis testi yapabilmeli
- Kruskal Wallis testi SPSS çıktısını yorumlayabilmeli

Diyabet hastalarımızın eğitim durumlarına göre boylarının farklılık gösterip göstermediğini araştırmak istiyoruz (H_0 : Eğitim durumları farklı olan diyabet hastalarının boyları arasında fark yoktur). Ölçmek istediğimiz değişken (boy) nümerik sürekli bir değişkendir. Eğitimsiz, ilkokul, ortaokul ve lise olmak üzere dört grubumuz var.

Yukarıdaki hipotezimizi test etmek için uygulayabileceğimiz test tek yönlü varyans analizidir (Analysis of Variance; ANOVA). Bu testi uygulayamamamız halinde nonparametrik alternatifi olan Kruskal Wallis testini yapabiliriz.

ANOVA

Varsayımları

Tek yönlü varyans analizinde gruplar tek bir faktörle belirlenir (örn. babanın mesleğinin öğrencinin ÖSS sınav notuna etkisi). Birden fazla faktörün birbiriyle ilişkisinin incelendiği durumlarda (örn. babanın mesleğinin VE okul türünün öğrencinin ÖSS sınav notuna etkisi) çok yönlü varyans analizi uygulanır.

- İncelenen toplumda değişken, grupların hepsinde normal dağılmalıdır.
- Değişkenin gruplar arasındaki varyansı eşit olmalıdır.
- Örneklem sayısı bu varsayımları kontrol edecek kadar büyük olmalıdır.

Mantığı

Bu test grup içi ve gruplar arası varyansların oranına dayanmaktadır. Sıfır hipotezinin doğru olması halinde değişkenin grupların genelindeki varyasyonu ile gruplar arasındaki varyasyonu aynı olmalıdır. Testin adının varyans analizi olması da buradan gelmektedir.

Şimdi, diyabet.sav verisetimizde “height” değişkenimizin eğitim grupları açısından durumunu incelemek için ANOVA testini uygulayalım:

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

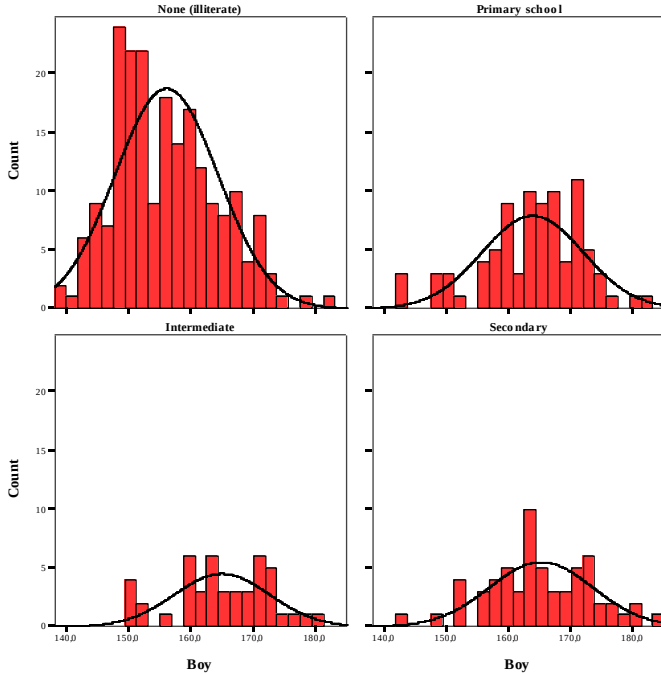
H_0 : Eğitim durumları farklı olan diyabet hastalarının boyları arasında fark yoktur.

H_1 : Eğitim durumları farklı olan diyabet hastalarının boyları arasında fark vardır.

2. Verilerin toplanması:

Veriler diyabet.sav verisetinde “height” ve “educat” değişkenlerinin altında kaydedildi.

Örneklem sayısı yeterlidir. Değişkenimizin gruplar açısından histogram grafiğine bakıldığında normal dağıldığı görülmektedir:



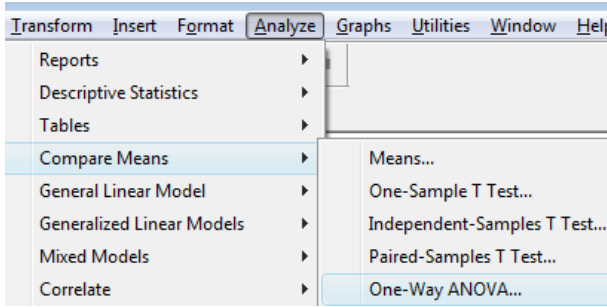
Varyansların homojenliği varsayımı ANOVA testi sırasında incelenecektir.

3. İlgili sıfır hipotezi için *test istatistiğinin* hesaplanması:

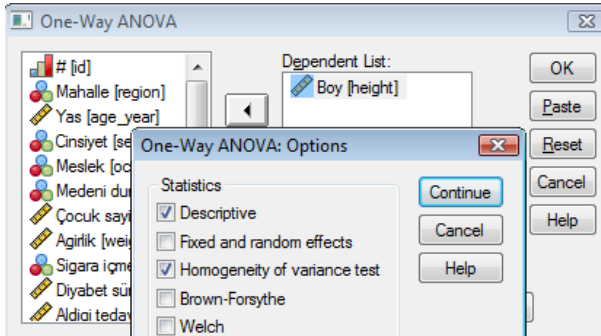
ANOVA testi grup için ve gruplar arası varyansların birbirine oranlanmasından (F oranı) oluşur. Bu oran F dağılımına uymaktadır.

SPSS ile yapacak olursak:

Analyze>Compare Means>One-Way ANOVA>[“height” değişkenini “Dependent List” alanına, “educat” değişkenini de “Factor” alanına geçirelim



>Options >[“Descriptive” ve “Homogeneity of variance test” kutucuklarını işaretleyelim]>Continue>ok.



Aşağıdaki çıktıları elde ederiz:

Descriptives

Boy

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
None (illiterate)	208	155,923	8,3130	,5764	154,787	157,059	138,0	182,0
Primary school	86	163,686	8,1418	,8780	161,940	165,432	142,0	183,0
Intermediate	46	164,826	7,6545	1,1286	162,553	167,099	150,0	180,0
Secondary	61	165,049	8,2954	1,0621	162,925	167,174	142,0	185,0
Total	401	159,998	9,2137	,4601	159,093	160,902	138,0	185,0

Test of Homogeneity of Variances

Boy

Levene Statistic	df1	df2	Sig.
,434	3	397	,729

ANOVA

Boy

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	7252,244	3	2417,415	35,938	,000
Within Groups	26704,754	397	67,266		
Total	33956,998	400			

Birinci tablo, değişkenin gruplara göre tanımlayıcı istatistiklerini vermektedir.

İkinci tabloda varyansların homojenliği test edilmektedir. p (Sig.) $>0,05$ olduğundan eğitim gruplarına göre varyanslar arasında fark olmadığı sonucunu çıkarabiliriz.

Üçüncü tablo F istatistiğini veriyor. Gruplar arası varyansların kareleri ortalamasının (2417,415) gruplar içi varyansların kareleri ortalamasına (67,266) bölünmesi F değerini vermektedir.

4. Test istatistiğinden elde edilen değer bilinen bir olasılık dağılımı ile karşılaştırılması:

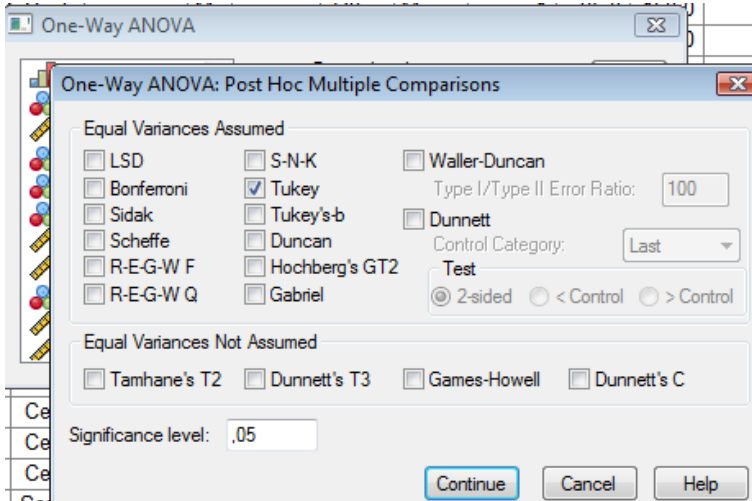
F oranı için sayfa 461'ye baktığımızda p değerinin 0,05'ten küçük olduğunu görürüz. Zaten SPSS'te p 'yi sıfıra yakın bir değer olarak vermiştir.

5. P değerinin ve sonuçların yorumlanması

P değeri 0,05'ten küçük olduğu için H_0 hipotezini reddederiz. Dolayısıyla diyabetik bireylerin eğitim durumlarına göre boylarının farklılık gösterdiği sonucuna varabiliriz.

ANOVA bize incelenen gruplar arasında bağımlı değişken açısından fark olduğunu söylemektedir. Grupların hepsinin birbirinden farklı olması gerekmez. Örneğin, sadece eğitimsiz bireylerle lise mezunları arasında bir fark olabilir. Farkın hangi gruplar arasında olduğunu anlamak için ANOVA'da Post Hoc analiz yapabiliriz. Bunun için yukarıdaki işlemi şu şekilde değiştirmeliyiz:

Analyze>Compare Means>One-Way ANOVA>[“height” değişkenini “Dependent List” alanına, “educat” değişkenini de “Factor” alanına geçirelim>Options >[“Descriptive” ve “Homogeneity of variance test” kutucuklarını işaretleyelim]>Post Hoc>[“Tukey” kutucuğunu işaretleyelim]>Continue>OK.



Önceki çıktıya ek olarak aşağıdaki tablo oluşacaktır:

Multiple Comparisons

Dependent Variable: Boy

Tukey HSD

(I) Eğitim durumu	(J) Eğitim durumu	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
None (illiterate)	Primary school	-7,7630*	1,0515	,000	-10,476	-5,050
	Intermediate	-8,9030*	1,3363	,000	-12,351	-5,455
	Secondary	-9,1261*	1,1942	,000	-12,207	-6,045
Primary school	None (illiterate)	7,7630*	1,0515	,000	5,050	10,476
	Intermediate	-1,1400	1,4982	,872	-5,005	2,725
	Secondary	-1,3631	1,3729	,754	-4,905	2,179
Intermediate	None (illiterate)	8,9030*	1,3363	,000	5,455	12,351
	Primary school	1,1400	1,4982	,872	-2,725	5,005
	Secondary	-,2231	1,6016	,999	-4,355	3,909
Secondary	None (illiterate)	9,1261*	1,1942	,000	6,045	12,207
	Primary school	1,3631	1,3729	,754	-2,179	4,905
	Intermediate	,2231	1,6016	,999	-3,909	4,355

*. The mean difference is significant at the .05 level.

Bu tablodan görüldüğü gibi esas fark, eğitimsiz bireylerle diğer gruplar arasındadır. Diğer grupların kendi aralarında fark yoktur.

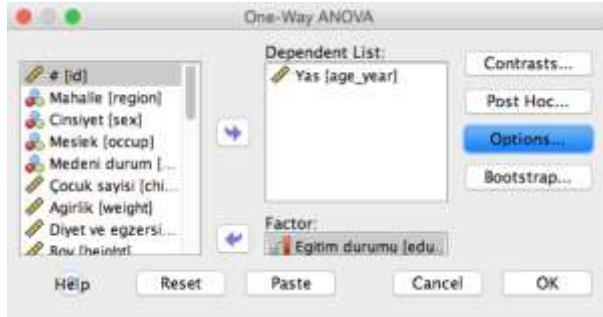
ANOVA Post Hoc analiz penceresine bakıldığında çok sayıda seçenek olduğu görülmektedir. Bunlardan üst kısımdakiler varyansların homojen olması durumunda alttakiler ise varyansların homojenliği varsayımı karşılanamadığında kullanılır. Hangi post hoc analizin seçileceği kullanıcının ihtiyacına göre belirlenir. Çoklu hipotez testlerinden kaynaklanan hatalar açısından düzeltme yapılan Tukey ve Bonferroni gibi testler tercih edilebilir.

Tek yönlü ANOVA ve lineerite testi

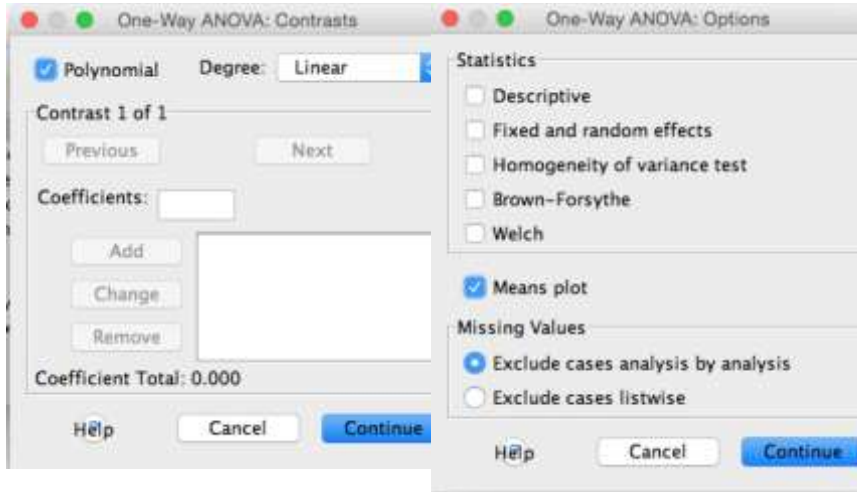
Hastalarımızın eğitim durumlarına göre yaşlarının farklılık gösterip göstermediğini araştırmak istiyoruz. Ayrıca arada bir ilişki varsa yaşın eğitim durumu arttıkça lineer bir şekilde değişip değişmediğini belirlemek istiyoruz.

(H₀: Eğitim durumları farklı olan hastalarının yaşları arasında lineer bir ilişki yoktur). Ölçmek istediğimiz değişken (yaş) nümerik bir değişkendir. Eğitim durumu kategorik sıralı bir değişkendir.

Analyze>Compare Means>One-Way ANOVA>[“age_year” değişkenini “Dependent List” alanına, “educat” değişkenini de “Factor” alanına geçirelim>



Contrasts >[“Polynomial” kutucuğunu işaretleyelim]>Options[“Means Plot” kutucuğunu işaretleyelim]>Continue>OK.

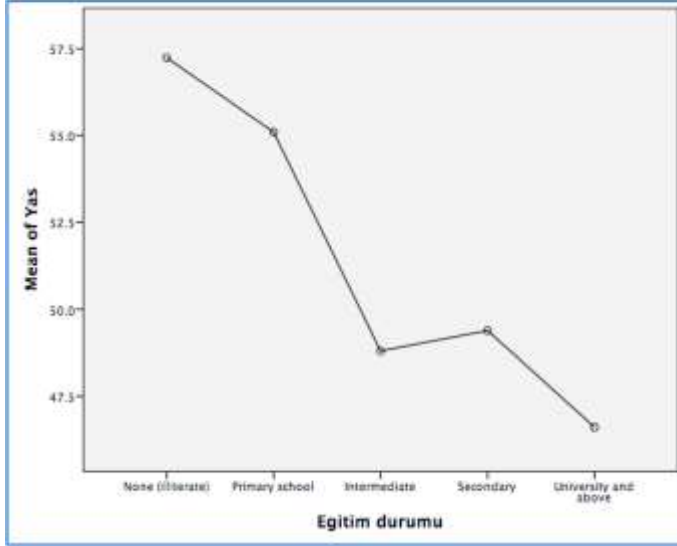


Aşağıdaki çıktıları elde ederiz:

ANOVA						
Yas						
			Sum of Squares	df	Mean Square	F
Between Groups	(Combined)		5654.054	4	1413.514	9.716
	Linear Term	Unweighted	2661.208	1	2661.208	18.293
		Weighted	5197.587	1	5197.587	35.727
		Deviation	456.468	3	152.156	1.046
Within Groups			60955.547	419	145.479	
Total			66609.601	423		

Eğitim grupları arasında yaş açısından anlamlı bir fark vardır ($F=9.716$; $p<0,001$). Aradaki fark lineer bir şekilde değişmektedir (Weighted Linear Term $F=35.727$; $p<0,001$).

Not: Gruplar arasında birey sayıları farklılık gösteriyorsa ve gruplar arasında nümerik değişkenin dağılım aralıkları aynı değilse “Weighted Linear Term”, aksi halde “Unweighted Linear Term” raporlanır.



Ortalamaların çizgi grafiğine baktığımızda da eğitim durumu arttıkça yaş ortalamasının düşme meylinde olduğunu ve bu düşüşün lineer bir trend izlediğini söyleyebiliriz.

Varsayımlar karşılanamıyorsa

ANOVA testi normallikten sapmalara kısmen dayanıklı olmakla beraber, varyansların homojen olmaması durumuna dayanıklı değildir. Verilerimizin varsayımları karşılamaması halinde veri dönüşümü uygulayabiliriz veya bu testin nonparametrik versiyonundan yararlanmamız gerekir.

Kruskal Wallis Testi

Bu test ANOVA'nın nonparametrik alternatifidir. Gruplar arasında sıraların toplamını dikkate alır. ANOVA varsayımlarını karşılamayan bir değişken için yapacak olursak.

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

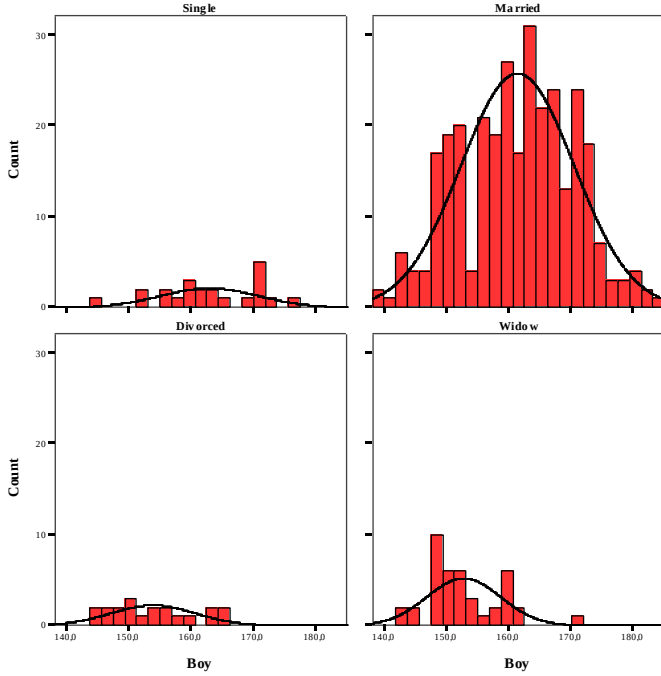
H_0 : Medeni durumları farklı olan diyabet hastalarının boyları arasında fark yoktur.

H_1 : Medeni durumları farklı olan diyabet hastalarının boyları arasında fark vardır.

2. Verilerin toplanması:

Veriler diyabet.sav verisetinde “height” ve “marital” değişkenlerinin altında kaydedildi.

Örneklem sayısı bazı gruplarda 30’un altındadır. Değişkenimizin gruplar açısından histogram grafiğine bakıldığında da bazı gruplarda normal dağılmadığı görülmektedir:



ANOVA yapıldığında da varyansların homojen dağılmadığı görülmektedir:

Test of Homogeneity of Variances

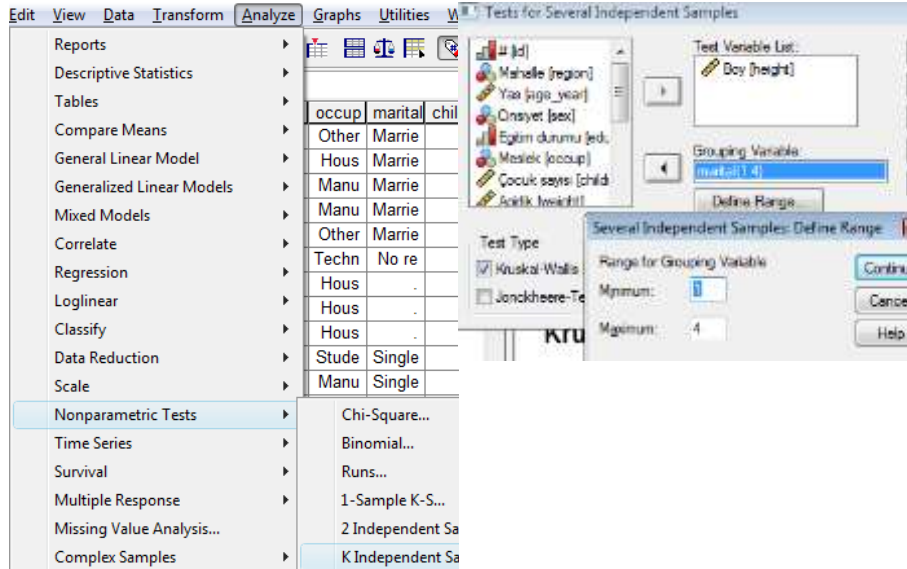
Boy

Levene Statistic	df1	df2	Sig.
4,852	3	392	,003

3. İlgili sıfır hipotezi için test istatistiğinin hesaplanması:

Analyze>Nonparametric Tests >K Independent Samples>[“height” değişkenini “Test Variable List” alanına, “marital” değişkenini de “Grouping Variable” alanına geçirelim

>Define Range >[Bu değişken için 1’den 4’e kadar 4 grubumuz olduğundan “Minimum” alanına 1, “Maximum” alanına 4 girelim]>Continue>OK.



Aşağıdaki çıktıları elde ederiz:

Ranks

	Medeni durum	N	Mean Rank
Boy	Single	22	236,02
	Married	313	213,72
	Divorced	20	117,35
	Widow	41	101,73
	Total	396	

Test Statistics(a,b)

	Boy
Chi-Square	47,341
Df	3
Asymp. Sig.	,000

a Kruskal Wallis Test

b Grouping Variable: Medeni durum

Birinci tablo, değişkenin gruplara göre sıralanmış verilerinin sıra ortalamalarını vermektedir.

İkinci tabloda Kruskal Wallis testi sonuçları görülmektedir.

4. Test istatistiğinden elde edilen değerin bilinen bir olasılık dağılımı ile karşılaştırılması:

Kruskal-Wallis testi *Ki Kare* dağılımı göstermektedir. 47,3 değerini Tablo A3'ten bakacak olursak:

<i>sd</i>	<i>İki yönlü p</i>			
	0.10	0.05	0.01	0.001
1	2.706	3.841	6.635	10.827
2	4.605	5.991	9.210	13.815
3	6.251	7.815	11.345	16.266
4	7.779	9.488	13.277	18.466

→47,34

SD=3 için (serbestlik derecesi burada grup sayısı – 1 olarak hesaplanır) iki yönlü *p* değerinin 0,001'den küçük olduğunu görürüz. Zaten SPSS te *p*'yi sıfıra yakın bir değer olarak vermiştir.

5. *P* değerinin ve sonuçların yorumlanması

P değeri 0,05'ten küçük olduğu için H_0 hipotezini reddederiz.

Dolayısıyla medeni durumları farklı olan diyabet hastalarının boyları arasında fark olduğu sonucuna varabiliriz.

Bölüm 28: Nümerik Veri İki Kategorik Bağımsız Değişken (İki Yönlü ANOVA)

Hamit ACEMOĞLU

Amaç: Bu konu sonunda okuyucunun iki kategorik bağımsız değişkenden elde edilen nümerik verilerin ortalamalarının karşılaştırılmasında kullanılan iki yönlü varyans analizi (ANOVA) testi hakkında bilgi sahibi olması ve bu testleri SPSS ile yapabilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Nümerik veri, kategorik bağımsız değişkeni tanımlayabilmeli
- İki yönlü ANOVA varsayımlarını sayabilmeli
- SPSS'te İki yönlü ANOVA testi yapabilmeli
- İki yönlü ANOVA için Post Hoc analiz yapabilmeli
- İki yönlü ANOVA testi SPSS çıktısını yorumlayabilmeli

*Diyabet hastalarımızın cinsiyet ve sağlık hizmetlerinden memnuniyet durumlarına göre açlık kan şekerlerinin farklılık gösterip göstermediğini araştırmak istiyoruz (H_0 : Cinsiyetleri ve memnuniyet durumları farklı olan diyabet hastalarının açlık kan şekerleri arasında fark yoktur). Ölçmek istediğimiz değişken (açlık kan şekerleri) nümerik sürekli bir değişkendir. Birinci kategorik bağımsız değişkenimiz **cinsiyet**, erkek ve kadınlardan oluşan iki bağımsız grubu bulunurken, ikinci kategorik bağımsız değişkenimiz olan **memnuniyet durumu**; Hiç memnun değil, az memnun, memnun ve çok memnun olmak üzere dört bağımsız kategorik oluşmaktadır.*

Yukarıdaki hipotezimizi test etmek için uygulayabileceğimiz test iki yönlü varyans analizidir (Analysis of Variance; ANOVA). İki yönlü ANOVA analizi ile iki ya da daha fazla bağımsız değişkenin, bir bağımlı değişken üzerindeki etkisi incelenmektedir. Bağımlı değişkenler nümerik verilerden oluşurken (boy, ağırlık gibi), bağımsız değişkenler cinsiyet ve eğitim durumu gibi kategorik değişkenlerden oluşmaktadır. Varyans analizinde bağımsız değişkenler faktör olarak isimlendirilmektedir.

İki yönlü ANOVA

Varsayımları

Tek yönlü varyans analizinde gruplar tek bir faktörle belirlenir (örn. babanın mesleğinin öğrencinin ÖSS sınav notuna etkisi). Birden fazla faktörün birbiriyle ilişkisinin incelendiği durumlarda (örn. babanın mesleğinin ve okul türünün öğrencinin ÖSS sınav notuna etkisi) iki yönlü varyans analizi uygulanır.

- İncelenen bağımlı değişken (sonuç ölçütü, çıktı), grupların ait olduğu toplumda normal dağılmalıdır.
- Değişkenin gruplar arasındaki varyansı eşit olmalıdır.
- Ölçümler birbirlerinden bağımsız gruplardan elde edilmiş olmalıdır.

Mantığı

Bu test grup içi ve gruplar arası varyansların oranına dayanmaktadır. Sıfır hipotezinin doğru olması halinde değişkenin grupların genelindeki varyasyonu ile gruplar arasındaki varyasyonu aynı olmalıdır. Testin adının varyans analizi olması da buradan gelmektedir.

Şimdi, diyabet.sav verisetimizde “akş” değişkenimizi cinsiyet ve memnuniyet durumları açısından durumunu incelemek için iki yönlü ANOVA testini uygulayalım:

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

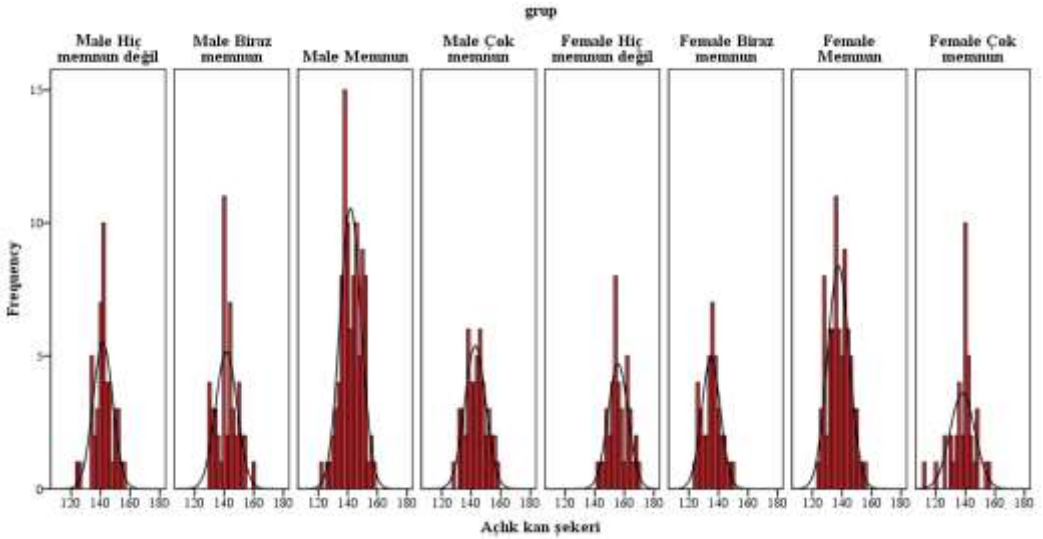
H_0 : Cinsiyetleri ve memnuniyet durumları farklı olan diyabet hastalarının açlık kan şekerleri arasında fark yoktur

H_1 : Cinsiyetleri ve memnuniyet durumları farklı olan diyabet hastalarının açlık kan şekerleri arasında grupların en az birinde fark vardır

Verilerin toplanması:

Veriler diyabet.sav verisetinde “akş”, “sex” ve “memnun1” değişkenlerinin altında kaydedildi.

Örneklem sayısı yeterlidir. Değişkenimizin gruplar açısından histogram grafiğine bakıldığında normal dağıldığı görülmektedir:



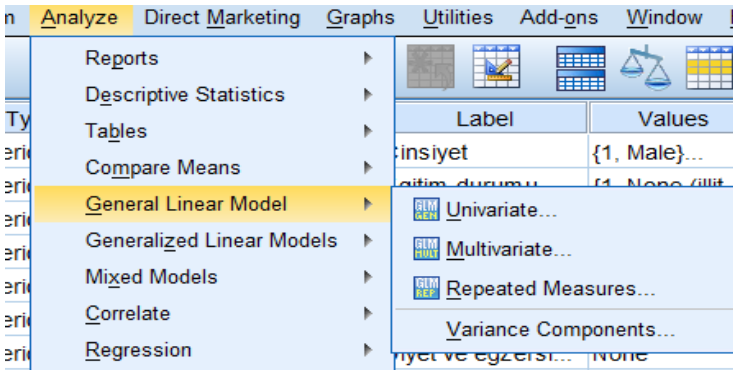
Varyansların homojenliği varsayımı ANOVA testi sırasında incelenecektir.

2. İlgili sıfır hipotezi için *test istatistiğinin* hesaplanması:

ANOVA testi grup için ve gruplar arası varyansların birbirine oranlanmasından (F oranı) oluşur. Bu oran F dağılımına uymaktadır.

SPSS ile yapacak olursak:

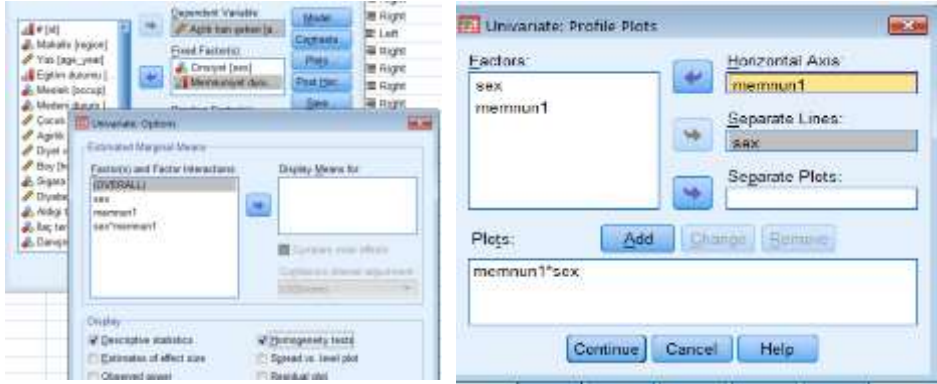
Analyze>General Linear Model>Univariate>[“akş” değişkenini “Dependent List” alanına, “sex” ve “memnun1” değişkenlerini de “Fixed Factor(s)” alanına geçirelim



>Options >[“Descriptive statistics” ve “Homogeneity test” kutucuklarını işaretleyelim]>Continue>

Grafik çizimi için;

>Plots> [Horizontal Axis kutucuğuna “memnun1”’i Separate Lines kutucuğuna alalım]>Add>Continue



Post Hoc testler için Post Hoc butonunu tıklayalım.

>[Factors alanından “memnun1” değişkenini Post Hoc Test for alanına geçirelim]>Post Hoc>[“Tukey” kutucuğunu işaretleyelim]>Continue> ok.



Aşağıdaki çıktıları elde ederiz:

Univariate Analysis of Variance

Descriptive Statistics

Dependent Variable: Açlık kan şekeri

Cinsiyet	Memnuniyet durumu	Mean	Std. Deviation	N
Male	Hiç memnun değil	141,4	6,7	46
	Az memnun	141,4	7,3	47
	Memnun	141,9	7,2	95
	Çok memnun	142,6	7,1	48
	Total	141,8	7,1	236
Female	Hiç memnun değil	155,7	6,8	40
	Az memnun	135,0	6,3	39
	Memnun	137,6	7,3	77
	Çok memnun	137,9	8,4	38
	Total	140,9	10,5	194
Total	Hiç memnun değil	148,1	9,8	86
	Az memnun	138,5	7,5	86
	Memnun	140,0	7,5	172
	Çok memnun	140,6	8,0	86
	Total	141,4	8,8	430

Levene's Test of Equality of Error Variances^a

Dependent Variable: Açlık kan şekeri

F	df1	df2	Sig.
,545	7	422	,80

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + sex + memnun1 + sex * memnun1

Tests of Between-Subjects Effects

Dependent Variable: Açlık kan şekeri

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	11477,3 ^a	7	1639,6	31,9	,000
Intercept	7819753,8	1	7819753,8	152279,6	,000
sex	6,9	1	6,9	0,1	,715
memnun1	5872,5	3	1957,5	38,1	,000
sex * memnun1	6365,7	3	2121,9	41,3	,000
Error	21670,2	422	51,4		
Total	8631116,0	430			
Corrected Total	33147,6	429			

a. R Squared = ,346 (Adjusted R Squared = ,335)

Homogeneous Subsets

Açlık kan şekeri

Tukey HSD^{a,b,c}

Memnuniyet durumu	N	Subset	
		1	2
Az memnun	86	138,5	
Memnun	172	140,0	
Çok memnun	86	140,6	
Hiç memnun değil	86		148,1
Sig.		,2	1,0

Means for groups in homogeneous subsets are displayed.

Based on observed means.

The error term is Mean Square(Error) = 51,351.

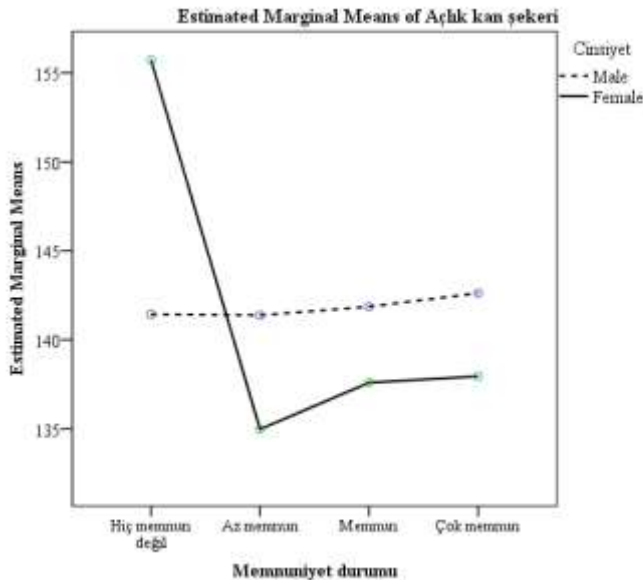
a. Uses Harmonic Mean Sample Size = 98,286.

b. The group sizes are unequal. The harmonic mean of the group sizes is used.

Type I error levels are not guaranteed.

c. Alpha = ,05.

Profile Plots



Birinci tablo, değişkenin gruplara göre tanımlayıcı istatistiklerini vermektedir. Toplamda Hiç memnun olmayanların ortalamasının diğer gruplardan daha yüksek olduğu bu yüksekliğin kadınlarda daha da belirginleştiği görülmektedir. Erkelerde memnuniyet durumuna göre açlık kan şekeri ortalaması birbirine yakındır.

İkinci tabloda varyansların homojenliği test edilmektedir. p (Sig.) $>0,05$ olduğundan varyanslar arasında fark olmadığı sonucunu çıkarabiliriz.

Tests of Between-Subjects Effects tablosu F istatistiğini veriyor. Cinsiyete göre açlık kan şekeri ortalamaları arasında bir farklılık olmadığı ($p=0.715$), Memnuniyet durumu açısından gruplar arasında farklılık olduğu, aynı zamanda cinsiyet*memnuniyet etkileşiminin de olduğu ($p=0.000$) görülmektedir. Grafikte incelendiğinde Hiç memnun olmayan kadınların açlık kan şekeri ortalamalarının en yüksek olduğu ve cinsiyet*memnuniyet etkileşiminin bu nedenle olduğu görülmektedir.

Homogeneous Subsets tablosunda Hiç memnun değil grubunun ayrı bir alt grupta olduğu diğer üç grubun aynı alt grup altında toplandıkları görülmektedir.

3. Test istatistiğinden elde edilen değerlerin bilinen bir olasılık dağılımı ile karşılaştırılması:

F oranı için sayfa 461'ye baktığımızda p değerinin cinsiyet için 0,05'ten büyük, memnuniyet durumu ve cinsiyet*memnuniyet etkileşimi için 0,05'ten küçük olduğunu görürüz. Zaten SPSS'te p 'yi sıfıra yakın bir değer olarak vermiştir.

4. P değerinin ve sonuçların yorumlanması

Cinsiyet için P değeri 0,05'ten büyük olduğu için H_0 hipotezini kabul ederiz. Dolayısıyla diyabetik bireylerin cinsiyetlerine göre açlık kan şekeri ortalamalarının farklılık göstermediği sonucuna varabiliriz.

Memnuniyet durumu ve cinsiyet*memnuniyet etkileşimi için P değeri 0,05'ten küçük olduğu için H_0 hipotezini reddederiz. Dolayısıyla diyabetik bireylerin memnuniyet durumları ve cinsiyet*memnuniyet etkileşimi durumlarına göre açlık kan şekeri ortalamalarının farklılık gösterdiği sonucuna varabiliriz.

ANOVA bize incelenen gruplar arasında bağımlı değişken açısından fark olduğunu söylemektedir. Grupların hepsinin birbirinden farklı olması gerekmez. Farkın hangi gruplar arasında olduğunu anlamak için ANOVA'da Post Hoc analiz sonuçlarını incelemek gerekir.

Multiple Comparisons
Açlık kan şekeri
Tukey HSD

(I) Memnuniyet durumu	(J) Memnuniyet durumu	Mean Difference (I- J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Hiç memnun değil	Az memnun	9,6 [*]	1,1	,000	6,8	12,4
	Memnun	8,1 [*]	,9	,000	5,7	10,6
	Çok memnun	7,5 [*]	1,1	,000	4,7	10,3
Az memnun	Hiç memnun değil	-9,6 [*]	1,1	,000	-12,4	-6,8
	Memnun	-1,5	,9	,403	-3,9	1,0
	Çok memnun	-2,1	1,1	,228	-4,9	0,7
Memnun	Hiç memnun değil	-8,1 [*]	,9	,000	-10,6	-5,7
	Az memnun	1,5	,9	,403	-1,0	3,9
	Çok memnun	-,6	,9	,919	-3,0	1,8
Çok memnun dimension3	Hiç memnun değil	-7,5 [*]	1,1	,000	-10,3	-4,7
	Az memnun	2,1	1,1	,228	-0,7	4,9
	Memnun	-,6	,9	,919	-1,8	3,0

Based on observed means.

The error term is Mean Square(Error) = 51,351.

*. The mean difference is significant at the ,05 level.

Bu tablodan görüldüğü gibi esas fark Hiç memnun değil grubu ile diğer gruplar arasındadır. Diğer grupların kendi aralarında fark yoktur.

Varsayımlar karşılanamıyorsa

ANOVA testi normallikten sapmalara kısmen dayanıklı olmakla beraber, varyansların homojen olmaması durumuna dayanıklı değildir. Verilerimizin varsayımları karşılamaması halinde veri dönüşümü uygulayabiliriz.

Bölüm 29: Kategorik Veri Tek Grup (Tek Orantı Testi)

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun tek bir gruptan elde edilen kategorik verilerin belli bir toplum orantısı ile karşılaştırılmasında kullanılan tek orantıyı toplumla karşılaştırma ve Sign testinin orantılar için kullanımı konusunda bilgi sahibi olması ve bu analizleri yapabilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Kategorik veri, tek grubu tanımlayabilmeli
- Tek orantı testinin varsayımını ve mantığını açıklayabilmeli
- Orantılar için işaret testinin kullanım yerini açıklayabilmeli
- Orantılar için işaret testinin uygulanma yöntemini açıklayabilmeli
- SPSS’te binomiyal testi yapabilmeli
- Binomiyal test SPSS çıktısını yorumlayabilmeli

Bölgemizdeki diyabet hastalarında sigara içme durumunu araştırdık. Örneklemimizdeki sigara içme durumunun toplumdakinden farklı olup olmadığına bakmak istiyoruz (H_0 : Diyabet hastalarının sigara içme durumu toplumdakinden farklı değildir).

Ölçmek istediğimiz değişken (sigara içme durumu) “evet” ve “hayır” şeklinde kodlanan ikili (dikotom) bir değişkendir. Bu tür değişkenlerin binomiyal dağılım gösterdiklerinden daha önce bahsedilmişti. Verilerimizi tek bir gruptan elde ettik ve belli bir toplum ortalaması ile karşılaştırmak istiyoruz

Tek orantı testi (test of a single proportion)

Varsayımlar

Örneklemimiz karşılaştırma için kullanılacak toplumdan alınmış olmalıdır. Her bireyde iki durum söz konusu olabilir: araştırılan özellik “var” veya “yok”.

Not

n sayıdaki örneklemimizde r sayıda bireyde incelenen durum “var” dır. İncelenen durumun orantısı $p = r/n$ olacaktır. İncelenen durumun toplumdaki π_1 orantısı ile aynı olup olmadığını araştırıyoruz.

Mantığı

Teorik dağılımlar bahsinde belirtildiği gibi, bu tür değişkenler binomiyal dağılım gösterir. Ancak, np ve $n(1-p)$ ’nin her ikisinin de 5’ten büyük olması

halinde dağılım normale yakın olur. Bu durumda orantının ortalaması p , standart hatası ise

$$= \sqrt{\frac{p(1-p)}{n}}$$

olacaktır.

Diyabet.sav (<https://drive.google.com/open?id=0B0ieO-Fy3z-5a3cydjIPQ0YtYk0>) veri setimizde “smoking” değişkenini kullanarak bir örnek yapabiliriz. Daha önceki araştırmalardan toplumumuzun %88’inin (π_1) sigara içmediğinin bulunduğunu varsayalım:

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

H_0 : Örneklemimizdeki bireylerin sigara içmeme durumu %88’e eşittir.

H_1 : Örneklemimizdeki bireylerin sigara içmeme durumu %88’den farklıdır.

2. Verilerin toplanması:

Veriler diyabet.sav veri setinde “smoking” değişkeni altında kaydedildi.

Verilerimiz karşılaştırmının yapılacağı toplumdaki elde edilmiştir.

3. İlgili sıfır hipotezi için *test istatistiğinin* hesaplanması:

$$z = \frac{|p - \pi_1| - \frac{1}{2n}}{\sqrt{\frac{p(1-p)}{n}}}$$

Bu değişken için $p = \%91,6$ ($373/407$), örneklem sayısı (n) 407’dir.

Formülde yerine konulursa; $\{[(0,916-0,88) - (1/(2 \times 407))]/\sqrt{[(0,916 \times 0,084)/407]}\}$

$z = 2,53$ olarak hesaplanır.

Not: $np = 407 \times 0,91 = 370,37 > 5$ ve $n \times (1-p) = 407 \times (1-0,91) = 36,63 > 5$. Her iki değer de 5’ten büyük olduğundan dağılımın normal dağılıma yakın olduğunu varsayıyoruz.

4. Test istatistiğinden elde edilen değerlerin bilinen bir olasılık dağılımı ile karşılaştırılması:

z tablosundan karşılaştıracak olursak:

<i>z</i>	<i>İki yönlü p</i>
0.0	1.000
0.1	0.920
2.4	0.016
2.5	0.012
2.6	0.009

p değerinin 0,009 ile 0,012 arasında olduğu görülmektedir.

5. *P* değerinin ve sonuçların yorumlanması

P değeri 0,05'ten küçük olduğu için H_0 hipotezini reddederiz.

Dolayısıyla örneklemimizdeki bireylerin toplumdakilere oranla daha az sigara içtiği yorumunu yapabiliriz.

$$p \pm 1.96 \sqrt{\frac{p(1-p)}{n}}$$

Dilersek formülünden %95 güven aralığını da hesaplayabiliriz.

$$0,916 \pm 1,96\sqrt{((0,916 \times 0,084)/407)}$$

$$[0,889 - 0,970]$$

Hesapladığımız güven aralığı toplum ortalaması olan %88'i içermediğinden örneklemimizin toplumdan farklı olduğu sonucunu çıkarabiliriz.

Kategorik veri tek grup için SPSS 15'te analiz seçeneği bulunmamaktadır. Bu test MedCalc™¹ veya MINITAB™² ile yapılabilir.

Eğer bir tercih durumunu araştırıyorsak (Binomiyal test)

Eğer incelediğimiz ikili değişken bir tercih durumunu belirtiyorsa işaret testini veya binomiyal testi kullanabiliriz. Örneğin diyabetli hastalarımıza iki farklı firmanın ürettiği biguanid türü ilacı verdiğimiz ve kullanım kolaylığı, tadı, yan etkileri vs. açısından hangi ilacı tercih ettiklerini sorduğumuzu düşünelim. Tek bir gruptan (diyabetli hastalarımız) elde ettiğimiz kategorik bir değişkenimiz (ilaç tercihi; A ilacı – B ilacı) olacaktır. İlaçlar arasında bir fark olmaması durumunda her iki ilacı da tercih eden birey sayısı eşit olacaktır.

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

¹ <http://www.medcalc.be/>

² <http://www.minitab.com/en-TR/default.aspx>

H_0 : Diyabetik hastalarda A ve B ilacı tercihi arasında fark yoktur.

H_1 : Diyabetik hastalarda A ve B ilacı tercihleri arasında fark vardır.

2. Verilerin toplanması:

Veriler diyabet.sav verisetinde “preference” değişkeni altında kaydedildi.

3. İlgili sıfır hipotezi için *test istatistiğinin* hesaplanması:

Kullanacağımız formül nümerik verilerdeki işaret testine benzer olacaktır.

Herhangi bir tercihte bulunmayan bireyler çıkarılır. Geriye kalan örneklem sayısı n' olur. $p = r/n'$ olarak hesaplanır. Burada r A durumunu tercih eden birey sayısıdır.

- $n' \leq 10$ olması durumunda r dikkate alınır.
- $n' > 10$ olması durumunda z değeri hesaplanır.

$$z' = \frac{\left| p - \frac{1}{2} \right| - \frac{1}{2n'}}{\sqrt{\frac{p(1-p)}{n'}}$$

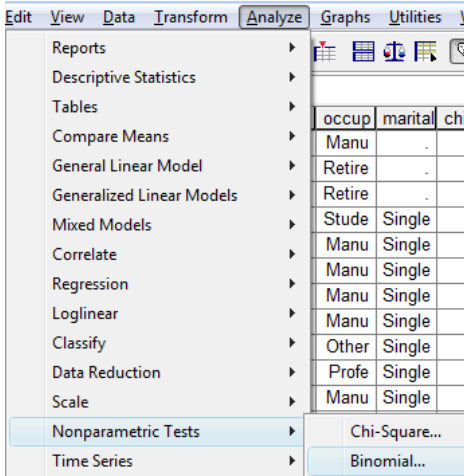
(burada π_1 yerine %50 yani $\frac{1}{2}$ alındığına dikkat ediniz).

Örneğimizde $p = 192/430 = 0,447$, $n' = 430$ 'dur.

Formülde yerine konduğunda $z = 2,16$ bulunur.

SPSS ile yapacak olursak:

Analyze>Nonparametric Tests>Binomial>



[“*preference*” değişkenini “*Test Variable List*” alanına geçirelim.
“*Test Proportion*” kutusuna 0,5 girelim]>ok.

Aşağıdaki çıktıyı elde ederiz:

Binomial Test						
		Category	N	Observed Prop.	Test Prop.	Asymp. Sig. (2-tailed)
İlaç tercihi	Group 1	Prefers drug B	238	,55	,50	,030 ^a
	Group 2	Prefers drug A	192	,45		
	Total		430	1,00		

a. Based on Z Approximation.

SPSS, sigara içme ve içmeme durumlarının eşit olduğu hipotezi kurulduğunda ($\pi_1 = \%50$) çift yönlü p’yi hesaplamaktadır. Aksi halde, toplumda sigara içenlerle içmeyenlerin birbirinden farklı olduğunu baştan kabul ettiğimizde (Test Proportion alanına %50’den farklı bir değer girdiğimizde) hipotez tek yönlü olmaktadır.

4. Test istatistiğinden elde edilen değer bilinen bir olasılık dağılımı ile karşılaştırılması:

- $n' \leq 10$ olması durumunda sayfa sayfa 463’ten r değerine bakılır.
- $n' > 10$ olması durumunda sayfa 458’den z değerine bakılır.

z değerini baktığımızda:

z	İki yönlü p
0.0	1.000
0.1	0.920
2.0	0.046
2.1	0.036
2.2	0.028

p değerinin 0,028 ile 0,036 arasında olduğunu görürüz. Zaten SPSS’te de $p = 0,030$ olarak verilmişti.

5. P değerinin ve sonuçların yorumlanması

P değeri 0,05’ten küçük olduğu için H_0 hipotezini reddederiz. Dolayısıyla diyabetik hastaların ilaç tercihleri arasında fark vardır. Hastalar istatistiksel olarak anlamlı derecede daha çok B ilacını tercih etmektedir.

Bölüm 30: Kategorik Veri İki Bağımlı Grup (McNemar Testi)

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun iki bağımlı gruptan elde edilen kategorik verilerin karşılaştırılmasında kullanılan McNemar testi hakkında bilgi sahibi olması ve bu testleri uygulayabilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Kategorik veri iki oranı açıklayabilmek
- 2x2 çapraz tablo oluşturabilmek
- Kategorik veri iki bağımlı gruba örnek verebilmek
- SPSS’te McNemar testini yapabilmek

Diyabetik hastalarımıza kan şekeri kontrolünün önemi ve gerektiğinde insülin kullanılması konusunda bir eğitim verdik. Örneklemimizdeki bireylerin ilaç tercihlerinin eğitimden sonra değişip değişmediğini incelemek istiyoruz. İncelediğimiz değişken [ilaç tercihi; 1-insülin, 2-oral antidiyabetik (AD)] kategorik bir değişkendir. Danışmanlık öncesi ve danışmanlık sonrasında aynı bireylerden veri topladık. Dolayısıyla iki bağımlı grup söz konusudur (H_0 : Diyabet hastalarının ilaç tercihlerinde danışmanlık öncesi ve sonrasında fark yoktur).

McNemar testi

Konu başında verdiğimiz ikinci senaryomuzdaki gibi bağımlı gruplarda kategorik bir değişkeni karşılaştırmak için McNemar testi kullanılabilir.

Testin mantığı

		Danışmanlık sonrası ilaç tercihi		Toplam
		İnsülin	Oral AD	İnsülin
Danışmanlık öncesi ilaç tercihi	İnsülin	159 (a)	6 (b)	165
	Oral AD	23 (c)	242 (d)	265
Toplam		182	248	430

Tablo: Gözlenen frekanslar.

İki ölçüm arasındaki fark, danışmanlık öncesinde ve sonrasındaki ölçümlerde farklılık gösteren **b** ve **c** hücreleridir. **a** ve **d** hücrelerinde danışmanlık öncesi ve

sonrasında bir değişiklik olmamıştır. McNemar testi bu farklılık gösteren b ve c hücrelerdeki frekansları dikkate alır.

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

H_0 : Diyabet hastalarının ilaç tercihlerinde danışmanlık öncesi ve sonrasında fark yoktur.

H_1 : Diyabet hastalarının ilaç tercihlerinde danışmanlık öncesi ve sonrasında fark vardır.

2. Verilerin toplanması:

Veriler diyabet.sav verisetinde “preference1” ve “preference2” değişkenleri altında kaydedildi.

3. İlgili sıfır hipotezi için *test istatistiğinin* hesaplanması:

$$\chi^2 = \frac{(|b - c| - 1)^2}{b + c}$$

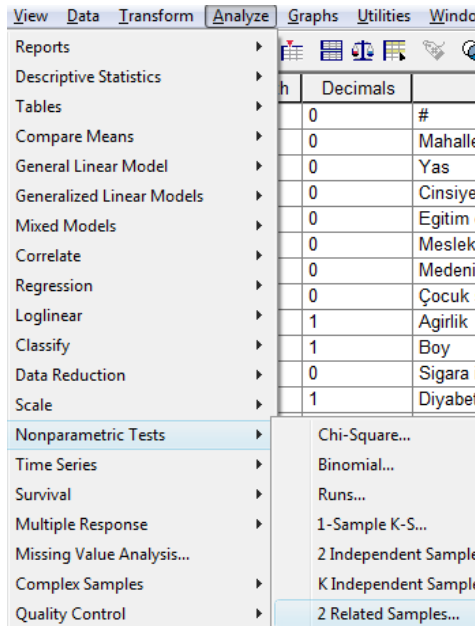
Bu test istatistiği, serbestlik derecesi 1 olacak şekilde Ki Kare dağılımı göstermektedir.

$$\chi^2 = (|6 - 23| - 1)^2 / (6 + 23)$$

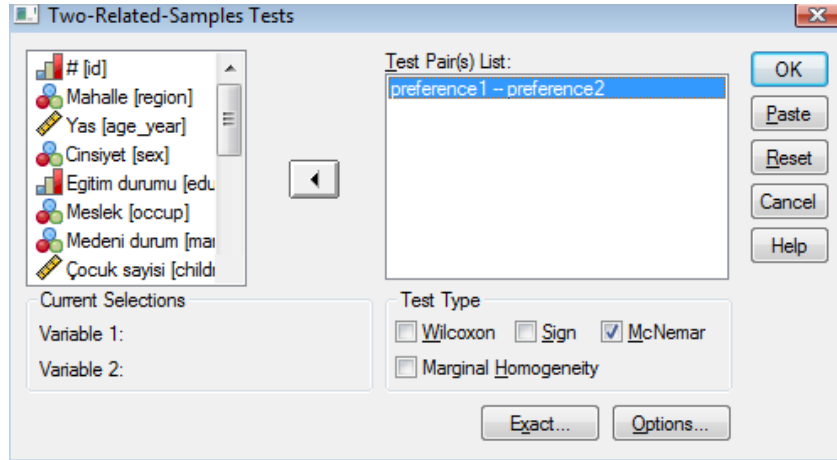
$$= 256 / 29$$

$$= 8,83$$

SPSS ile yapacak olursak: *Analyze > Nonparametric Tests > 2 Related Samples >*



[“*preference1*” ve “*preference2*” değişkenlerini işaretleyip “*Test Pair(s) List*” alanına geçirelim. “*McNemar*” kutusunu işaretleyelim]>ok.



Aşağıdaki çıktıları elde ederiz:

Danışmanlık öncesi ilaç tercihi & Danışmanlık sonrası ilaç tercihi

Danışmanlık öncesi ilaç tercihi	Danışmanlık sonrası ilaç tercihi	
	1	2
1	159	6
2	23	242

Test Statistics(b)

	Danışmanlık öncesi ilaç tercihi & Danışmanlık sonrası ilaç tercihi
N	430
Chi-Square(a)	8,828
Asymp. Sig.	0,003

a Continuity Corrected

b McNemar Test

Birinci tablo dört gözlü Ki Kare tablosunu vermektedir.

İkinci tabloda test istatistiği sonuçları görülüyor. Ki Kare değeri **8,828** olarak hesaplanmıştır.

4. Test istatistiğinden elde edilen değerin bilinen bir olasılık dağılımı ile karşılaştırılması:

Sayfa 460'dan χ^2 değerini baktığımızda:

<i>SD</i>	<i>İki yönlü p değeri</i>			
	0.10	0.05	0.01	0.001
1	2.706	3.841	6.635	10.827
2	4.605	5.991	9.210	13.815

SD=1'de Ki kare = 8,8 için p değerinin **0,01 ile 0,001** arasında bir değer olduğunu görüyoruz. Zaten SPSS'te de p değeri **0,003** olarak belirtilmişti.

5. P değerinin ve sonuçların yorumlanması

P değeri 0,05'ten küçük olduğu için H_0 hipotezini reddederiz.

Dolayısıyla diyabetik bireylerde danışmanlık öncesi ve sonrasında ilaç tercihleri arasında fark vardır. Danışmanlıktan sonra hastaların insülin tercihleri artmaktadır.

Bölüm 31: İkiden fazla nominal kategorisi olan bağımlı iki ölçüm (McNemar Bowker testi)

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun iki bağımlı gruptan elde edilen ve ikiden fazla nominal kategorisi olan kategorik verilerin karşılaştırılmasında kullanılan McNemar Bowker testi hakkında bilgi sahibi olması ve bu testi uygulayabilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- İkiden fazla nominal kategorisi olan iki bağımlı ölçüme örnek verebilmek
- SPSS’te McNemar Bowker testini yapabilmek

Kliniğimizde takip edilen diyabet hastalarının aldıkları tedavi çeşitlerinde bir ay öncesine göre anlamlı bir değişim olup olmadığını araştırmak istiyoruz. Her hastanın aldığı tedavi nominal olarak (tedavi almıyor/sadece sulfonilüre/ sulfonilüre+biguanid/insülin) kodlanmış ve birer ay arayla iki kez kaydedilmiştir (iki bağımlı ölçüm < 9. (H_0 : Diyabet hastalarının aldıkları tedavi çeşitlerinde bir ay öncesine göre anlamlı bir değişim olmamıştır).

McNemar-Bowker testi

Bağımlı iki ölçüm arasında ikiden fazla nominal kategorisi olan kategorik bir değişkeni karşılaştırmak için McNemar-Bowker testi kullanılabilir.

Testin mantığı

		Bir ay sonraki tedavi			Toplam
		Sulf	Sulf.+B.	İnsülin	
Bir ay önceski tedavi	Sulfonilüre	a (214)	b (5)	c (5)	224
	Sulfonilüre+Biguanid	d (8)	e (87)	f (9)	104
	İnsülin	g (10)	h (7)	i (8)	25
Toplam		232	99	22	353

Tablo: Gözlenen frekanslar.

İki ölçüm arasındaki fark, bir ay öncesine göre farklılık gösteren koyu renkli boyanmış b, c, d, f, g ve h hücreleridir. McNemar-Bowker testi bu farklılık gösteren hücreleri esas alır.

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

H_0 : Diyabet hastalarının aldıkları tedavi çeşitlerinde bir ay öncesine göre anlamlı bir değişim olmamıştır.

H_1 : Diyabet hastalarının aldıkları tedavi çeşitlerinde bir ay öncesine göre anlamlı bir değişim olmuştur.

2. Verilerin toplanması:

Veriler diyabet.sav verisetinde “therapy” ve “therapy2”değişkenleri altında kaydedildi.

3. İlgili sıfır hipotezi için test istatistiğinin hesaplanması¹:

$$B = \sum_{i < j} \frac{(n_{ij} - n_{ji})^2}{n_{ij} + n_{ji}} \sim \chi^2 \left(\frac{k(k-1)}{2} \right)$$

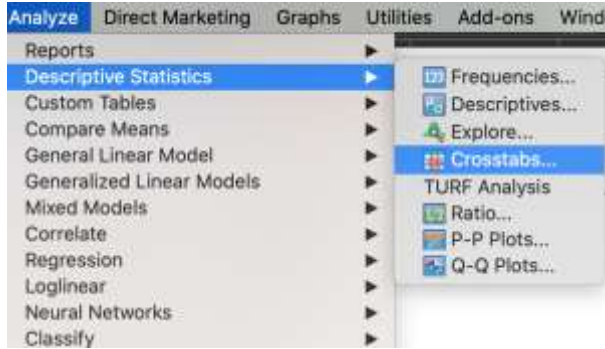
Bu test istatistiği k (kategori sayısı)=3, serbestlik derecesi ise $k(k-1)/2=3$ olacak şekilde Ki Kare dağılımı göstermektedir.

$$\chi^2 = [(5-8)^2 + (5-10)^2 + (9-7)^2] / [(5+8) + (5+10) + (9+7)]$$

$$= 38/44$$

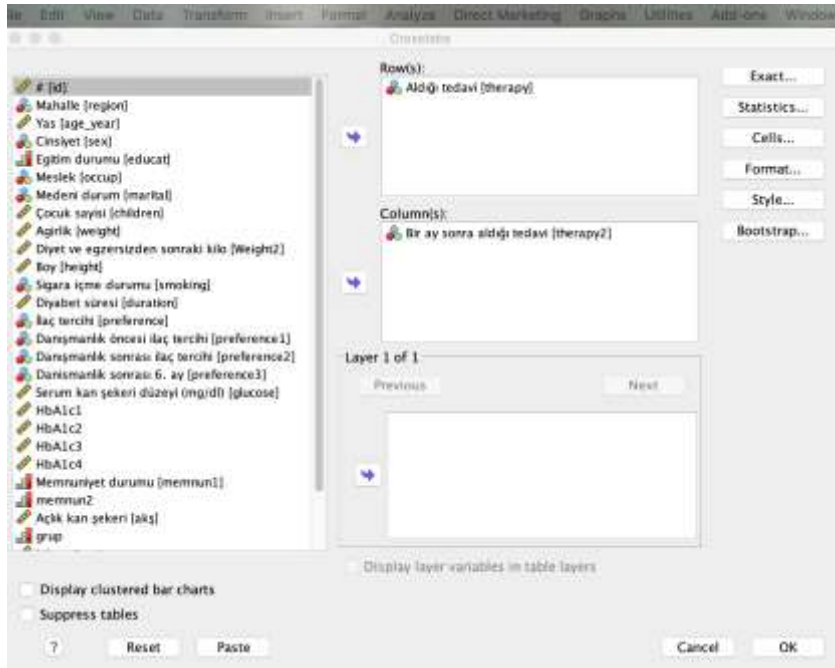
$$= 0,863$$

SPSS ile yapacak olursak: *Analyze>Descriptive Statistics>Crosstabs>*

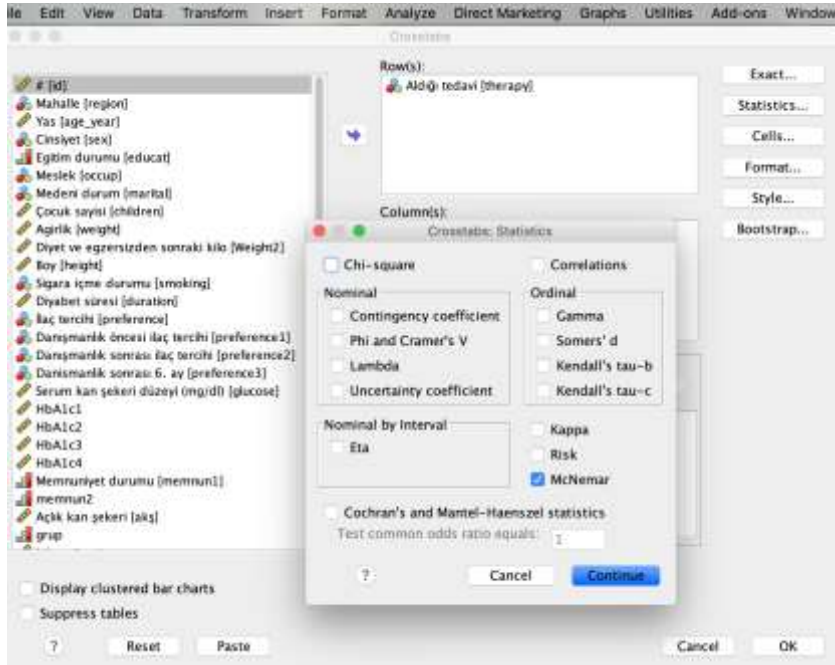


[Ros(s) kutusuna “therapy”, Column(s) kutusuna ise “therapy2” değişkenlerini alalım]

¹ Charles Zaiontz. McNemar-Bowker Test. <http://www.real-statistics.com/non-parametric-tests/mcnemar-bowker-test/>



>Statistics>[McNemar kutusunu işaretleyelim]



>Cells [Row kutucuğunu işaretleyelim]>Continue>ok.

Aşağıdaki çıktıları elde ederiz:

Aldığı tedavi * Bir ay sonra aldığı tedavi Crosstabulation

			Bir ay sonra aldığı tedavi		
			Sulphonylurea only	Sulphonylurea + Biguanides	Insulin
Aldığı tedavi	Sulphonylurea only	Count % within Aldığı tedavi	214 95.5%	5 2.2%	5 2.2%
	Sulphonylurea + Biguanides	Count % within Aldığı tedavi	8 7.7%	87 83.7%	9 8.7%
	Insulin	Count % within Aldığı tedavi	10 40.0%	7 28.0%	8 32.0%
Total		Count % within Aldığı tedavi	232 65.7%	99 28.0%	22 6.2%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
McNemar-Bowker Test	2.609	3	.456
N of Valid Cases	353		

Birinci tablo dokuz gözlü Ki Kare tablosunu vermektedir.

İkinci tabloda ise test istatistiği sonuçları görülüyor. Ki Kare değeri **2,609** olarak hesaplanmıştır (SPSS kullanılan formül hesaplama elle yaptığımızdan farklılık göstermektedir).

4. Test istatistiğinden elde edilen değer bilinen bir olasılık dağılımı ile karşılaştırılması:

Sayfa 460'den χ^2 değerini baktığımızda:

<i>sd</i>	İki yönlü P değeri			
	0.10	0.05	0.01	0.001
1	2.706	3.841	6.635	10.827
2	4.605	5.991	9.210	13.815
3	6.251	7.815	11.345	16.266
4	7.779	9.488	13.277	18.466

SD=3'de Ki kare = 2,6 için p değerinin **0,1'den** büyük olduğunu görüyoruz. SPSS'te de p değeri **0,456** olarak belirtilmişti.

5. P değerinin ve sonuçların yorumlanması

P değeri 0,05'ten büyük olduğu için H_0 hipotezini kabul ederiz. Dolayısıyla diyabet hastalarının aldıkları tedavi çeşitlerinde bir ay öncesine göre anlamlı bir değişim olmamıştır.

Bölüm 32: Kategorik Veri İki Bağımlı Grup (Kappa Uyum Analizi)

Hamit ACEMOĞLU

Amaç: Bu konu sonunda okuyucunun iki gözlemciden elde edilen kategorik verilerin uyum analizinde kullanılan Kappa uyum analizi hakkında bilgi sahibi olması ve bu testi uygulayabilmesi amaçlanmıştır.

Not: Cohen'in kappa ölçüsü sadece iki değerlendiriciyi ele alır. Eğer değerlendirici sayısı ikiden çoksa Kappa'ya benzeyen Fleiss'in kappa katsayısı¹ kullanılmalıdır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Testin hangi durumlarda kullanılacağını açıklayabilmek
- Testin mantığını açıklayabilmek
- κ katsayısını elle hesaplayabilmek
- SPSS'te Kappa uyum analizini yapabilmek
- SPSS çıktısını yorumlayabilmek

Kliniğimizde takip edilen diyabet hastalarının sağlık durumları iki farklı asistan tarafından değerlendirilmiştir. Asistanların hastaların durumlarını iyi, orta ve kötü olarak kodlamaları istenmiştir. Aynı hastaları değerlendiren asistanlar arasında uyumu araştırmak istiyoruz. İncelediğimiz değişken kategorik bir değişkendir (H_0 : İki asistanın hasta değerlendirme kararları arasında farklı yoktur).

Altın bir standardın olmadığı durumlarda bir hastaya ait materyal (görüntüleme yöntemleri, EKG vb tanı yöntemleri) üzerinde hastalığın varlığı veya yokluğu ya da derecesine karar veren birden fazla değerlendiricinin uyumu oldukça önemlidir.

Yukarıdaki senaryoda olduğu gibi iki farklı gözlemcinin elde ettiği kategorik değişkenlerin uyum analizinde sıklıkla Cohen'in Kappa uyum testi kullanılır. Gözlemciler arası uyumun ölçümünde basit uyum katsayısı (P_0) da hesaplanabilir. Ancak basit uyuma göre κ katsayısının avantajı, uyumun şans ile ortaya çıkması beklenen kısmını düzeltilmesidir.

¹ https://tr.wikipedia.org/wiki/Fleiss%27in_kappa_katsay%C4%B1s%C4%B1

Uyum analizinde eşit satır ve sütuna sahip (2x2, 3x3 gibi) çapraz tablolar kullanılarak basit uyum katsayısı (P_0) ve Cohen'in Kappa (κ) katsayısı hesaplanır. Aşağıda 3x3 çapraz tablo için kullanılan eşitlikler verilmiştir:

Uyum analizi için örnek tablo				
	Asistan 1			
Asistan 2	İyi	Orta	Kötü	Toplam
İyi	a	b	c	(a+b+c)
Orta	d	e	f	(d+e+f)
Kötü	g	h	k	(g+h+k)
Toplam	(a+d+g)	(b+e+h)	(c+f+k)	n

Basit uyum katsayısı eşitliği

$$(P_0) = \frac{a+e+k}{n}$$

P_0 'ın değeri 0 ile 1 arasında değişmekte ve 1'e yakın olması istenir. 0.5'in altında elde edilen P_0 değeri genellikle "şansa bağlı uyum" olarak kabul edilir. Basit uyum katsayısının zayıf yönü "a" veya "d" gözlemlerinden hangisinde yoğunlaşma olduğu hakkında bilgi vermemesidir.

Cohen'in kappa katsayısı (κ) eşitliği

$$\kappa = \frac{P_0 - P_e}{1 - P_e}$$

P_e şansa bağlı uyumu gösterir ve aşağıdaki eşitlikle hesaplanır.

$$P_e = \frac{[(a+d+g)(a+b+c) + (b+e+h)(d+e+f) + (c+f+k)(g+h+k)]}{n^2}$$

Kappa katsayısı -1 ile +1 arasında değişir. Tam uyum olduğunda $\kappa = 1$ olur. Gözlenen uyumun şansa bağlı uyuma eşit ya da ondan büyük olması durumunda $\kappa \geq 0$ iken, gözlenen uyumun şansa bağlı uyumundan küçük olması durumunda $\kappa < 0$ olur. Kappa katsayısının yorumlanabilir aralığı 0 ile +1 arasında olup, negatif ($\kappa < 0$) değerlerinin güvenilirlik açısından bir anlamı yoktur.

Kappa katsayısına göre uyumun yorumlanması

κ	Yorum
<0,00	Kötü, uyum yok
0,00-0,20	Zayıf
0,21-0,40	Fena değil
0,41-0,60	Orta
0,61-0,80	İyi
>0,80	Neredeyse mükemmel

Bu bölümün başında verilen senaryo için örnek verilerin aşağıdaki gibi olduğunu düşünelim:

Uyum analizi için örnek tablo				
	Asistan 1			
Asistan 2	İyi	Orta	Kötü	Toplam
İyi	193	20	5	218
Orta	15	123	5	143
Kötü	4	11	54	69
Toplam	212	154	64	430

Basit uyumu hesaplayacak olursak:

$$(P_0) = \frac{193+123+54}{430} = \mathbf{0,86}$$

Cohen'in kappa katsayısı (κ) için önce P_e yi hesaplayalım

$$P_e = \frac{[(212)(218)+(154)(143)+(64)(69)]}{430^2} = 0,393$$

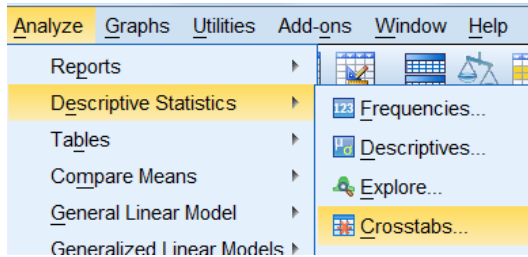
$$\kappa = \frac{0,86-0,393}{1-0,393} = \mathbf{0,77}$$

$\kappa = \mathbf{0,77}$ olması iki asistan arasında hastaların tedaviye uyumları iyi düzeyde olduğu şeklinde yorumlanabilir.

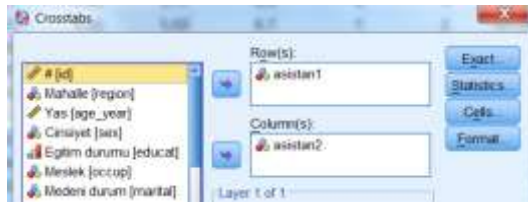
Örneğin SPSS ile yapılması

<https://drive.google.com/file/d/0B0ieO-Fy3z-5RjBzVkFsYWQ1MGs/view?usp=sharing> adresinden Diyabet.sav SPSS dosyasını indiriniz ve açınız.

Analyze>Descriptive Statistics>Crosstabs>



[“asistan1” değişkenini “Row(s)” alanına geçirelim. “asistan2” değişkenini ise “Column(s)” alanına geçirelim]>Statistics [Kappa kutusunu işaretleyelim]>Continue> Continue>ok.



Aşağıdaki çıktıları elde ederiz:

asistan1 * asistan2 Crosstabulation

Count		asistan2			Total
		İyi	Orta	Kötü	
asistan1	İyi	193	20	5	218
	Orta	15	123	5	143
	Kötü	4	11	54	69
Total		212	154	64	430

Symmetric Measures

	Value	Asymp. Std. Error ^a	Approx. T ^b	Approx. Sig.
Measure of Agreement Kappa	,770	,027	21,271	,000
N of Valid Cases	430			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

3x3 gözlü tabloda asistan1 ve asistan2'nin aynı hastaların tedaviye uyum değerlendirmeleri verilmiştir. Hemen altındaki tabloda Kappa katsayısı 0,77 olarak hesaplanmış ve elle hesaplamayla aynıdır. İki asistan arasında hastaların tedaviye uyum değerlendirmeleri “iyi” düzeydedir. P değeri <0,001 olarak hesaplanmış olup iki asistanının uyumlarının “0” dan farklı olduğu görülmektedir. Ancak p değeri kappa uyum analizlerinde nadiren raporlanmaktadır.

Alıştırma 11

Aşağıdaki tablo için basit uyum ve κ katsayılarını hesaplayarak yorumlayınız.

	Gözlemci1			
Gözlemci2	1	2	3	4
1	880	66	0	17
2	146	145	18	15
3	0	0	30	5
4	3	1	25	14

Çözüm için bakınız s. 477.

Bölüm 33: Kategorik Veri İki Bağımsız Grup (Ki Kare Analizi)

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun iki gruptan elde edilen kategorik verilerin karşılaştırılmasında kullanılan Ki Kare (χ^2) testi hakkında bilgi sahibi olması ve bu testleri uygulayabilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Kategorik veri iki oranı açıklayabilmek
- 2x2 çapraz tablo oluşturabilmek
- 2x2 çapraz tabloda beklenen frekansları hesaplayabilmek
- Ki kare testinin varsayımlarını açıklayabilmek
- SPSS’te Ki Kare testi yapabilmek
- Fisher exact testinin kullanım alanını açıklayabilmek

Kliniğimizde takip edilen diyabet hastalarının ilaç tercihlerinde cinsiyet açısından bir farklılık olup olmadığını araştırmak istiyoruz. İncelediğimiz değişken (ilaç tercihi; 1-insülin, 2-oral antidiyabetik) kategorik bir değişkendir. Erkekler ve bayanlar olmak üzere iki grubumuz söz konusu (H_0 : Diyabetik erkek ve bayanların ilaç tercihleri arasında fark yoktur).

Ki Kare (χ^2) testi

Birinci senaryomuzun analizi için Ki Kare testi kullanılabilir. Ki Kare testinde **2x2 frekans tablosu** söz konusudur. H_0 ’ın doğru olması halinde, tablonun her gözündeki **gözlenen frekanslarla beklenen frekanslar** birbirine yakın olmalıdır. Ki Kare, bu frekansları karşılaştırarak hesap yapar.

Varsayımlar

- Bağımsız iki gruptan n_1 ve n_2 sayıda örneklem olmalı.
- Her bireyden sadece bir kez veri alınmış olmalı.
- Tablodaki satırlar ve sütunlar birbirini dışlamalıdır (mutually exclusive). Yani, kategorilerden ikisi aynı anda seçilememelidir (örn.: birey ya insülin, ya oral antidiyabetiği tercih etmeli, ikisini birden seçememeli).
- Biraz muhafazakâr da olsa, genel yaklaşım, beklenen frekansların tablonun her bir gözü için 5’ten büyük olması şeklindedir.

Testin mantığı

Ki Kare testi gözlenen frekanslarla beklenen frekanslar arasındaki farka bakar:

Özellik	Grup 1	Grup 2	Toplam
Var	a	b	a+b
Yok	c	d	c+d
Toplam	$n_1=a+c$	$n_2=b+d$	$n=a+b+c+d$
Özelliğin orantısı	$p_1=a/n_1$	$p_2=b/n_2$	$p=(a+b)/n$

Diyabet.sav verisetinde “cinsiyet” ve “preference1” değişkenlerini kullanarak bir örnek yapacak olursak, verisetimizdeki frekanslar aşağıdaki gibidir:

Özellik	Erkek	Bayan	Toplam
İnsülin	71	94	165
Oral AD	165	100	265
Toplam	236	194	430
Özelliğin orantısı	$p_1=0,30$	$p_2=0,48$	$p=0,38$

Eğer gruplar arasında fark olmasaydı frekansların aşağıdaki gibi olmasını beklerdik:

Özellik	Erkek	Bayan	Toplam
İnsülin	90,6*	74,4	165
Oral AD	145,4	119,6	265
Toplam	236	194	430

* Beklenen frekansların hesaplanması:

430 kişiden 236’sı erkek ise;

165 insülin kullananandan kaç erkek olmalıdır?

$$430 \quad 236$$

$$165 \quad x$$

$$x = 165 \times 236 / 430 = 90,6 \text{ kişi.}$$

Görünüşe göre, gözlenen frekanslarla beklenen frekanslar arasında fark vardır. Bu farkın yüksekliği Ki Kare değerinin büyüklüğüyle doğru orantılıdır. Şimdi birinci senaryomuz için uygulama yapalım:

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

H_0 : Diyabetik erkek ve bayanların ilaç tercihleri arasında fark yoktur.

H_1 : Diyabetik erkek ve bayanların ilaç tercihleri arasında fark vardır.

2. Verilerin toplanması:

Veriler diyabet.sav verisetinde “sex” ve “preference1” değişkenleri altında kaydedildi.

3. İlgili sıfır hipotezi için test istatistiğinin hesaplanması:

$$\chi^2 = \sum \frac{\left(|G-B| - \frac{1}{2} \right)^2}{B}$$

G = Dört gözlü tablonun her gözündeki gözlenen frekanslar

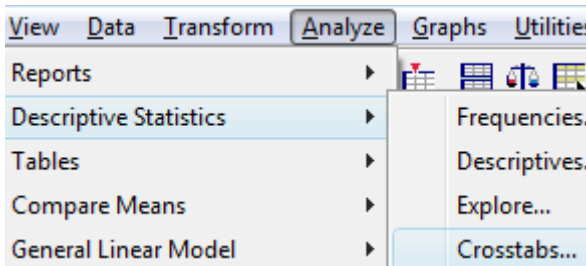
B = Dört gözlü tablonun her gözündeki beklenen frekanslar

Bu test istatistiği, serbestlik derecesi 1 olacak şekilde Ki Kare dağılımı göstermektedir.

$$\begin{aligned} \chi^2 &= ((|71-90,6| - 1/2)^2/90,6) + ((|94-74,4| - 1/2)^2/74,4) + ((|165-145,4| - 1/2)^2/145,4) + ((|100-119,6| - 1/2)^2/119,6) \\ &= (19,6 - 0,5)^2/90,6 + (19,6 - 0,5)^2/74,4 + (19,6 - 1,5)^2/145,4 + (19,6 - 0,5)^2/119,6 \\ &= 364,81/90,6 + 364,81/74,4 + 364,81/145,4 + 364,81/119,6 \\ &= 4,02 + 4,90 + 2,51 + 3,05 \\ &= \mathbf{14,48} \end{aligned}$$

SPSS ile yapacak olursak:

Analyze>Descriptive Statistics>Crosstabs>



[“preference1” değişkenini “Row(s)” alanına geçirelim. “sex” değişkenini ise “Column(s)” alanına geçirelim]>Statistics [Chi-square

kutusunu işaretleyelim]>Continue>Cells [“Observed”, “Expected” ve “Row” kutularını işaretleyelim]>Continue>ok.

Aşağıdaki çıktıları elde ederiz:

Danışmanlık öncesi ilaç tercihi * Cinsiyet Crosstabulation

			Cinsiyet		Total
			Male	Female	
Danışmanlık öncesi ilaç tercihi	İnsülin	Count	71	94	165
		Expected Count	90,6	74,4	165,0
		% within Danışmanlık öncesi ilaç tercihi	43,0%	57,0%	100,0%
	Oral AD	Count	165	100	265
		Expected Count	145,4	119,6	265,0
		% within Danışmanlık öncesi ilaç tercihi	62,3%	37,7%	100,0%
Total	Count		236	194	430
	Expected Count		236,0	194,0	430,0
	% within Danışmanlık öncesi ilaç tercihi		54,9%	45,1%	100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	15,192 ^b	1	,000		
Continuity Correction ^a	14,425	1	,000		
Likelihood Ratio	15,215	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	15,157	1	,000		
N of Valid Cases	430				

a. Computed only for a 2x2 table

b. 0 cells (0%) have expected count less than 5. The minimum expected count is 74,44.

Birinci tablodan görüldüğü gibi gözlenen frekanslar ve beklenen frekanslar arasında belirgin farklılık vardır. Erkeklerin %43'ü insülini tercih ederken bayanların %57'si tercih etmesi, arada bir fark olduğunu düşündürüyor.

İkinci tabloda Ki Kare değerine bakıldığında “continuity correction” (normal dağılıma yaklaşması için süreklilik düzeltmesi) sonrasında değerin **14,425** olduğu görülüyor. Serbestlik derecesi (df) 1 olacak şekilde iki yönlü *p* değeri sıfıra yakın (0,0001) olarak bildirilmiştir.

Ki kare varsayımlarından beklenen frekansların hepsinin 5'ten büyük olması kuralının karşılanamaması halinde **Fisher exact** testi uygulanır.

Ki kare analizi yapıldığında ikinci tabloda da görüldüğü gibi Fisher exact test sonucu otomatik olarak verilir.

4. Test istatistiğinden elde edilen değerin bilinen bir olasılık dağılımı ile karşılaştırılması:

Sayfa 460'dan χ^2 değerini baktığımızda:

<i>SD</i>	<i>İki yönlü p değeri</i>			
	0.10	0.05	0.01	0.001
1	2.706	3.841	6.635	10.827
2	4.605	5.991	9.210	13.815

SD=1'de Ki kare = **14,4** için p değerinin **0,001**'den daha küçük olduğunu görürüz. Zaten SPSS'te de p sıfıra yakın olarak belirtilmişti.

5. P değerinin ve sonuçların yorumlanması

P değeri 0,05'ten küçük olduğu için H_0 hipotezini reddederiz.

Dolayısıyla diyabetik erkekler ve kadınlar arasında antidiyabetik ilaç tercihleri açısından fark vardır.

Bölüm 34: Kategorik Veri İkiden Fazla Bağımsız Grup (Ki Kare Analizi)

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun ikiden fazla bağımsız grubu bir kategorik değişken açısından karşılaştırmak için Ki Kare testinin kullanımı ve ikiden fazla grubu olan kategorik değişkenlerde artan veya azalan bir ilişki (trend) olup olmadığını incelemek için Ki Kare testinin kullanımı hakkında bilgi sahibi olması ve bu analizleri yapabilmesi amaçlanmıştır

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Kategorik veri ikiden fazla grubu açıklayabilmek
- Ki Kare testi varsayımlarını açıklayabilmek
- Ki Kare testinin mantığını açıklayabilmek
- Fisher exact testinin uygulama alanını açıklayabilmek
- Ki kare trend testinin kullanım alanını açıklayabilmek
- SPSS'te ki kare testi yapabilmek
- SPSS'te ki kare ve ki kare trend çıktısını yorumlayabilmek

İki farklı senaryo düşünelim:

- *Senaryo 1: Farklı mahallelerden gelen diyabet hastalarımızın medeni durumları arasında fark olup olmadığını inceleyelim. İncelediğimiz değişkenin (medeni durum) dört kategorisi vardır (1-bekar, 2-evli, 3-boşanmış, 4-dul). Karşılaştırmayı 3 farklı grupta yapıyoruz (1-güney mah., 2-doğu mah., 3-merkez mah.). (H_0 : Mahalleler arasında bireylerin medeni durumu açısından fark yoktur.)*
- *Senaryo 2: Örneklemimizdeki iki kategorik değişken arasında lineer bir ilişki olup olmadığına bakmak istiyoruz. Değişkenlerimizden biri ikili kategorik bir değişken (1-evet, 2-hayır gibi), diğeri ise kategorik sıralı (ordinal) bir değişken olmalıdır (yaş grupları, eğitim durumları gibi). Örneğin, örneklemimizde eğitim durumu ile sigara içme durumu arasında bir trend (lineer ilişki) olup olmadığını inceleyelim (H_0 : Eğitim durumu ile sigara içme durumu arasında lineer bir ilişki yoktur).*

Çoklu gruplarda Ki Kare (χ^2) testi

Birinci senaryomuzda 4x3 gözlü bir Ki Kare tablosu oluşacaktır.

Varsayımlar

- Her bireyden sadece bir kez veri alınmış olmalı.

- Tablodaki satırlar ve sütunlar birbirini dışlamalıdır (mutually exclusive). Yani, kategorilerden ikisi aynı anda seçilememelidir.
- Beklenen frekansların en az %80'i 5 veya 5'ten büyük olmalıdır.

Testin mantığı

Çok gözlü Ki Kare de 2x2 durumunda olduğu gibi gözlenen ve beklenen frekanslar arasındaki farka dayalıdır.

Varsayımların karşılanamaması durumunda

Tablodaki gözlerden %20'den fazlasında beklenen frekansın 5'ten küçük olması durumunda gruplar birleştirilmeye çalışılır (örn. lise mezunları ve üniversite mezunları bir grup yapılabilir). Yine de varsayım karşılanamazsa gruplar birleştirilmeye devam edilir. Ta ki, 2x2'lik bir Ki Kare tablosu elde edene kadar. Yine de varsayım karşılanmıyorsa Fisher exact testi hesaplanır. Ancak, bazı istatistik programları büyük Ki Kare tabloları için de Fisher exact testi sonuçlarını vermektedir.

Birinci senaryomuzu diyabet.sav veri setine göre uygulayacak olursak:

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

H_0 : Mahalleler arasında bireylerin medeni durumu açısından fark yoktur

H_1 : Mahalleler arasında bireylerin medeni durumu açısından fark vardır.

2. Verilerin toplanması:

Veriler diyabet.sav veri setinde “mahalle” ve “medeni” değişkenleri altında kaydedildi.

3. İlgili sıfır hipotezi için *test istatistiğinin* hesaplanması:

$$\chi^2 = \sum \frac{(G - B)^2}{B}$$

G = Ki Kare tablosunun her gözündeki gözlenen frekanslar

B = Ki Kare tablosunun her gözündeki beklenen frekanslar

Bu test istatistiği, serbestlik derecesi (sıra-1)x(sütun-1) olacak şekilde Ki Kare dağılımı göstermektedir.

Ki Kare tablosuna baktığımızda;

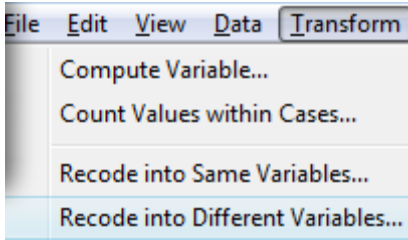
Mahalle * Medeni durum Crosstabulation

			Medeni durum				Total
			Bekar	Evlü	Başanmış	Dul	
Mahalle	Güney	Count	6	151	10	20	187
		Expected Count	13,4	142,0	11,1	20,5	187,0
		% within Mahalle	3,2%	80,7%	5,3%	10,7%	100,0%
	Doğu	Count	7	23	5	5	40
		Expected Count	2,9	30,4	2,4	4,4	40,0
		% within Mahalle	17,5%	57,5%	12,5%	12,5%	100,0%
	Merkez	Count	17	145	10	21	193
		Expected Count	13,8	146,6	11,5	21,1	193,0
		% within Mahalle	8,8%	75,1%	5,2%	10,9%	100,0%
Total	Count		30	319	25	46	420
	Expected Count		30,0	319,0	25,0	46,0	420,0
	% within Mahalle		7,1%	76,0%	6,0%	11,0%	100,0%

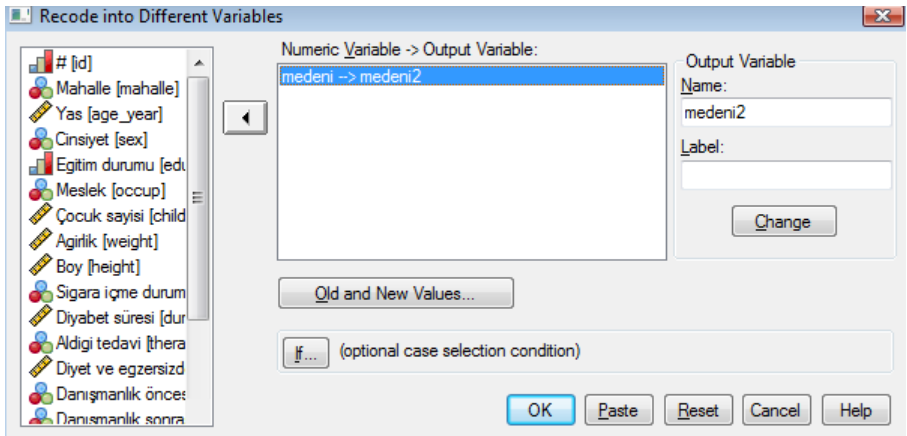
Üç hücrede (%25) beklenen frekansın 5'in altında olduğunu görüyoruz. Dolayısıyla Ki Kare için gerekli varsayım karşılanmamaktadır.

“Başanmış” ve “Dul” gruplarını tek bir grupta birleştirmemiz çözüm olabilir:

Transform > Recode into Different Variables >



[“medeni” değişkenini “Numeric Variable” alanına geçirelim. “Name” alanına “medeni2” yazalım] > Change > Old and New Values >



[“Old Value” kutusuna 4, “New Value” kutsuna 3 yazıp “Add” butonuna tıklayalım. Sonra “Old Value” alanından “All other values”, “New Value” alanından “Copy old value(s)” işaretleyip “Add” butonuna tıklayalım]

The dialog box 'Recode into Different Variables: Old and New Values' is shown. On the left, under 'Old Value', the 'Value:' radio button is selected with the value '4' entered. On the right, under 'New Value', the 'Value:' radio button is selected with the value '3' entered. Below these, the 'Old -> New:' list shows '4 -> 3' and 'ELSE -> Copy'. The 'Add' button is highlighted. At the bottom right, there are checkboxes for 'Output variables are strings' (unchecked) and 'Convert numeric strings to numbers' (unchecked), along with a 'Width' field set to '8'. The 'Continue', 'Cancel', and 'Help' buttons are at the bottom.

Continue>OK

Veri setinde “medeni2” adında yeni bir değişken oluşacaktır.

“Variable View” penceresinden “medeni2” değişkeni için gerekli ayarları yapalım (“missing” alanına 9 yazmayı unutmayınız):

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	id	Numeric	8	0	#	None	None	3	Right	Ordinal
2	mahalle	Numeric	8	0	Mahalle	(1, Güney)...	None	5	Right	Nominal
3	age_year	Numeric	8	0	Yas	None	None	7	Right	Scale
4	sex	Numeric	8	0	Cinsiyet	(1, Male)...	None	3	Right	Nominal
5	educat	Numeric	8	0	Eğitim durumu	(1, None (ilker)...	None	6	Right	Ordinal
6	occup	Numeric	8	0	Meslek	(1, Profession...	None	5	Right	Nominal
7	medeni	Numeric	8	0	Medeni durum	(1, Bekar)...	9	5	Right	Nominal
8	children	Numeric	8	0	Çocuk sayısı	None	None	6	Right	Scale
9	weight	Numeric	8	1	Ağırlık	None	None	5	Right	Scale
10	height	Numeric	8	1	Boy	None	None	5	Right	Scale
11	smoking	Numeric	8	0	Sigara içme du	(1, Yes)...	None	7	Right	Nominal
12	duration	Numeric	8	1	Dişabet süresi	None	None	6	Right	Scale
13	therapy	Numeric	8	0	Aldığı tedavi	(0, No pharma...	None	15	Left	Nominal
14	Weight2	Numeric	8	2	Diyet ve egzers	None	None	8	Left	Scale
15	preference	Numeric	8	0	Danışmanlık o	(1, Ynsülin)...	None	12	Right	Nominal
16	preference	Numeric	8	0	Danışmanlık s	(1, Ynsülin)...	None	12	Right	Nominal
17	yearsnp	Numeric	8	0	Yapı grupları	(1, 39 yapı ve	None	10	Right	Ordinal
18	medeni2	Numeric	8	0	Medeni durum	(1, Bekar)...	9	10	Right	Nominal

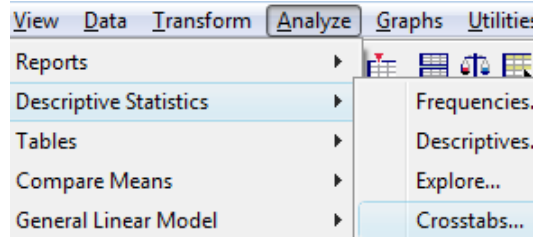
Yeni “medeni2” değişkenimizde sadece 1 hücrede (%11,1) beklenen frekans 5’in altında olduğundan işlem yapabiliriz.

$$\chi^2 = (6-13,4)^2/13,4 + (151-142)^2/142 + (30-31,6)^2/31,6 + (7-2,9)^2/2,9 + (23-30,4)^2/30,4 + (10-6,8)^2/6,8 + (17-13,8)^2/13,8 + (145-146,6)^2/146,6 + (31-32,6)^2/32,6$$

=14,68

SPSS ile yapacak olursak:

Analyze>Descriptive Statistics>Crosstabs>



[“mahalle” değişkenini “Row(s)” alanına geçirelim. “medeni2” değişkenini ise “Column(s)” alanına geçirelim]>Statistics [Chi-square kutusunu işaretleyelim]>Continue>Cells [“Observed”, “Expected” ve “Row” kutularını işaretleyelim]>Continue>OK

Aşağıdaki çıktıları elde ederiz:

Mahalle * Medeni durum Crosstabulation

			Medeni durum			Total
			Bekar	Evlü	Bosanmış	
Mahalle	Güney	Count	6	151	30	187
		Expected Count	13,4	142,0	31,6	187,0
		% within Mahalle	3,2%	80,7%	16,0%	100,0%
	Dogu	Count	7	23	10	40
		Expected Count	2,9	30,4	6,8	40,0
		% within Mahalle	17,5%	57,5%	25,0%	100,0%
	Merkez	Count	17	145	31	193
		Expected Count	13,8	146,6	32,6	193,0
		% within Mahalle	8,8%	75,1%	16,1%	100,0%
Total	Count	30	319	71	420	
	Expected Count	30,0	319,0	71,0	420,0	
	% within Mahalle	7,1%	76,0%	16,9%	100,0%	

Chi-Square Tests

	Value	Df	Asymp. Sig. (2-sided)
Pearson Chi-Square	14,900(a)	4	,005
Likelihood Ratio	14,114	4	,007
Linear-by-Linear Association	1,274	1	,259
N of Valid Cases	420		

a 1 cells (11,1%) have expected count less than 5. The minimum expected count is 2,86.

4. Test istatistiğinden elde edilen değerin bilinen bir olasılık dağılımı ile karşılaştırılması:

Sayfa 460'dan χ^2 değerini baktığımızda:

<i>SD</i>	<i>İki yönlü p değeri</i>			
	0.10	0.05	0.01	0.001
1	2.706	3.841	6.635	10.827
2	4.605	5.991	9.210	13.815
3	6.251	7.815	11.345	16.266
4	7.779	9.488	13.277	18.466
5	9.236	11.070	15.086	20.515

$SD=(3-1) \times (3-1)=4$ 'te Ki kare = **14,6** için p değerinin **0,01** ile **0,001** arasında olduğunu görürüz. SPSS'te de $p = 0,005$ olarak verilmiştir.

5. P değerinin ve sonuçların yorumlanması

P değeri 0,05'ten küçük olduğu için H_0 hipotezini reddederiz.

Dolayısıyla mahalleler arasında medeni durum açısından farklılıklar vardır. SPSS çıktısını incelediğimizde bekârların doğuda, evlilerin ise güneyde fazla olduğunu görüyoruz.

Ki Kare Trend testi

Çok gruplu kategorik bir değişkenimizin ikili bir değişken açısından artan veya azalan bir trend (eğilim) izleyip izlemediğini araştırmak için (yani lineer bir ilişki olup olmadığına bakmak için) Ki Kare Trend testi yapılabilir. Örneğin, bir ilaca cevap verenlerin orantısının yaşla birlikte artıp artmadığını incelemek isteyebiliriz.

SPSS’te Ki Kare analizi yapıldığında çıktılar arasında lineer ilişki de otomatik olarak verilmektedir.

Konu başında verdiğimiz senaryo için bir uygulama yapacak olursak:

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

H_0 : Eğitim durumu ile sigara içme durumu arasında lineer bir ilişki (trend) yoktur.

H_1 : Eğitim durumu ile sigara içme durumu arasında lineer bir ilişki vardır.

2. Verilerin toplanması:

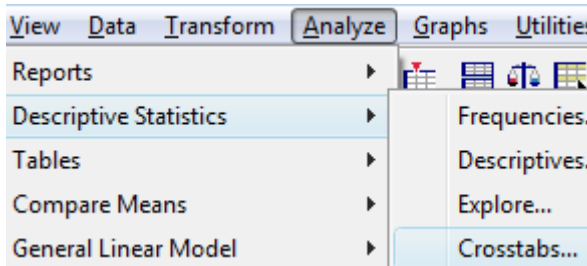
Veriler diyabet.sav verisetinde “educat” ve “smoking” değişkenleri altında kaydedildi.

3. İlgili sıfır hipotezi için test istatistiğinin hesaplanması:

Test istatistiği formülü karmaşık olduğundan burada verilmeyecektir.

SPSS ile yapacak olursak:

Analyze>Descriptive Statistics>Crosstabs>



[“educat” değişkenini “Row(s)” alanına geçirelim. “smoking” değişkenini ise “Column(s) alanına geçirelim]>Statistics [Chi-square kutusunu işaretleyelim]>Continue>Cells [“Observed”, “Expected” ve “Row” kutularını işaretleyelim]>Continue>OK

Aşağıdaki çıktıları elde ederiz:

Eğitim durumu * Sigara içme durumu Crosstabulation

			Sigara içme durumu		Total
			Yes	No	
Eğitim durumu	Yok	Count	8	201	209
		Expected Count	17,1	191,9	209,0
		% within Eğitim durumu	3,8%	96,2%	100,0%
	İlkokul	Count	7	80	87
		Expected Count	7,1	79,9	87,0
		% within Eğitim durumu	8,0%	92,0%	100,0%
	Ortaokul	Count	11	37	48
		Expected Count	3,9	44,1	48,0
		% within Eğitim durumu	22,9%	77,1%	100,0%
	Lise ve üzeri	Count	7	52	59
		Expected Count	4,8	54,2	59,0
		% within Eğitim durumu	11,9%	88,1%	100,0%
Total	Count	33	370	403	
	Expected Count	33,0	370,0	403,0	
	% within Eğitim durumu	8,2%	91,8%	100,0%	

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	20,199 ^a	3	,000
Likelihood Ratio	17,133	3	,001
Linear-by-Linear Association	11,440	1	,001
N of Valid Cases	403		

a. 2 cells (25,0%) have expected count less than 5. The minimum expected count is 3,93.

Birinci tablo çok gözlü Ki Kare tablosunu vermektedir. Görüldüğü gibi sigara içme durumu eğitim durumuyla birlikte giderek artmaktadır. Lise ve üzeri bireylerde tekrar düşme görülmekle birlikte lineerlik açısından yine de anlamlı olmaktadır. Zira, ikinci tablodaki test istatistiğinde “Linear-by-Linear Association” 11,4; p değeri ise 0,001 olarak verilmiştir (tam değeri 0,000718).

4. Test istatistiğinden elde edilen değerlerin bilinen bir olasılık dağılımı ile karşılaştırılması:

Sayfa 460’dan χ^2 değerini baktığımızda:

<i>SD</i>	<i>İki yönlü p değeri</i>			
	0.10	0.05	0.01	0.001
1	2.706	3.841	6.635	10.827
2	4.605	5.991	9.210	13.815

SD=1'de Ki kare = 11,4 için p değerinin **0,001**'den küçük bir değer olduğunu görüyoruz.

5. *P değeri* ve sonuçların yorumlanması

P değeri 0,05'ten küçük olduğu için H_0 hipotezini reddederiz.

Dolayısıyla diyabetik bireylerde yaşla birlikte sigara içme oranları artmaktadır.

Ham veriler olmadan SPSS ile Ki Kare analizi

Bazen verilerin tamamı elimizde olmayıp sadece özetlenmiş verilerden analiz yapmamız veya yapılmış bir analizin sağlamasını yapmamız gerekebilir.

Örneğin Konuralp Tıp Dergisinde yayınlanan, Aydan ve arkadaşlarının “Evde sağlık hizmeti alan yaşlı hastalarda beslenme durumunun değerlendirilmesi” başlıklı makalelerinde yaptıkları Ki Kare analizini ele alalım:

http://konuralptipdergi.duzce.edu.tr/Dokumanlar/sayi6-3/07_KTD-2014-6-3.pdf

Yazarlar Tablo 2’de kadın ve erkekleri Mini Nutrisyonel Değerlendirme (MNA) formunun sonucuna göre karşılaştırmışlar ve p değerini 0,183 olarak rapor etmişler:

Tablo 2. MNA Toplamına Göre Cinsiyet Dağılımları Değerlendirilmesi

		MNA toplam			p
		Malnütrisyonlu (n=59) n(%)	Malnütrisyon Riski (n=70) n(%)	Normal Durum (n=49) n(%)	
Cinsiyet	Kadın	36 (61,0)	47 (67,1)	38 (77,6)	0,183
	Erkek	23 (39,0)	23 (32,9)	11 (22,4)	

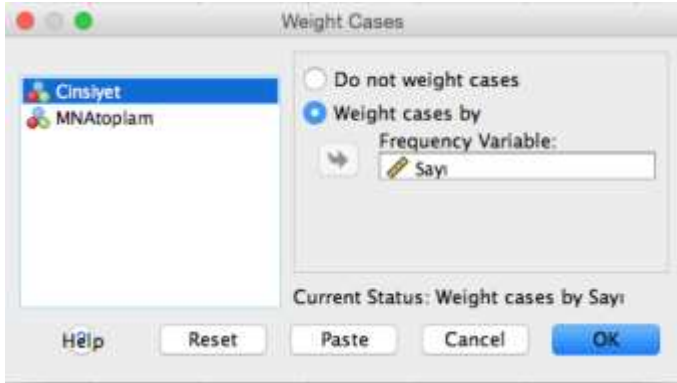
Pearson Chi-Square Test

Bu analizi SPSS ile yapmaya çalışalım:

“Cinsiyet” değişkeninin iki kategorisi, “MNA toplam” değişkeninin ise üç kategorisi var. Dolayısıyla, altı gözlü bir Ki Kare tablosu söz konusu. Bu tablo kombinasyonlarının değerlerini SPSS verisetimizde oluşturalım. Her hücrenin gözlenen frekans değerini gireceğimiz bir de “Sayı” değişkeni tanımlayalım:

	Cinsiyet	MNA toplam	Sayı	var
1	Kadın	Malnütrisyonlu	36	
2	Kadın	Malnütrisyon riski	47	
3	Kadın	Normal durum	38	
4	Erkek	Malnütrisyonlu	23	
5	Erkek	Malnütrisyon riski	23	
6	Erkek	Normal durum	11	

Şimdi verilerimizi “Sayı” değişkenine göre ağırlıklandıralım (Data>Weight Cases yolundan):



Şimdi Ki Kare analizi yapalım:

Analyze>Descriptive Statistics>Crosstabs>[“Cinsiyet” değişkenini “Row(s)” alanına geçirelim. “MNAtoplamlar” değişkenini ise “Column(s)” alanına geçirelim]>Statistics [Chi-square kutusunu işaretleyelim]>Continue>Cells [“Row” kutusunu işaretleyelim]>Continue>OK.

Aşağıdaki çıktıyı elde ederiz:

Cinsiyet * MNAtoplamlar Crosstabulation						
			MNAtoplamlar			Total
			Malnütrisyonlu	Malnütrisyon riski	Normal durum	
Cinsiyet	Kadın	Count	36	47	38	121
		% within Cinsiyet	29.8%	38.8%	31.4%	100.0%
	Erkek	Count	23	23	11	57
		% within Cinsiyet	40.4%	40.4%	19.3%	100.0%
Total		Count	59	70	49	178
		% within Cinsiyet	33.1%	39.3%	27.5%	100.0%

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3.399 ^a	2	.183
Likelihood Ratio	3.492	2	.175
Linear-by-Linear Association	3.291	1	.070
N of Valid Cases	178		
a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 15.69.			

Görüldüğü gibi, Pearson Ki Kare değeri 3,399, p değeri ise **0,183**'tür.

Başka bir örneğin SPSS Syntax ile yapılışını görmek isterseniz

<http://www.theanalysisfactor.com/chi-square-test-proportions/> adresine bakabilirsiniz.

Bölüm 35: Kategorik Veri İkiden Fazla Bağımlı Grup (Cochran Q Analizi)

Hamit ACEMOĞLU

Amaç: Bu konu sonunda okuyucunun ikiden fazla bağımlı gruptan elde edilen dikotom kategorik verilerin karşılaştırılmasında kullanılan Cochran's Q testi hakkında bilgi sahibi olması ve bu testleri uygulayabilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Cochran's Q testinin varsayımlarını açıklayabilmek
- Kategorik veri ikiden fazla bağımlı gruba örnek verebilmek
- SPSS'te Cochran's Q testi yapabilmek
- Cochran's Q testinin kullanım alanını açıklayabilmek

Diyabetik hastalarımıza kan şekeri kontrolünün önemi ve gerektiğinde insülin kullanılması konusunda bir eğitim verdik. Örneklemimizdeki bireylerin ilaç tercihlerinin eğitimden sonra birinci ve altıncı aylarda değişip değişmediğini incelemek istiyoruz. İncelediğimiz değişken (ilaç tercihi; 1-insülin, 2-oral antidiyabetik) kategorik dikotom bir değişkendir. Danışmanlık öncesi ve danışmanlık sonrası birinci ve 6. aylarda aynı bireylerden veri topladık. Dolayısıyla üç bağımlı grup söz konusudur.

Cochran's Q testi varsayımları

İkiden fazla bağımlı gruplardan elde edilen veri dikotom ise örneğin evet-hayır, var-yok gibi. Bu durumda $2 \times k$ gözlü çapraz tablo elde ederiz. Kullanılacak test ise Cochran's Q testidir.

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

H_0 : Diyabet hastalarının ilaç tercihlerinde eğitim öncesi, eğitim sonrası birinci ve altıncı aylar arasında fark yoktur

H_1 : Diyabet hastalarının ilaç tercihlerinde eğitim öncesi, sonrası birinci ve altıncı aylardan en az birinde fark vardır.

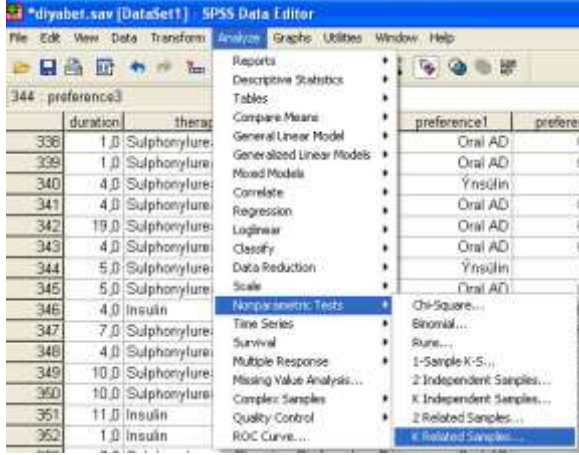
2. Verilerin toplanması:

Veriler diyabet.sav verisetinde “preference1”, “preference2” ve “preference3” değişkenleri altında kaydedildi.

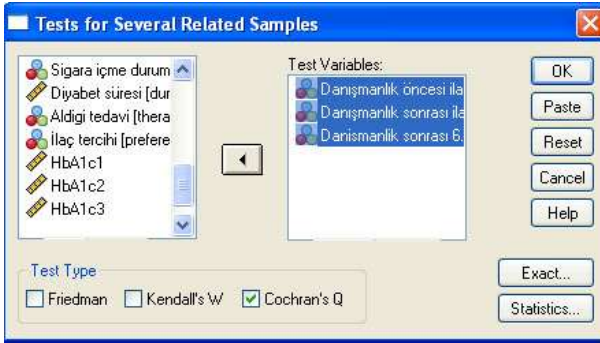
3. İlgili sıfır hipotezi için test istatistiğinin hesaplanması:

SPSS ile yapacak olursak:

Analyze>Nonparametric Tests>Related Samples>



[“preferenc1”, “preferenc2” ve “preferenc3” değişkenlerini işaretleyip “Test Variables” alanına geçirelim. “Cochran’s Q” kutusunu işaretleyelim]>OK



Aşağıdaki çıktıları elde ederiz:

Cochran Test

Frequencies

	Value	
	1	2
Danışmanlık öncesi ilaç tercihi	165	265
Danışmanlık sonrası ilaç tercihi	182	248
Danışmanlık sonrası 6. ay	192	238

Test Statistics

N	430
Cochran's Q	14,711(a)
df	2
Asymp. Sig.	,001

a 1 is treated as a success.

Birinci tablo 2*3 gözlü frekans tablosunu vermektedir.

İkinci tabloda test istatistiği sonuçları görülüyor. Cochran's Q değeri **14,711** olarak hesaplanmıştır.

4. Test istatistiğinden elde edilen değerin bilinen bir olasılık dağılımı ile karşılaştırılması:

14,711 olarak hesaplanan Ki Kare değeri teorik dağılımdan hesaplanan tablo değeri ile karşılaştırıldığında p değerinin 0,01'den küçük olduğu görülmektedir.

5. P değerinin ve sonuçların yorumlanması

P değeri 0,05'ten küçük olduğu için H_0 hipotezini reddederiz. Dolayısıyla diyabetik bireylerde danışmanlık öncesi sonrası birinci ve altıncı aylardan en az birinde ilaç tercihleri arasında fark vardır.

Bölüm 36: Kategorik Veri İkidenden Fazla Bağımlı Kategori (Marginal Homogeneity Testi)

Hamit ACEMOĞLU

Amaç: Bu konu sonunda okuyucunun iki bağımlı gruptan elde edilen ordinal kategorik verilerin karşılaştırılmasında kullanılan Marginal Homogeneity testi hakkında bilgi sahibi olması ve bu testi uygulayabilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Marginal Homogeneity testinin varsayımlarını açıklayabilmek
- Kategorik sıralı (ordinal) veri iki bağımlı gruba örnek verebilmek
- SPSS’te Marginal Homogeneity testini yapabilmek

Diyabetik hastalarımıza hizmet veren sağlık kuruluşu ile ilgili bir kalite iyileştirme programının uygulandığını varsayalım. Program uygulanmadan önce ve uygulandıktan sonra hasta memnuniyeti “hiç memnun değilim, az memnunum, memnunum ve çok memnunum” şeklinde sorulmuştur. Dolayısıyla ordinal kategorik veri elde edilecektir.

Marginal homogeneity test varsayımları

İki bağımlı gruplardan elde edilen veri ordinal ise örneğin hastalık şiddeti hafif, orta ve ağır gibi. Bu durumda $K \times K$ gözlü çapraz tablo elde ederiz. Analizde kullanılacak test ise marginal homogeneity testidir.

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

H_0 : Diyabet hastalarının kalite iyileştirme programı uygulanmadan önce ve sonra verilen sağlık hizmetlerinden memnuniyet dağılımları arasında fark yoktur

H_1 : Diyabet hastalarının kalite iyileştirme programı uygulanmadan önce ve sonra verilen sağlık hizmetlerinden memnuniyet dağılımları arasında fark vardır.

2. Verilerin toplanması:

Veriler diyabet.sav veri setinde “menuniyet1” ve “menuniyet2” değişkenleri altında kaydedildi. Veriler kaydedilirken azdan çoğa sıralı olmasına dikkat edilmelidir. Hiç memnun olmayanlar 1, az memnun olanlar 2 memnun olanlar 3 ve çok memnun olanlar 4 gibi.

3. İlgili sıfır hipotezi için test istatistiğinin hesaplanması:

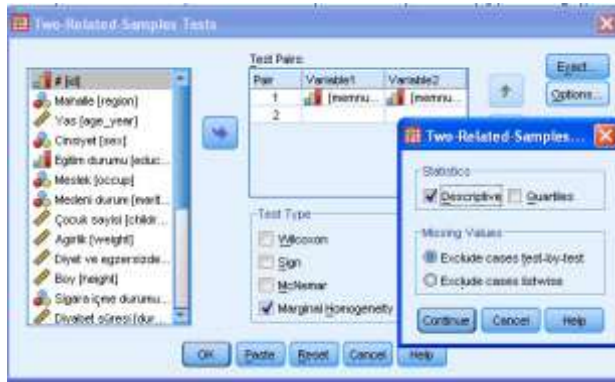
Veriler diyabet.sav veri setinde “therapy” değişkeni altında kaydedildi. Veriler kaydedilirken azdan çoğa sıralı olmasına dikkat edilmelidir. Hiç

memnun olmayanlar 1, az memnun olanlar 2 memnun olanlar 3 ve çok memnun olanlar 4 gibi.

SPSS ile yapacak olursak:

Analyze>Nonparametric Tests>2 Related Samples>

[“memnuniyet1” ve “memnuniyet2” değişkenlerini işaretleyip “Test Pair(s) List” alanına geçirelim. “Marginal Homogeneity” kutusunu ve , Options> Descriptive işaretleyelim]>ok.



Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
memnun2	430	3,00	1,001	1	4
memnun1	430	2,60	1,021	1	4

Marginal Homogeneity Test

	memnun2 & memnun1
Distinct Values	4
Off-Diagonal Cases	172
Observed MH Statistic	559,000
Mean MH Statistic	473,000
Std. Deviation of MH Statistic	6,557
Std. MH Statistic	13,115
Asymp. Sig. (2-tailed)	,000

4. Test istatistiğinden elde edilen değerin bilinen bir olasılık dağılımı ile karşılaştırılması:

577,9 olarak hesaplanan Ki Kare değeri teorik dağılımdan hesaplanan tablo değeri ile karşılaştırıldığında p değerinin 0,001’den küçük olduğu görülmektedir.

5. *P* değerinin ve sonuçların yorumlanması

P değeri 0,05'ten küçük olduğu için H_0 hipotezini reddederiz. Dolayısıyla diyabetik bireylerde kalite iyileştirme programı öncesi ve sonrasında memnuniyet dağılımları arasında fark vardır. Programdan sonra hastaların memnuniyet düzeyleri artmaktadır.

Bölüm 37: Kategorik Veri İkiden Fazla Bağımsız Kategori (Ki Kare)

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun ikiden fazla bağımsız kategoriden elde edilen kategorik verilerin karşılaştırılmasında kullanılan Ki Kare testi hakkında bilgi sahibi olması ve bu testi uygulayabilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Kategorik veri ikiden fazla grubu açıklayabilmek
- Ki Kare testi varsayımlarını açıklayabilmek
- Ki Kare testinin mantığını açıklayabilmek
- Fisher exact testinin uygulama alanını açıklayabilmek
- SPSS’te Ki Kare testi yapabilmek
- SPSS’te Ki Kare çıktısını yorumlayabilmek

Örneklelimizdeki bireylerin medeni durumlarının (evli, bekâr, boşanmış = ikiden fazla kategori) antidiyabetik ilaç tercihleri (oral antidiyabetik, insülin = iki grup) ile ilişkili olup olmadığını incelemek istiyoruz.

Bağımsız değişkenimiz ikiden fazla kategoriden oluştuğundan bu örnekte de örnek 8 ve 9’da olduğu gibi Ki kare testi uygulanmalıdır.

Kategori sayısı ikiden fazla olduğunda çapraz tablonun göz sayısı 4’ten fazla olacağından grup sayısının kaç adet olduğuna bakılmaksızın Ki Kare analizi yapılır.

Bakınız sayfa 298

Bölüm 38: Kategorik Veri İkiden Fazla Kategori – 1 Grup

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun bir grupta ikiden fazla kategorisi olan veriler için Ki Kare testi hakkında bilgi sahibi olması ve bu testi uygulayabilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Kategorik veri 1 grup – ikiden fazla kategoriyi açıklayabilmek
- Genetik Ki Kare testinin kullanım alanını açıklayabilmek
- SPSS’te ki Kare testi yapabilmek
- SPSS’te ki Kare çıktısını yorumlayabilmek

Örnekleminizdeki bireylerin kullandıkları tedaviler arasında anlamlı bir fark olup olmadığını incelemek istiyoruz. “Aldığı tedavi” (therapy) 6 kategorisi olan kategorik bir değişkendir.

İkiden fazla kategorisi olan tek bir değişkende kategoriler arasındaki farklılığın anlamlı olup olmadığı sağlık uygulamalarında sık karşılaşılabileceğimiz bir sorudur. Sağlık hizmetlerine başvuru yüzdesi açısından haftanın günleri arasında fark olup olmadığı, bir klinikteki ikiden fazla hekimin tercih edilme yüzdeleri arasında fark olup olmadığı gibi ek örnekler verilebilir. Burada yapılacak testin adı da Ki Kare olmakla birlikte, önceki konularda bahsedilen çapraz tablolarda (crosstabs) uygulanan Ki Kare testinden farklılık göstermektedir. SPSS’teki menüleri de farklıdır.

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

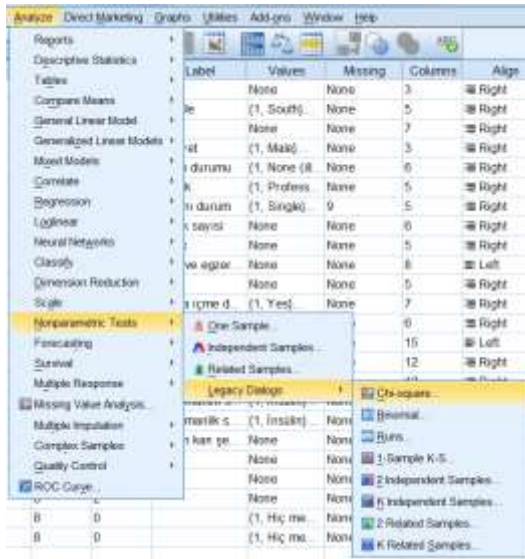
H_0 : Hastalarımızın diyabet için kullandıkları tedavi çeşitlerinin yüzdeleri arasında anlamlı bir fark yoktur.

H_1 : Hastalarımızın diyabet için kullandıkları tedavi çeşitlerinin yüzdeleri arasında anlamlı bir fark vardır .

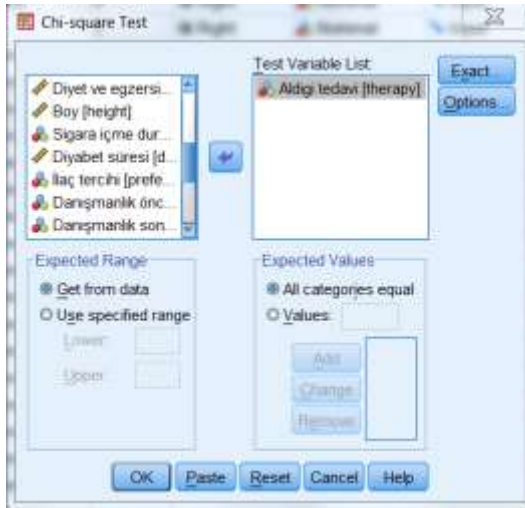
2. Verilerin toplanması:

Ki Kare testini SPSS ile yapacak olursak:

Analyze>Nonparametric Tests>Legacy dialogues>Chi-square>



[“therapy” değişkenini “Test Variable List” alanına geçirelim]>OK



Aldığı tedavi			
	Observed N	Expected N	Residual
No pharmacologic therapy	21	71,7	-50,7
Sulphonylurea only	233	71,7	161,3
Biguanides only	5	71,7	-66,7
Sulphonylurea + Biguanides	125	71,7	53,3
Insulin	43	71,7	-28,7
Insulin + oral antidiabetics	3	71,7	-68,7
Total	430		

Test Statistics	
	Aldığı tedavi
Chi-square	577,972 ^a
Df	5
Asymp. Sig.	,000

a. 0 cells (,0%) have expected frequencies less than 5. The minimum expected cell frequency is 71,7.

Elde edilen iki çıktıdan birincisinde kullanılan çeşitli tedavilerin tanımlayıcı istatistiği, diğerinde ise Ki Kare ve p değerleri sunulmaktadır.

P değerinin ve sonuçların yorumlanması

P değeri 0,005'ten küçük olduğu için H_0 hipotezini reddederiz. Dolayısıyla diyabetik bireylerde kullanılan tedavi çeşitlerinin yüzdeleri arasında anlamlı bir fark vardır. Hastalar daha çok tek başına sulfonilüre kullanmaktadır.

Bölüm 39: İki Nümerik Bağımlı Değişken (Bland-Altman Analizi)

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun iki bağımlı nümerik değişken arasındaki ilişkiyi incelemek için Bland-Altman analizi hakkında bilgi sahibi olması ve bu analizleri SPSS ile yapabilmesi amaçlanmıştır

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Bland-Altman analizinin ne zaman yapılacağını açıklayabilmek
- SPSS ile Bland-Altman grafiği yapabilmek
- Bland-Altman çıktısını yorumlayabilmek

Hastalarımızın glomeruler filtrasyon hızlarını (GFR) [nümerik değişken] iki farklı yöntemle (A ve B yöntemi) ölçtük. Bu yöntemlerin birbirleri ile ne kadar uyumlu olduklarını araştırmak istiyoruz. Klasik olarak kliniğimizde yıllardır A yöntemi kullanılmaktadır. A yöntemi yerine yeni geliştirilen ve daha hızlı ve ekonomik olan B yöntemini kullanıp kullanamayacağımızı araştırmak istiyoruz. H_0 : GFR hesaplamada A yöntemi ile B yöntemi arasında fark yoktur.

Bland-Altman analizinin kullanım yeri

Bland-Altman analizi, iki nümerik değişken arasındaki uyumu göstermek üzere geliştirilmiştir. (Hatırlanacağı üzere, Kappa uyum analizinde ise kategorik değişkenler arasındaki uyum araştırılmaktadır). Temelde bir saçılma grafiği (Bakınız S. 67) olan Bland-Altman analizinde aynı kişilerde iki farklı yöntemle nümerik ölçümler yapılmaktadır (bu nedenle bağımlı iki grup). Ölçülen iki nümerik değişken arasındaki fark X ekseninde, iki nümerik değişkenin ortalaması ise Y ekseninde gösterilir.

Bland-Altman grafiğinde ayrıca ortalama farkları gösteren bir yatay çizgi ve ortalama + 1,96 standart sapma ve ortalama – 1,96 standart sapma seviyelerini gösteren iki yatay çizgi olmak üzere üç adet yatay çizgi bulunur.

MedCalc (Bakınız S. 412) gibi bazı istatistik paket programları Bland-Altman analizini direkt olarak yapmaktadırlar. SPSS’te bu analiz bulunmamakla birlikte, saçılma grafikleri ve tek örnekleme t testini kullanarak Bland-Altman analizinin nasıl yapılacağı gösterilecektir.

SPSS ile Bland-Altman analizinin yapılması

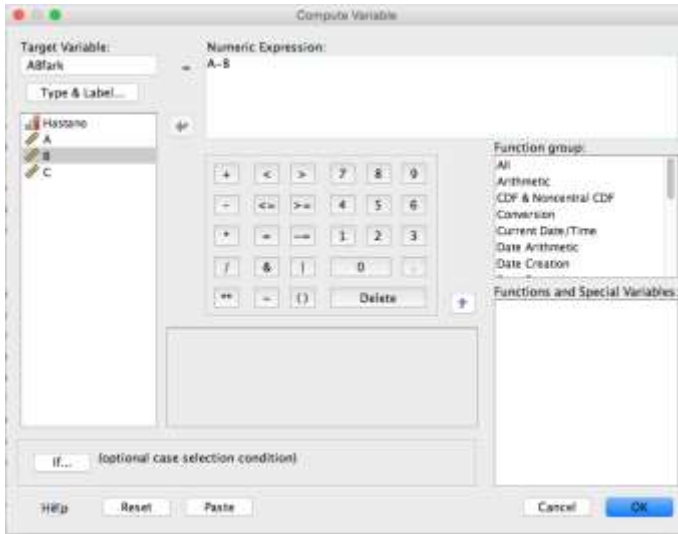
<https://drive.google.com/file/d/0B0ieO-Fy3z-5YTc5UUNibUZmajg/view?usp=sharing>

adresinden BlandAltman.sav SPSS veri setini indirip açınız.

A ile B ölçümleri arasındaki uyuma bakalım:

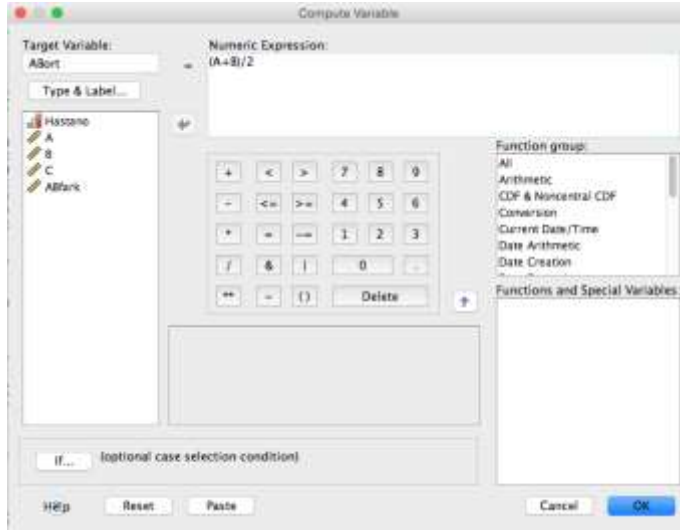
Öncelikle A-B ve (A+B)/2 değişkenlerini oluşturmamız gerekecek.

Transform>Compute Variable [Taget Variable alanına yeni bir değişken ismi (Örn. ABfark) yazalım. Numeric Expression kısmına A-B yazalım]



>OK

Transform>Compute Variable [Taget Variable alanına yeni bir değişken ismi (Örn. ABort) yazalım. Numeric Expression kısmına (A+B)/2 yazalım]



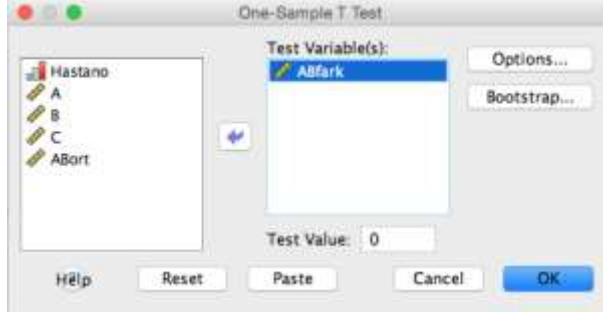
>OK

Şimdi veri setimizde iki yeni değişken oluştu:

Hasta...	A	B	C	ABfark	ABort
1	107.61	126.17	108.01	-19	117
2	91.92	87.62	76.15	4	90
3	93.77	147.21	103.37	-53	120
4	81.02	90.64	95.71	-10	86
5	46.21	77.70	84.11	-31	62
6	89.57	125.97	79.71	-36	108
7	100.76	130.59	105.48	-30	116
8	81.71	91.30	52.02	-10	87
9	97.29	114.99	113.32	-18	106
1	95.10	97.41	124.95	-2	96
11	93.28	89.61	103.88	4	91
12	95.05	104.96	109.37	-10	100
13	81.73	118.78	92.00	-37	100
14	112.80	111.99	81.65	1	112
15	81.10	98.90	105.15	-18	90
16	79.88	100.66	68.10	-21	90
17	80.45	118.55	80.94	-38	100
18	82.85	104.86	101.30	-22	94
19	90.97	106.88	79.55	-16	99
20	62.33	112.28	118.67	-50	87
21	89.56	107.13	73.69	-18	98

Öncelikle ABfark değişkeninin sıfırdan farklı olup olmadığına tek örneklemde t testi ile (Bakınız S. 196) bakalım.

Analyze > Compare Means > One-Sample T Test [ABfark değişkenini “Test Variable(s)” analına geçirelim] > OK



Aşağıdaki çıktıları elde ederiz:

One-Sample Statistics				
	N	Mean	Std. Deviation	Std. Error Mean
ABfark	30	-16.60	18.260	3.334

Ortalama fark = -16,6; standart sapması ise 18,26 olarak hesaplanmıştır.

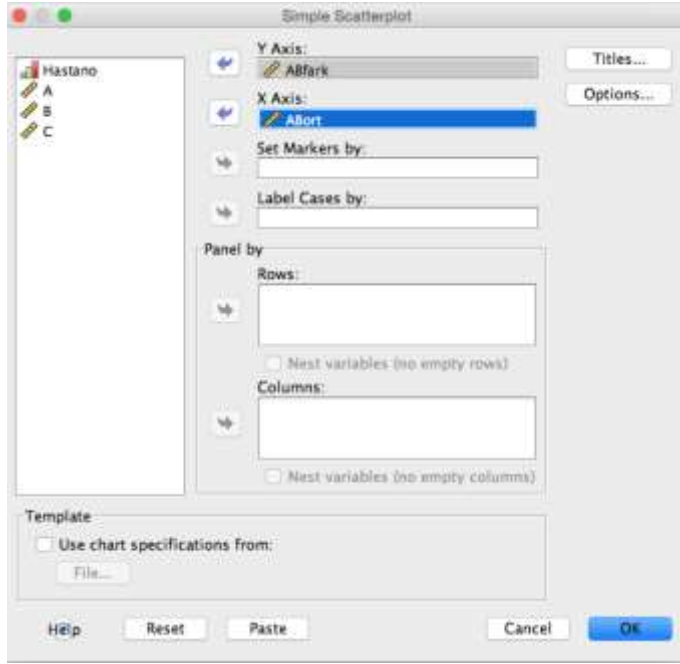
One-Sample Test					
Test Value = 0					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference Lower Upper
ABfark	-4.980	29	.000	-16.602	-23.42 -9.78

p değeri $< 0,001$ olduğundan iki yöntemin ölçümleri arasındaki fark sıfırdan farklıdır (H_0 hipotezi reddedilmelidir). Dolayısıyla GFR ölçümünde A yöntemi yerine B yöntemini kullanmak doğru olmayacaktır.

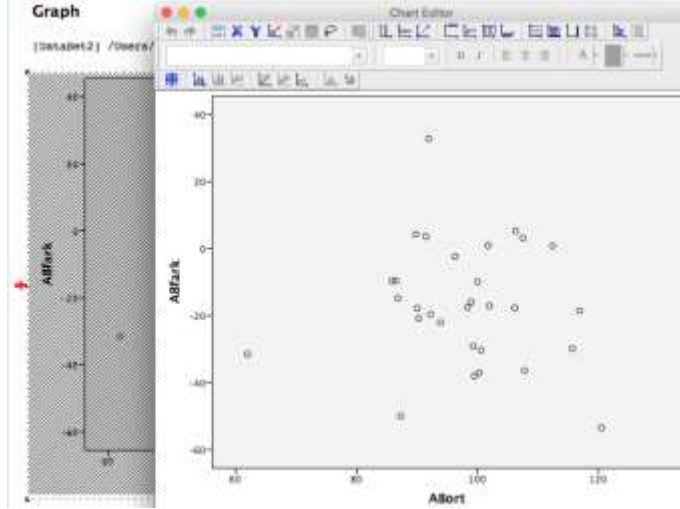
Şimdi Bland-Altman grafiği ile uyum durumunu inceleyim:

ABfark ve ABort değişkenlerini kullanarak saçılma grafiği çizelim:

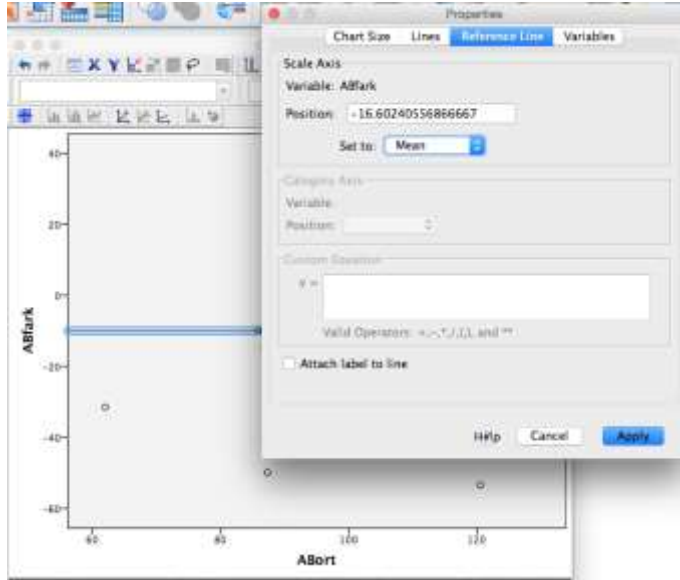
Graphs>Legacy Dialogues>Scatter/Dot>Simple Scatter>Define [ABfark değişkenini Y Axis, ABort değişkenini X Axis alanına geçirelim]>OK



[Oluşan grafiği çift tıklayarak açalım]

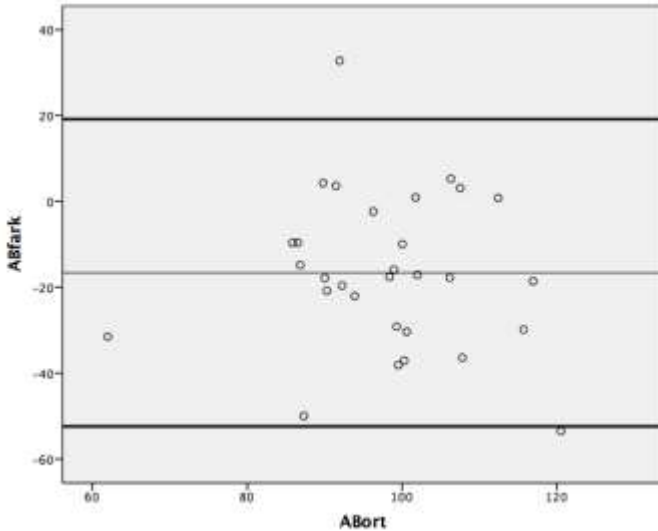


Options>Y Axis Reference Line [Set to: Mean seçelim]>Apply



>Close

Örneğimizde Ortalama fark = -16,6; standart sapması ise 18,26 olarak hesaplanmıştı. Grafiğimizi tekrar çift tıklayarak (Ortalama + $1,96 \cdot SS$) = **19,18** ve (Ortalama - $1,96 \cdot SS$) = **-52,38** olacak şekilde iki yatay çizgi daha çizeceğiz.



Bland-Altman grafiğini şu şekilde yorumlayabiliriz: farkların ortalaması (ince çizgi) sıfıra yakın değildir. Aynı zamanda A ve B ölçümlerinin

farklı ortalamalarına karşılık gelen A-B farklarının dağılımı da varyasyon göstermektedir (kalıç çizgiler birbirinden uzaktır). Bu nedenle klasik olarak kullanageldiğimiz A testi yerine B testinin kullanılmaya başlanması doğru olmayacaktır.

Alıştırma 12

Hastalarımızın glomeruler filtrasyon hızlarını (GFR) [nümerik değişken] iki farklı yöntemle (A ve C yöntemi) ölçtük. Bu yöntemlerin birbirleri ile ne kadar uyumlu olduklarını araştırmak istiyoruz. Klasik olarak kliniğimizde yıllardır A yöntemi kullanılmaktadır. A yöntemi yerine yeni geliştirilen ve daha hızlı ve ekonomik olan C yöntemini kullanıp kullanamayacağımızı araştırmak istiyoruz. H_0 : GFR hesaplamada A yöntemi ile C yöntemi arasında fark yoktur.

<https://drive.google.com/file/d/0B0ieO-Fy3z-5YTc5UUNibUZmaj/view?usp=sharing> adresinden BlandAltman.sav SPSS veri setini kullanarak yukarıdaki sıfır hipotezini test ediniz.

Çözüm için bakınız s. 477

Bölüm 40: İki Nümerik Bağımsız Değişken (Korelasyon)

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun iki nümerik değişken arasındaki lineer ilişkiyi incelemek için Pearson ve Spearman korelasyon hesapları hakkında bilgi sahibi olması ve bu analizleri SPSS ile yapabilmesi amaçlanmıştır

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Korelasyonun ne zaman yapılacağını açıklayabilmek
- Pearson ve Spearman korelasyonu için varsayımları sayabilmek
- SPSS’te Pearson korelasyonu yapabilmek
- SPSS’te Spearman korelasyonu yapabilmek
- Korelasyon katsayısı r ’yi yorumlayabilmek
- r^2 ’yi açıklayabilmek
- Serpiştirme grafiği (scatter plot) çizebilmek

Diyabetli hastalarımızda diyabet süresi (x) uzadıkça HbA1c düzeylerinin (y) de arttığından şüpheleniyoruz. Diyabet süresi ile HbA1c arasında doğrusal bir ilişki olup olmadığına bakmak istiyoruz.

Korelasyon analizi, her ikisi de nümerik olan iki değişken arasındaki ilişkinin derecesini ölçer. Değişkenleri x ve y olarak tanımlayıp bir serpiştirme grafiği (scatter diagram) üzerinde gösterebiliriz. Böylece iki değişken arasındaki ilişkinin yönünü ve derecesini anlamaya çalışırız.

Pearson korelasyon katsayısı

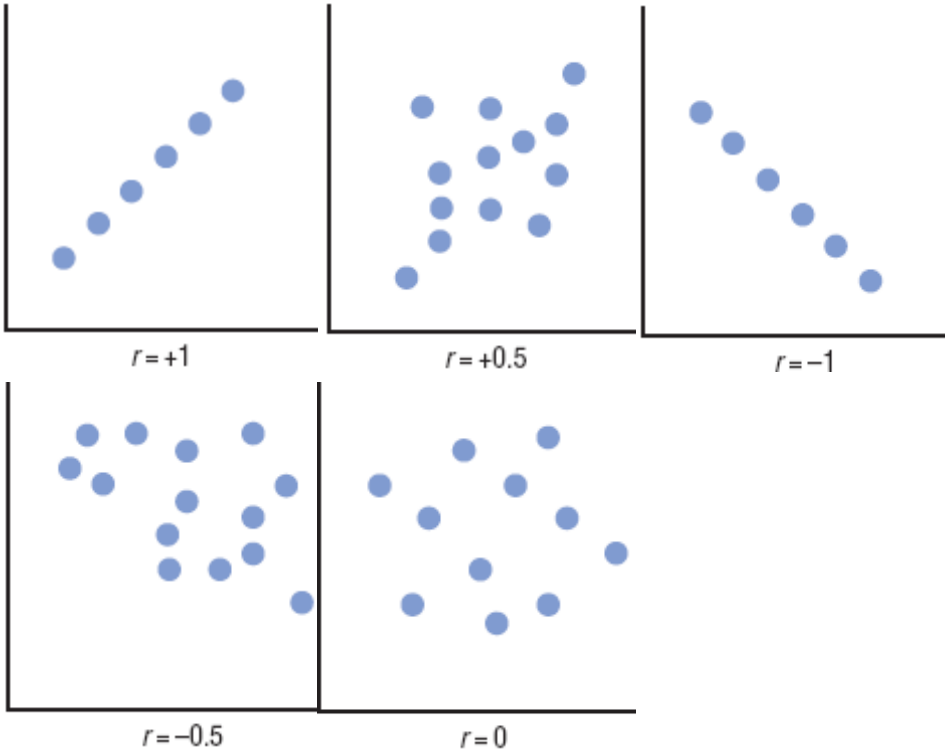
Grafikteki noktaların düz bir hat üzerinde dağılmaları halinde değişkenlerin arasında doğrusal bir ilişki olduğundan bahsederiz. **Pearson Product Moment Korelasyon Katsayısı**’nı (basitçe **korelasyon katsayısı** denir) hesaplayarak doğrusal ilişkiyi belirleyen çizgiyi bulabiliriz. Toplumdaki değeri “rho” ile gösterilen korelasyon katsayısının örneklemimizdeki değeri r ile gösterilir ve aşağıdaki formülle hesaplanır:

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

Özellikleri

- r ’nin değeri -1 ile +1 arasında değişir
- r ’nin işareti değişkenlerden biri artarken ötekinin de arttığını (pozitif) veya azaldığını (negatif) gösterir.

- r 'nin büyüklüğü serpiştirme grafiğinde noktaların düz bir hatta ne kadar yakın olduğunu gösterir. +1 veya -1 olması arada mükemmel bir ilişki olduğunu gösterir (pratikte genelde mümkün değil). 0 olması ise arada doğrusal bir ilişki olmadığını gösterir. r 'nin mutlak değeri arttıkça iki değişken arasındaki ilişki de daha kuvvetli demektir.
- r 'nin herhangi bir birimi yoktur.
- r 'nin değeri sadece örneklemdeki değişkenlerin en az ve en çok değerleri arasında geçerlidir. Örneklemdekinden daha küçük veya büyük değerler için bir yorum yapılamaz.
- x ve y 'nin yerlerinin değiştirilmesi r 'nin değerini değiştirmez.
- x ve y arasında bir korelasyonun olması, bir sebep-sonuç ilişkisi olduğu anlamına gelmez.
- r^2 , y 'deki değişimin x ile olan doğrusal ilişkisiyle açıklanabilen kısmını gösterir.

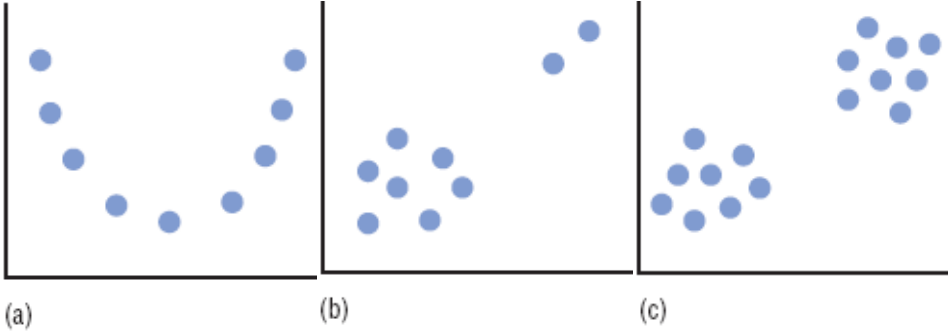


Şekil: r 'nin değişik değerlerini gösteren grafikler.

***r* ne zaman hesaplanmamalıdır?**

Aşağıdaki durumlarda *r*'yi hesaplamak yanıltıcı olabilir:

- iki değişken arasında doğrusal olmayan bir ilişki olması durumunda
- Verinin her birey hakkında birden fazla ölçüm içermesi durumunda.
- Uç değerlerin bulunması durumunda.
- Verinin alt gruplardan oluşması durumunda.



Şekil: korelasyon katsayısının hesaplanmasının uygun olmadığı durumlar. (a) doğrusal olmayan bir ilişki bulunması; (b) uç değerlerin bulunması; (c) verinin alt gruplardan oluşması.

Pearson korelasyon katsayısı için hipotez testi

Konu başındaki senaryomuzu diyabet.sav verisetine göre uygulayacak olursak:

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

H_0 : Diyabet süresi ile HbA1c arasında bir korelasyon yoktur (korelasyon katsayısı sıfırdır)

H_1 : Diyabet süresi ile HbA1c arasında bir korelasyon vardır (korelasyon katsayısı sıfırdan farklıdır)

2. Verilerin toplanması:

Veriler diyabet.sav verisetinde “duration” ve “hba1c” değişkenleri altında kaydedildi.

3. İlgili sıfır hipotezi için test istatistiğinin hesaplanması:

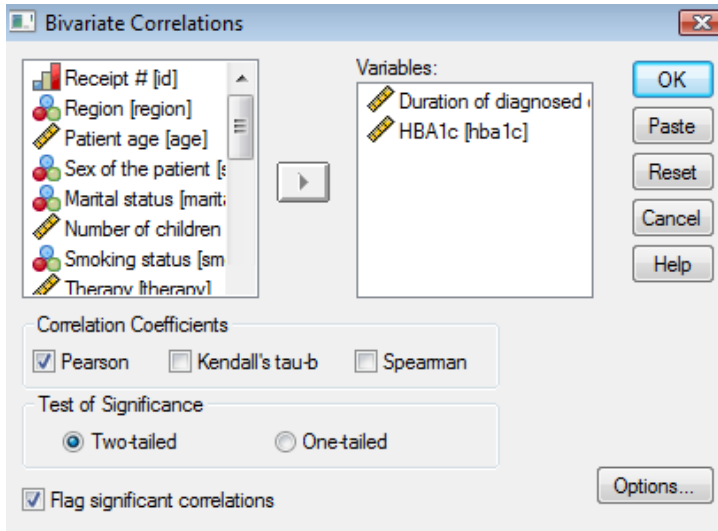
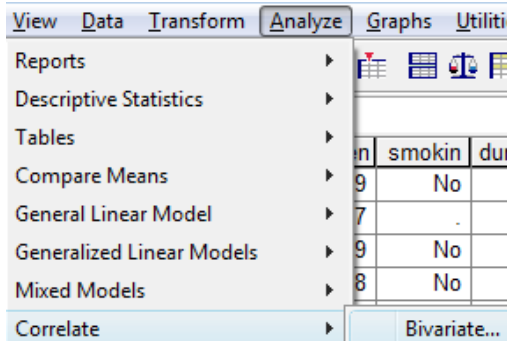
$n \leq 150$ ise r test istatistiği olarak kullanılır.

$n > 150$ ise T hesaplanır:

$$T = \sqrt{\frac{(n-2)}{(1-r^2)}}$$

SPSS’te:

Analyze>Correlate>Bivariate>[“duration” ve “hba1c” değişkenlerini “Variables” alanına geçirelim. “Pearson” kutucuğu işaretli]>ok



Aşağıdaki çıktıyı elde ederiz:

Correlations

		Duration of diabetes	HBA1c
Duration of diabetes	Pearson Correlation	1	,036
	Sig. (2-tailed)		,677
	N	165	133
HBA1c	Pearson Correlation	,036	1
	Sig. (2-tailed)	,677	
	N	133	162

4. Test istatistiğinden elde edilen değerin bilinen bir olasılık dağılımı ile karşılaştırılması:

$n \leq 150$ ise sayfa 462'ten r bakılır.

$n > 150$ ise 459'den T bakılır.

SPSS'te Pearson korelasyon katsayısı 0,036; p ise 0,677 olarak verilmiştir.

5. P değerinin ve sonuçların yorumlanması

P değeri 0,05'ten büyük olduğu için H_0 hipotezini kabul ederiz. Dolayısıyla diyabet süresi ile hba1c değerleri arasında anlamlı bir doğrusal korelasyon yoktur. Yani diyabet süresi arttıkça HbA1c artmamakta veya azalmamaktadır.

H_0 kabul edildiği için önemi kalmamaktadır. Ancak, r^2 'yi hesaplayacak olsaydık, şöyle bir yorum yapabilirdik: $r^2 = 0,001296$. Dolayısıyla HbA1c'deki değişimin %0,1296'sı diyabet süresiyle açıklanabilir olacaktı.

Spearman korelasyon katsayısı

Aşağıdaki durumlardan birinin söz konusu olması halinde Spearman korelasyon katsayısını (r_s) hesaplarız:

- Değişkenlerden birisinin ordinal olması
- Her iki değişkenin de normal dağılmaması
- Örneklem sayısının küçük olması
- x ve y arasında doğrusal bir ilişkinin olmaması

Farkları

- r_s x ve y arasında bir ilişki olduğunu gösterir. Bu ilişkinin doğrusal olması gerekmez
- sıfır hipotezi test edilirken örneklem sayısı 10 veya altında ise sayfa 463'e bakılır
- r_s^2 hesaplanamaz.

Bölüm 41: Lineer Regresyon

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun çeşitli bağımsız değişkenleri kullanarak bir nümerik değişkenin değerini tahmin etmek için lineer regresyon analizi hakkında bilgi sahibi olması ve bu analiz SPSS ile yapabilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Lineer regresyon analizinin endikasyonlarını açıklayabilmek
- Lineer regresyon modeline konacak değişkenlerin nasıl seçileceğini açıklayabilmek
- “Dummy değişken” terimini açıklayabilmek
- SPSS’te Lineer regresyon analizi yapabilmek
- SPSS’te Lineer regresyon çıktısını yorumlayabilmek

Kişilerin vücut ağırlıklarını bazı bağımsız değişkenleri kullanarak tahmin etmek istiyoruz. Araştırmamızda kişilerin yaşları, boyları, cinsiyetleri gibi değişkenler hakkında veri toplanmış. Bağımsız değişkenlerden faydalanarak kişilerin vücut ağırlıklarını tahmin edebilir miyiz?

Regresyon analizleri sağlık hizmetlerinde sıklıkla kullanılmaktadır. Tahmin etmek istediğimiz değişkenin nümerik olması durumunda Lineer (İng.=doğrusal) regresyon (İng.=gerileme) analizi yapmalıyız. Örneğimizde vücut ağırlığını (nümerik değişken) tahmin etmek istiyoruz. Tahmin ettirici (bağımsız) değişkenlerin de nümerik olması gerekir.

Varsayımlar

Lineer regresyon analizini yapabilmemiz için gerekli şartlar şunlardır:

- **Lineer ilişki:** Değişkenler arasındaki ilişki lineer olmalıdır. İlişkinin olmaması veya üstsel bir ilişkinin olması durumunda lineer regresyon analizi yapamayız. Lineer ilişkiyi görmek için saçılma grafiği veya korelasyon analizinden faydalanabiliriz.
- **Normal dağılım:** Modele koyduğumuz değişkenlerin normal dağılımı gerekir; lineer regresyon analizi uç değerlerden kolay etkilendir. Normalliği incelemek için histogram grafiklerini kullanabiliriz.
- **Multi kolinearite** olmamalı: Yani bağımsız değişkenler birbiriyle bağımlı olmamalıdır. Örneğin, tansiyonu tahmin etmek için kan şekeri ve hemoglobin A1c değişkenlerini kullanacak olsak, HbA1c direkt olarak kan şekerinden etkilendiğinden modele konması uygun değildir.

- **Otokorelasyon** olmamalı: Bağımlı değişkenin bir ölçümü ötekileri etkilememeli. Örneğin bir firmanın borsadaki değeri bağımlı değişken olsa, firmanın önceki fiyatı sonraki değerini etkileyeceğinden otokorelasyondan söz edilebilir. Otokorelasyon da saçılma grafiği veya Durbin-Watson (d) testi ile incelenebilir. Kural olarak 1,5-2,5 arasındaki d değerleri otokorelasyon olmadığını gösterir.
- Varyansların homojenliği (**Homoscedasticity**): Saçılma grafiği bu amaçla da kullanılabilir. Bağımlı değişkenin küçük değerleri için bağımsız değişkendeki varyasyon, bağımlı değişkenin büyük değerlerine göre farklı ise bu durumda heteroscedastisiteden söz edilir. Örneğin, gelir düzeyi ile lüks ürünlere yapılan harcamaları tahmin ettiğimizi düşünelim. Düşük gelir düzeylerindeki kişilerin yaptıkları harcamaların varyasyonu nispeten aynı iken yüksek gelir düzeylerinde varyasyon artmaktadır, yani zenginlerin bazıları lüks harcama yaparken bazıları yapmamaktadır. Bu durum heteroscedastisite olarak bilinir.

Modelin oluşturulması

Modele dahil edeceğimiz değişkenlerin seçimini sayfa 352’de lojistik regresyon bahsinde anlatıldığı şekilde yapmak gerekir.

SPSS ile uygulama

Konu başında verilen senaryomuzu SPSS ile analiz edelim:

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

H_0 : Oluşturulan modeldeki değişkenlerin kişilerin ağırlıklarını tahmin etmede etkisi yoktur.

H_1 : Oluşturulan modeldeki değişkenler kişilerin ağırlıklarını tahmin ettirebilir.

2. Verilerin toplanması:

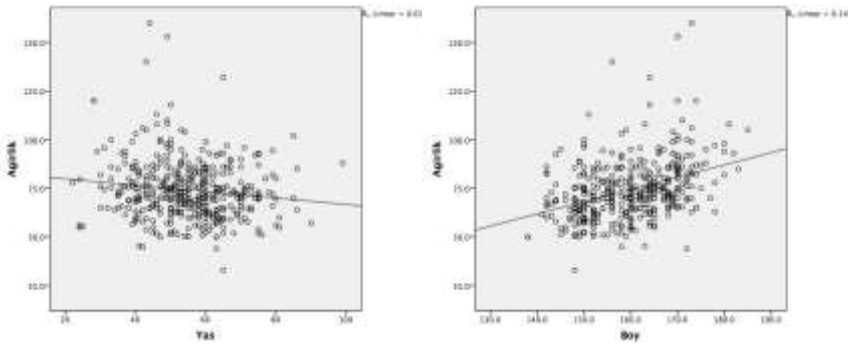
Veriler diyabet.sav verisetinde kaydedildi. Bağımlı değişkenimiz “weight” nümerik olarak kodlanmıştır. Bağımsız değişkenler olarak incelenecek “age_year” ve “height” değişkenleri de nümeriktir.

Korelasyon analizinde ağırlık ile yaş ve boy arasında anlamlı bir korelasyon olduğu ama yaş ile boy arasında (bağımsız değişkenler) korelasyon olmadığı görülmektedir:

Correlations				
		Yas	Ağırlık	Boy
Yas	Pearson Correlation	1	-.135	.001
	Sig. (2-tailed)		.005	.976
	N	429	423	405
Ağırlık	Pearson Correlation	-.135**	1	.377**
	Sig. (2-tailed)	.005		.000
	N	423	424	404
Boy	Pearson Correlation	.001	.377**	1
	Sig. (2-tailed)	.976	.000	
	N	405	404	406

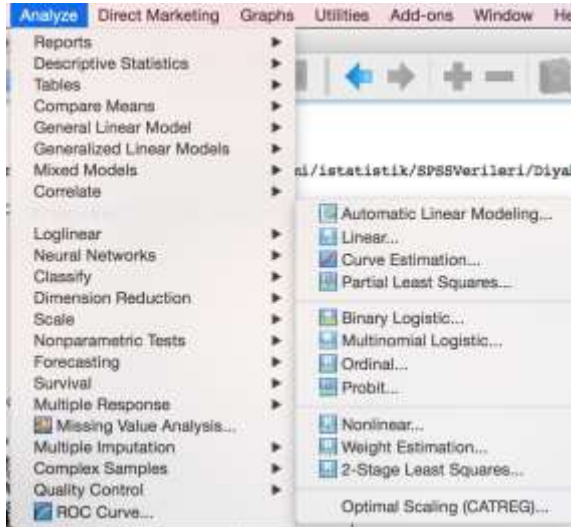
** . Correlation is significant at the 0.01 level (2-tailed).

Saçılma grafiklerinde de ilişkinin – zayıf da olsa – linear olduğu görülmektedir:

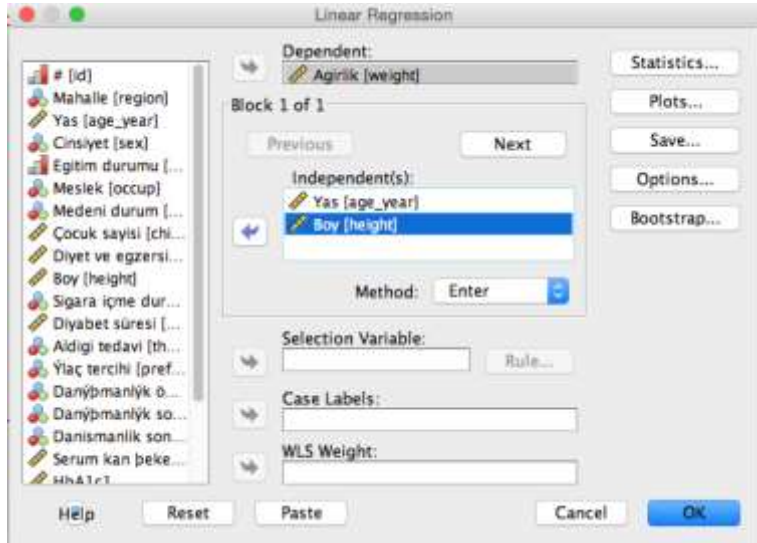


3. İlgili sıfır hipotezi için test istatistiğinin hesaplanması:

Analyze>Regression>Linear>



[“weight” değişkenini “Dependent” alanına, “age_year” ve “height” değişkenlerini “Independent(s)” alanına geçirelim.]>



Statistics>[Confidence intervals ve Durbin-Watson kutucuklarını işaretleyelim]>Continue>Save>[Unstandardized kutucuğunu işaretleyelim. Böylece modelimizin tahmin değerleri veri setimizde ayrı bir sütun olarak hesaplanıp yazılacaktır.]>Continue>OK

Aşağıdaki çıktıları elde ederiz:

Variables Entered/Removed ^a			
Model	Variables Entered	Variables Removed	Method
1	Boy, Yaş ^b	.	Enter

a. Dependent Variable: Ağırlık
b. All requested variables entered.

Bu tablo, analize dahil edilen değişkenleri göstermektedir.

Model Summary ^a					
Model	R	R Square	Adjusted R Square	Sd. Error of the Estimate	Durbin-Watson
1	.404 ^a	.163	.159	14,6509	1.919

a. Predictors: (Constant), Boy, Yaş
b. Dependent Variable: Ağırlık

Bu tablo modelimizin özetini göstermektedir.

R kare

R kare değeri 16,3 olarak hesaplanmıştır. Bu durumda “Modelimizin ağırlığa etki eden faktörlerin %16,3’ünü açıklayabildiğini”

söyleyebiliriz. Durbin-Watson değerinin 1,9 olması ilgili varsayımın karşılandığını göstermektedir.

4. Test istatistiğinden elde edilen değer bilinen bir olasılık dağılımı ile karşılaştırılması:

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	15375.411	2	7687.705	38.940	.000 ^b
	Residual	78970.685	400	197.427		
	Total	94346.095	402			

a. Dependent Variable: Ağırlık
b. Predictors: (Constant), Boy, Yaş

F dağılımı kullanılarak yapılan çıkarımda ANOVA tablosunda *p* değerinin <0,001 olduğu görülmektedir. Bu durumda modelimizin bütün olarak bakıldığında anlamlı olduğunu söyleyebiliriz.

5. *P* değerinin ve sonuçların yorumlanması

Coefficients ^a								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-16.666	12.570		-1.326	.186	-41.377	8.045
	Yaş	-.177	.057	-.143	-3.128	.002	-.289	-.066
	Boy	.630	.076	.378	8.269	.000	.480	.779

a. Dependent Variable: Ağırlık

Bu tabloda değişkenlerin tek tek anlamlılık düzeyleri verilmiştir. *p* değerlerine (Sig.) bakıldığında hem yaş ($p=0,002$), hem de boy ($p<0,001$) değişkeninin **birbirinden bağımsız olarak** ağırlığa etki ettikleri söylenebilir. Bu bilgiler kullanılarak bir lineer regresyon eşitliği yazılabilir.

Lineer regresyon eşitliği

Lineer regresyon eşitliğini aşağıdaki gibi oluşturabiliriz.

Unutmamalıyız ki, modelimiz ağırlık değişkeninin yaklaşık %16'sını tahmin edebilmektedir.

$$\text{Ağırlık} = -16,66 - (0,177 * \text{yaş}) + (0,630 * \text{boy})$$

SPSS verisetimizi incelediğimizde PRE_1 adı altında yeni bir değişken tanımlanıp modelin tahmin değerlerinin bu alana yazıldığı görülmektedir:

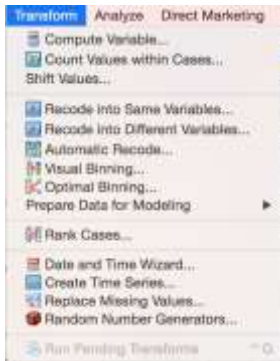
weight	PRE_1
113.0	70.25219
84.0	82.39375
102.0	75.29423
118.0	77.72818
88.0	76.76048
75.0	72.90209
71.0	75.93634
75.0	79.50306
86.0	73.17680
80.0	.
99.0	81.40909
75.0	79.02148

Bu tablodan da anlaşılacağı üzere modelimizin tahmin gücü zayıftır. Bazı bireylerin durumları kısmen tahmin edilebilse de birçok değişken iyi tahmin edilememiştir.

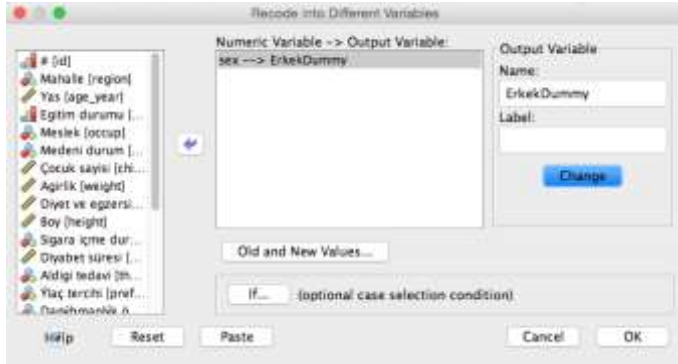
Kategorik değişkenlerin modele konulması

Daha önce bahsedildiği üzere, lineer regresyon analizinde modelimize sadece nümerik değişkenleri dahil edebiliriz. Bununla birlikte, kategorik değişkenlerin modele eklenme zorunlulu hissedildiğinde değişkenler ikili olarak kodlanmak suretiyle modele eklenebilirler. Oluşturulan bu yeni değişkenlere **Dummy değişken** denir.

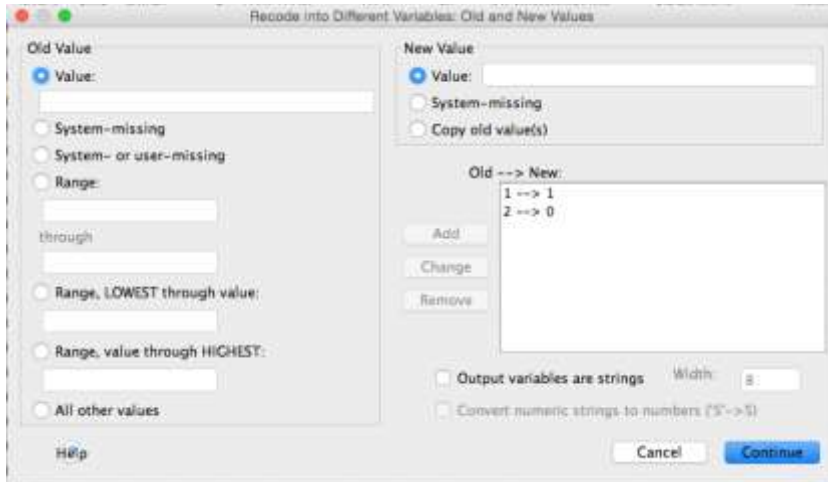
Örneğin, ağırlığı tahmin etmede cinsiyet değişkenini de modele eklemek istesek cinsiyeti erkek=1; kadın=2 yerine erkek (1=evet; 0=hayır), kadın (1=evet; 0=hayır) şeklinde kodlamamız gerekir. Yeniden kodlamayı şu şekilde yapabiliriz: *Transform>Recode into Different Variables*



[“sex” değişkenini Numeric variable -> Output Variable alanına geçirelim]
>[“Name” kısmına “ErkekDummy” yazalım]>Change

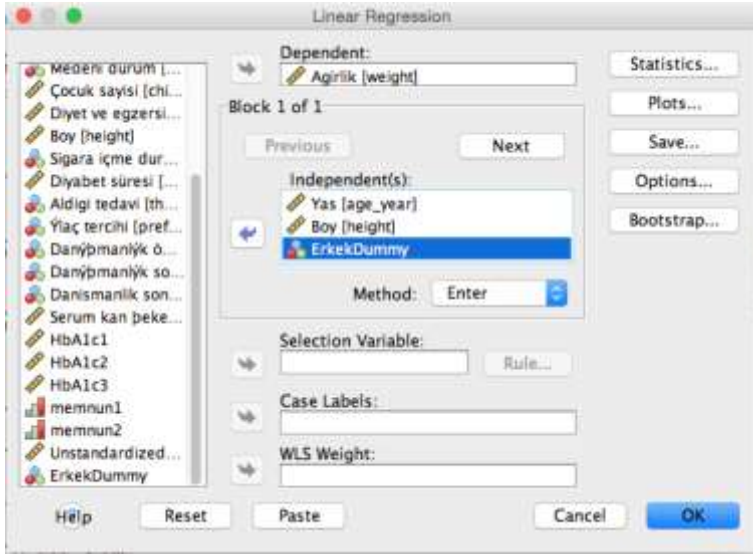


>Old and New Values>[Old value kısmına 1, New Value kısmına 1 yazalım]>Add>>[Old value kısmına 2, New Value kısmına 0 yazalım]>Add



>Continue>OK

Şimdi yeni değişkenimizi de ekleyerek analizimizi tekrar edebiliriz:



SPSS çıktısına baktığımızda cinsiyetin anlamlı olmadığını görüyoruz:

Coefficients ^a								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-29.396	17.482		-1.681	.093	-63.765	4.973
	Yas	-.166	.058	-.134	-2.864	.004	-.279	-.052
	Boy	.712	.110	.428	6.497	.000	.497	.928
	ErkekDummy	-2.137	2.040	-.070	-1.048	.295	-6.147	1.873

a. Dependent Variable: Agirlik

a. Dependent Variable: Ağırlık

Bu noktada, modele yeni bir değişken eklemek suretiyle diğer değişkenlerin katsayılarının ve lineer regresyon formülünün de değiştiğine dikkat çekmek gerekir.

Aıştırma 13

Adli tıp uzmanları için ölen kişinin ne kadar zaman önce öldüğü önem arz eder. Bir araştırmada farklı zamanlarda ölen farelerin ölüm zamanlarının kanda yapılan bazı analizlerle tahmin edilmesi amaçlanıyor (Verileri sağladığı için Doç. Dr. Talip ŞENER'e teşekkür ederiz). SPSS verisetini şu bağlantıdan indirebilirsiniz: <https://drive.google.com/file/d/0B0ieO-Fy3z-5M1BldG9XLWt6Tlk>

İlgili veri setini kullanarak şu sorulara cevap veriniz:

1. Araştırmanın bağımlı değişkeni nedir?
2. Araştırmanın bağımsız değişkenleri hangileridir?
3. Bağımsız değişkenleri kullanarak bağımlı değişkeni tahmin etmek için bir lineer regresyon analizi yapın.
4. Regresyon analizinin r kare değeri kaçtır?
5. Regresyon formülünü yazın.

Çözüm için bakınız s. 480.

Bölüm 42: Ortalamaya Doğru Gerileme

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun “ortalamaya doğru gerileme” kavramı hakkında bilgi sahibi olmaları ve günlük yaşamla bağlantısını kurabilmeleri amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- “ortalamaya doğru gerileme” kavramını açıklayabilmek
- Günlük yaşamdan “ortalamaya doğru gerileme” örnekleri verebilmek

1822-1911 yıllarında yaşayan Francis Galton, toplum ortalamasından çok uzun babaların çocuklarının da uzun olduğunu, ama toplum ortalamasına yaklaştığını, topluma göre çok kısa babaların çocuklarının ise kendilerinden uzun ve yine toplum ortalamasına yakın olduklarını tespit etmiş.

Daha sonra yapılan araştırmalarda bu prensibin hayatın çeşitli alanlarına da uygulanabileceği saptanmıştır. Örneğin, borsanın beklenenin üzerinde aşırı yükselmesi durumunda daha sonraki seansta ortalamaya gerileyeceğini tahmin edebiliriz.

Bir basketbol oyuncusunun ortalama bir maçta 10 sayı yaptığını varsayalım. Oyuncu bir maçta olağüstü başarı gösterip 16 puan yapsa, sonraki maçta büyük olasılıkla daha düşük yapacağını ama yine de ortalamaya yakın olacağını tahmin edebiliriz. Basketbol koçunun vazifesi, oyuncunun ortalamasını yükseltmeye çalışmak olmalıdır.

Ortalamaya doğru gerileme, akademisyenler için de önemli olabilir.

Öğrencilerin sınav puanlarından gelişimlerinin takip edildiğini düşünelim.

<https://drive.google.com/file/d/0B0ieO-Fy3z-5ZHJHcVNpTzRlQTQ/view?usp=sharing>

adresinden RegressiontotheMean.sav dosyasını indiriniz. Bu veri setinde altı farklı sınavdan öğrencilerin aldıkları puanlar, bu puanların ortalamaları (Mean), %5 alt sınır (P95Low) ve %5 üst sınır (P95Up) verilmiştir:

ID	Exam1	Exam2	Exam3	Exam4	Exam5	Exam6	Mean	P95Low	P95Up
120801012	46	44	27	56	45	74	46.7	43.8	53.5
120801020	70	80	62	57	62	77	68.0	61.2	74.8
120801030	54	58	56	51	63	76	63.0	56.7	69.3
120801038	88	88	64	73	71	61	75.8	63.7	77.9
120801040	54	73	58	66	63	74	64.7	58.2	71.1
120801045	53	49	44	40	67	77	55.0	49.5	60.5
120801046	60	64	45	50	50	74	57.8	52.1	63.6
130801004	76	79	77	75	74	85	77.7	69.9	85.4
130801005	38	63	57	62	69	85	62.3	56.1	68.6
130801006	46	79	58	52	57	75	61.3	55.2	67.5
130801007	54	49	40	57	44	69	52.2	46.9	57.4
130801008	65	35	41	40	58	70	51.5	46.4	56.7

Ortalamasının çok üzerinde nota alan bir öğrencinin bir sonraki notunun daha düşük olacağını tahmin edebiliriz. Örneğin, 120801020 nolu öğrenci ikinci sınavdan 80 puan almıştır. Bu puan öğrencinin kendi ortalaması olan 68 puanın çok üzerindedir (%95'lik üst sınırı olan 74,8 puanın da üzerinde). Bu öğrencinin üçüncü sınav puanının 80'in altında olacağını tahmin edebiliriz. Nitekim öğrenci üçüncü sınavdan 62 puan almıştır.

Ortalamaya doğru gerileme kuralı beklenenden çok düşük not alan öğrenciler için de geçerlidir. Örneğin 120801012 numaralı öğrenciyi inceleyelim: Öğrencinin ortalama puanı 48,7'dir. Üçüncü sınavdan çok düşük not alan öğrencinin (27 puan) sonraki sınavda bu puandan daha yüksek ve ortalama puanına yaklaşan bir puan alması beklenmelidir. Nitekim öğrenci dördüncü sınavdan 56 almıştır.

Ortalamaya doğru gerileme prensibi hayatın çeşitli alanlarına uyarlanabilir ve tahmin yapmamızda yol gösterebilir.

Bölüm 43: Lojistik Regresyon

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun ikili bir bağımlı değişkeni tahmin etmede birden fazla faktörden oluşan bir model oluşturulması ve lojistik regresyon analizi hakkında bilgi sahibi olması ve bu analizi SPSS ile yapabilmesi amaçlanmıştır

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Lojistik regresyonun kullanım alanlarını açıklayabilmek
 - Binomiyal ve
 - multinomiyal lojistik regresyonun kullanım alanlarını açıklayabilmek
- Lojistik regresyon modeline konacak değişkenlerin nasıl seçileceğini açıklayabilmek
 - Tek değişkenli model
 - İkili karşılaştırmalar
- -2 log-likelihood'u açıklayabilmek
- Odds ratio'nun lojistik regresyondaki yerini açıklayabilmek
- Lojistik regresyon modelinin bağımlı değişkeni tahmin etme sınıflandırma tablosunu yorumlayabilmek
- Lojistik regresyon modelinin duyarlılık (sensitivity), özgüllük (specifity) ve doğru sınıflandırma oranını açıklayabilmek
- SPSS kullanarak lojistik regresyon yapabilmek
- Değişik lojistik regresyon yöntemlerini açıklayabilmek
 - Enter
 - Forward Selection (Conditional).
 - Forward Selection (Likelihood Ratio)
 - Forward Selection (Wald).
 - Backward Elimination (Conditional).
 - Backward Elimination (Likelihood Ratio).
 - Backward Elimination (Wald).

Diyabet hastalarımızın sigara içme durumunu belirleyen faktörleri araştırmak istiyoruz. Böylece hangi hastalarımızın sigara içmeye daha meyilli olduğunu saptayabilir, gerekli klinik yaklaşımı sunabiliriz.

Hastalarımız hakkında sigara içme durumunun yanında cinsiyet, medeni durum, çocuk sayısı, diyabetin süresi, aldığı tedavi gibi değişkenler hakkında veriye sahibiz.

Birden fazla değişkeni kullanarak **ikili bir değişkenin** (sigara içme durumu, baş ağrısı olma durumu gibi) durumunu tahmin etmek için (içiyor/içmiyor veya var/yok gibi) lojistik regresyon analizinden yararlanabiliriz.

Bağımlı değişkenimizin ikiden fazla kategori içermesi halinde (örn. çeşitli değişkenleri kullanarak diyabetik bireylerin ilaç tercihlerini tahmin etmeye çalışmak) **multinomial lojistik regresyon** analizini kullanabiliriz.

Bağımlı değişkenimizin nümerik bir değişken olması halinde bireyin alacağı değeri tahmin etmek için (örn. diyabetik bireylerimizin kan şekerlerini tahmin etmek) **lineer regresyon** analizini kullanabiliriz.

Öncelikle bağımlı değişkenimizi 1 = var, 0 = yok şeklinde kodlamalıyız. Bunu yapmamız halinde SPSS’te otomatik olarak kategoriler 1 ve 0 olarak dönüştürüleceğinden yorumlama sırasında sorun yaşayabiliriz.

Bir bireyin incelenen duruma sahip olma olasılığına (örn. sigara içiyor) p diyelim. Regresyon eşitliğinde lojistik dönüşüm uygulanır. Olasılığın logit değeri e tabanına göre hastalığın odds oranının logaritmasıdır.

$$\text{logit}(p) = \ln \frac{p}{1-p}$$

Lojistik regresyon eşitliği

Lojistik regresyon eşitliği aşağıdaki formülle gösterilir:

$$\text{logit}(p) = a + b_1x_1 + b_2x_2 + \dots + b_kx_k$$

- x = açıklayıcı değişken (cinsiyet, yaş gibi)
- p , bireyin ilgili duruma sahip olma olasılığı
- a , sabit katsayı
- $b_1, b_2, \dots$ tahmin edilen lojistik regresyon katsayıları

İlgili tahmin ettirici değişkenin **odds oranı** e^{b_1} olarak hesaplanır. Hastalığın nadir olması halinde odds oranı **rölatif risk** olarak ta değerlendirilebilir.

Eğer lojistik regresyon eşitliğinden hastalığın meydana gelme olasılığını tahmin etmek istersek;

$$z = a + b_1x_1 + b_2x_2 + \dots + b_kx_k$$

buradan da;

$$p = \frac{e^z}{1 + e^z}$$

formülünü kullanarak olasılığı hesaplayabiliriz.

Modelin oluşturulması

İyi bir lojistik regresyon analizi için öncelikle sağlam bir model oluşturmalıyız. Yani, tahmin etmeye çalıştığımız bağımlı değişkenimiz etki edebilecek

faktörleri iyi belirlemeliyiz. Bu işlemi klinik bilgimizi ve literatür verilerini kullanarak yapmalıyız. Örn. diyabetik bireylerin sigara içme durumlarının tahmininde kullanacağımız modele yaş ve cinsiyeti eklemeliyiz. Çünkü, mevcut bilgilerimiz, bu değişkenlerin her ikisinin de sigara içiciliğini belirlemede önemli olduğunu gösteriyor. Diyabetin süresinin ve medeni durumun da sigara içiciliği üzerinde etkili olabileceğini varsayabiliriz. Ancak, çocuk sayısının veya total kolesterol düzeylerinin sigara içiciliğine etkisi klinik açıdan mantıklı gözükmemektedir.

Modelimize çok fazla değişken koymak yerine az değişkenle daha kesin bir tahmin yapmayı tercih etmeliyiz. Diğer taraftan, örneklem sayımızın yeterli olmadığı durumlarda çok sayıda değişkeni modelimize eklememiz halinde analiz yapılamayacaktır.

Modelin uygunluğunun test edilmesi

- **-2 log likelihood:** ki kareye yakın bir dağılım gösterir ve anlamlı (yüksek) olması kötü bir tahmine işaret eder.
- **R kare (R square):** bağımlı değişkendeki değişimin ne kadarının modele konulan değişkenler tarafından belirlendiğini gösterir. Örn. 0,218 olması durumunda bağımlı değişkenin %21,8'inin modele konan değişkenler tarafından belirlendiği anlaşılr.
- **Sınıflandırma tablosu (classification table):** bu tablo modelin incelenen durumu ne kadar doğru sınıflandırabildiğini gösterir. Buradan duyarlık (sensitivity), özgüllük (specifity) ve genel sınıflandırma durumu hesaplanabilir.

Örnek uygulama

Konu başında verilen senaryomuzu SPSS ile analiz edelim:

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

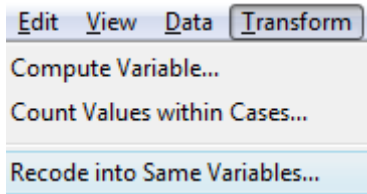
H_0 : Oluşturulan modeldeki değişkenlerin diyabet hastalarında sigara içme durumlarını tahmin etmede etkisi yoktur.

H_1 : Oluşturulan modeldeki değişkenler diyabet hastalarında sigara içme durumlarını tahmin ettirebilir.

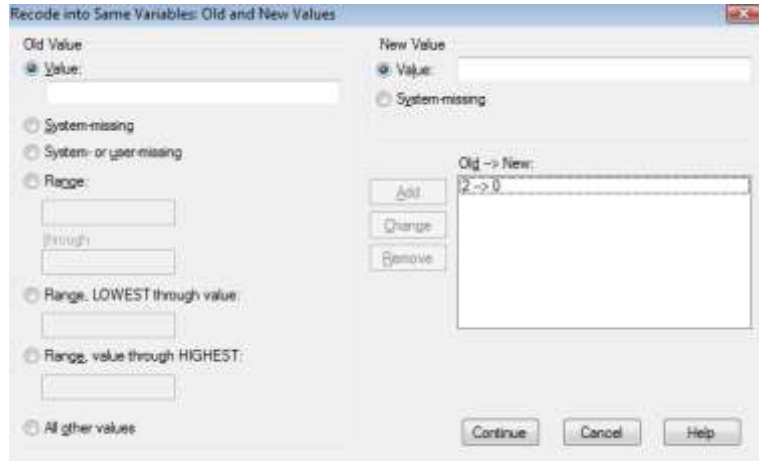
2. Verilerin toplanması:

Veriler diyabet.sav verisetinde kaydedildi. Bağımlı değişkenimiz “smoking” 1=evet, 2=hayır şeklinde kodlanmıştır. Öncelikle değişkeni 1=evet, 0=hayır olacak şekilde yeniden kodlamalıyız:

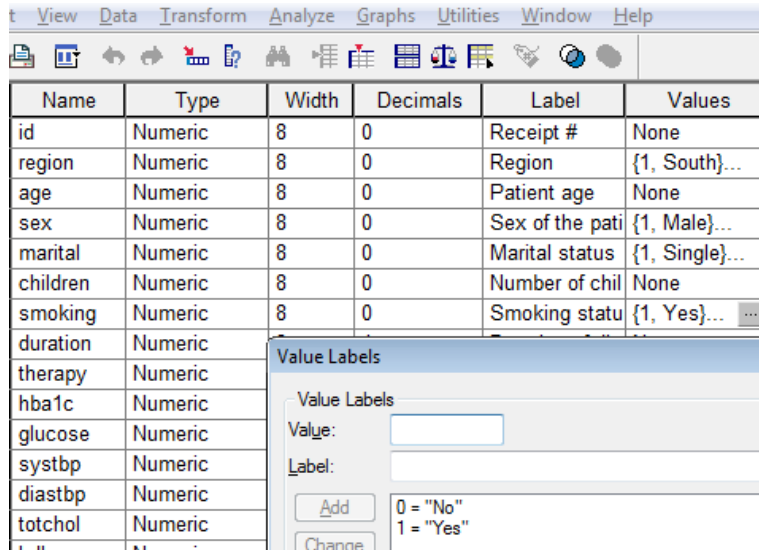
Transform>Recode into Same Variables [“smoking” değişkenini “Numeric Variables” alanına geçirelim]>



Old and New Values [Old Value = 2; New Value = 0]>Add>Continue>ok



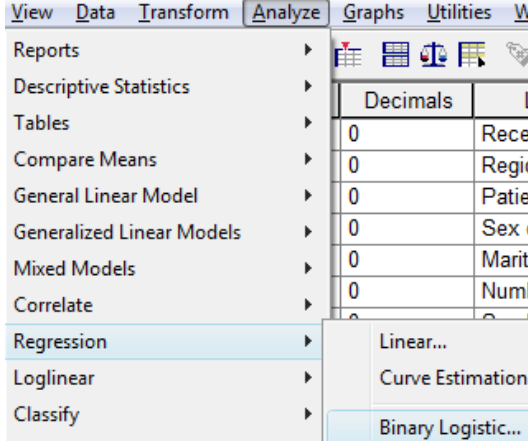
Şimdi SPSS verisetinde “Variable View” görünümünde “smoking” değişkeninin kodlamasını 1=evet, 0=hayır şeklinde değiştirmeliyiz.



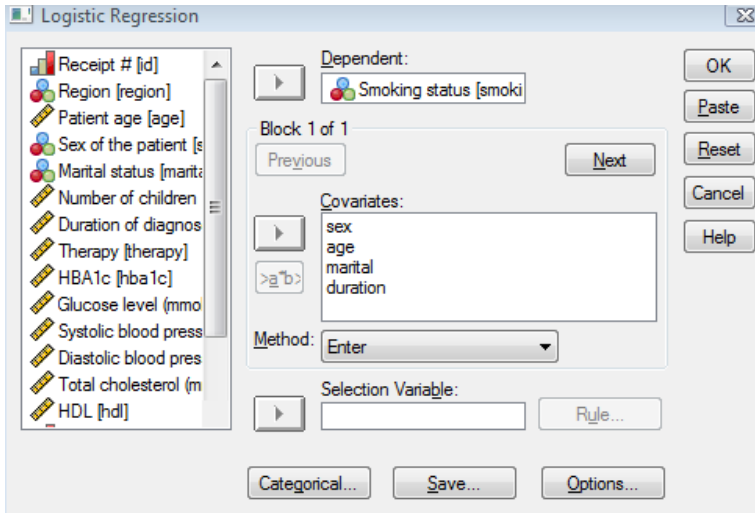
3. İlgili sıfır hipotezi için test istatistiğinin hesaplanması:

Öncelikle modelimizi belirlemeliyiz. Verisetimizdeki değişkenler içerisinde “age”, “sex”, “marital” ve “duration” değişkenlerinin sigara içme durumu üzerinde belirleyici olabileceğini varsayıyoruz (modeli doğru seçmemizin önemini burada bir kez daha vurgulamakta yarar var).

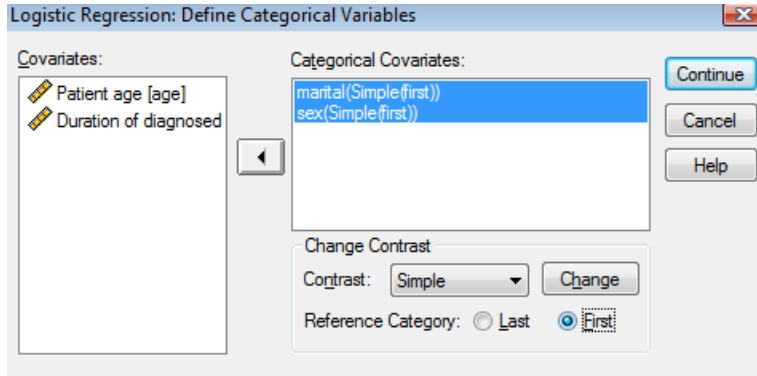
Analyze>Regression>Binary Logistic>



[“smoking” değişkenini “Dependent” alanına, “sex”, “age”, “marital” ve “duration” değişkenlerini “Covariates” alanına geçirelim.]>



Categorical>[Kategorik değişkenlerimizi (“marital” ve “sex”) “Categorical Covariates” alanına geçirelim]>[Değişkenleri işaretleyelim]>Simple>First>Change>



Continue>OK

Aşağıdaki çıktıları elde ederiz:

Case Processing Summary

Unweighted Cases(a)		N	Percent
Selected Cases	Included in Analysis	145	72,9
	Missing Cases	54	27,1
	Total	199	100,0
Unselected Cases		0	,0
Total		199	100,0

a If weight is in effect, see classification table for the total number of cases.

Verisetindeki 199 kişiden 145'i analize dahil edilmiştir. Modele eklenen değişkenlerden herhangi birinde veri eksikliği olması haline o birey analize dâhil edilmemektedir.

Dependent Variable Encoding

Original Value	Internal Value
No	0
Yes	1

Bağımlı değişkenin kodlanması “hayır=0”, “evet=1” olarak kodlanmıştır. Bilgisayarda kullanılan iç kodlama da 0 ve 1 olup bu

kodlamayla uyumludur (aksi olsaydı yorumlama sırasında kafa karışıklığı olabilirdi).

Categorical Variables Codings

		Frequency	Parameter coding		
		(1)	(2)	(3)	(1)
Marital status	Single	5	-,250	-,250	-,250
	Married	118	,750	-,250	-,250
	Divorced	6	-,250	,750	-,250
	Widow	16	-,250	-,250	,750
Sex of the patient	Male	84	-,500		
	Female	61	,500		

Kullanılan kategorik değişkenler ve kategorileri verilmiştir.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	80,607(a)	,104	,213

a Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.

-2 log likelihood yüksektir. Buradan, oluşturduğumuz modelin çok sağlam bir model olmadığını anlıyoruz. R kare değerinin 0,213 olmasından da sigara içiciliğini belirleyen faktörlerden sadece %21,3'ünün modelimizdeki değişkenler tarafından açıklanabildiğini anlıyoruz.

Classification Table^a

Observed			Predicted		
			Smoking status		Percentage Correct
			No	Yes	
Step 1	Smoking status	No	130	0	100,0
		Yes	15	0	,0
	Overall Percentage				89,7

a. The cut value is ,500

Sınıflandırma tablosu da modelin uygunluğu hakkında bilgi veriyor. Modelimiz sigara içmeyen 130 kişinin tamamını doğru olarak tahmin

ederken (özgüllük, specificity = %100), sigara içen 15 kişinin hiçbirini doğru olarak tahmin edememiştir (duyarlılık, sensitivity = %0). Modelin genel tahmin etme orantısı %89,7'dir.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1	sex(1)	-2,594	1,155	5,046	1	,025	,075
	age	-,050	,028	3,217	1	,073	,951
	marital			,176	3	,981	
	marital(1)	-,569	1,357	,176	1	,675	,566
	marital(2)	-17,469	15971,387	,000	1	,999	,000
	marital(3)	-17,481	9893,766	,000	1	,999	,000
	duration	,027	,095	,081	1	,776	1,027
	Constant	-8,721	4696,886	,000	1	,999	,000

a. Variable(s) entered on step 1: sex, age, marital, duration.

Bu tablo lojistik regresyon eşitliğini ve değişkenlerin anlamlılık düzeylerini vermektedir. Modelde kullanılan değişkenlerden sadece cinsiyet anlamlı bulunmuştur ($p=0,025$). Odds oranı cinsiyet için 0,075'tir. Cinsiyet 1=erkek, 2=kadın olarak kodlanmıştır. Bu durumda "Kadın olmak sigara içme açısından koruyucudur. Kadınların içiciliği erkeklerle karşılaştırıldığında 0,075 kattır" diyebiliriz.

4. Test istatistiğinden elde edilen değerlerin bilinen bir olasılık dağılımı ile karşılaştırılması:

Lojistik regresyonda modelin uygunluğu, lojistik regresyon eşitliği, değişkenlerin anlamlılığı bir arada değerlendirilmelidir.

50 yaşında, 10 yıldır diyabet hastası, evli bir bayan için lojistik regresyon eşitliğini yazacak olursak:

$$\text{logit}(p) = -8,7 - 0,05 \times 50 + 0,027 \times 10 - 0,569 \times 2 \text{ (evli)} - 2,594 \times 2 \text{ (erkek)} = -16,12$$

$$p = (e^{-16,12}) / (1 + e^{-16,12}) = 0,00000009312$$

Modelimize göre bu bireyin sigara içme olasılığı 100 milyonda 9'dur.

5. P değerinin ve sonuçların yorumlanması

Modelimizin uygunluğu gerek -2 log likelihood, gerekse sınıflandırma tablosuna göre iyi değildir. Bununla birlikte modele itibar edilip kullanılması düşünülürse dâhil edilen değişkenler içerisinde sigara içmeye anlamlı olarak etki eden değişkenin sadece cinsiyet olduğu görülmektedir. Bayan olmak sigara içimi açısından koruyucudur.

Alıştırma 14

Lojistik regresyon konusunu pekiştirmek için aşağıda bağlantısı verilen makaleyi okuyunuz ve soruları cevaplandırınız.

İyi N, Genç A. Lojistik Regresyon Analizi Yardımıyla Denekte Menopoz Evresine Geçişe İlişkin Bir Sınıflandırma Modelinin Elde Edilmesi. S Ü Fen Ed Fak Fen Derg 2005;25:19-27.

http://fen.selcuk.edu.tr/web/fendergi/onceki_sayilar/25/19-27.pdf veya

<http://www.aile.net/agep/istat/odev11.pdf>

1. Araştırmada incelenen bağımlı değişken nedir?
2. Bağımlı değişkeni tahmin etmede kullanılacak modele hangi değişkenlerin gireceği nasıl belirlenir?
3. Oluşturulan modelin menopozu tahmin etmedeki duyarlılık ve özgüllüğü kaçır? Doğru sınıflandırma oranı kaçır?
4. Modelin -2 log likelihoodu kaçır?
5. Modeldeki değişkenlerden menopozu tahmin etmede en etkili olanı hangisidir? Neden?
6. Tablo 5'teki sütunlardan B sütunu (ikinci sütun) neyi ifade eder?
7. Tablo 5'teki Wald değeri neyi ifade eder?
8. Tablo 5'te ATESTER değişkeni için odds oranı kaçır? Bunu klinik olarak nasıl yorumlarsınız?
9. Tablo 5'te YAŞ değişkeni için güven aralığını yorumlayınız.

Ayrıca aşağıdaki makaleyi ve konuyla ilgili benzer makaleleri de okuyunuz. Bu makalede lojistik regresyonun ayrıntılarının verilmediğine, sadece odds oranlarının belirtildiğine dikkat ediniz.

Campbell JW, Busby SC, Steyer TE. Attitudes and Beliefs About Emergency Contraception Among Patients at Academic Family Medicine Clinics. Annals of Family Medicine 2008;6(SUPPL 1):23-27.

http://www.annfammed.org/cgi/reprint/6/suppl_1/S23

Çözüm için bakınız s. 483

Alıştırma 15

Lineer regresyon ve Lojistik regresyon konuları, günlük pratiğimizde sık karşılaşılabileceğimizden ve şimdiye kadar işlediğimiz konuları anlamamızı gerektirmesi açısından önemlidir. O nedenle, bu konuyu biraz daha pekiştirmede yarar var. Elinizdeki SPSS verilerini kullanarak değişik modeller oluşturunuz ve pratik yapınız.

Aşağıda bağlantısı verilen makaleyi okuyarak soruları cevaplayınız:

Prasad R, Verma SK, Agrawal GG, Mathur N. Prediction Model for Peak Expiratory Flow in North Indian Population. The Indian Journal of Chest Diseases & Allied Sciences 2006;48:103-106.
<http://medind.nic.in/iae/t06/i2/iaet06i2p103.pdf>

1. Bu araştırmada Peak Expiratory Flow Rate'i (PEFR) tahmin etmek için oluşturulan model hangi değişkenlerden oluşmaktadır?
2. Model nasıl oluşturulur? Siz olsaydınız PEFR tahmininde nasıl bir model kurardınız?
3. Oluşturulan modele cinsiyetin dâhil edilmemesini nasıl açıklarsınız?
4. Modelde "goodness of fit (model fitting)" kaç olarak bulunmuş? Bunu nasıl yorumlarsınız?
5. 10-18 yaş grubu için PEFR tahmin formülü " $PEFR = -617.5 + 9.7 (Age) + 4.78 (Height) + 2.66 (Weight)$ " olarak verilmiş. Buna göre PEFR'e en fazla etki eden değişken hangisidir?

Ayrıca aşağıdaki ve benzeri makaleleri de incelemenizi öneririz:

Karadavut U, Genç A, Tozlu A, Kınacı İ, Aksoyak Ş, Palta Ç, Pekgör A. Nohut (Cicer arietinum L.) Bitkisinde Verime Etki Eden Bazı Karakterlerin Alternatif Regresyon Yöntemleriyle Karşılaştırılması. Tarım Bilimleri Dergisi 2005;11(3):328-333.
http://tarimbilimleri.agri.ankara.edu.tr/2005/111_3/makale18.pdf

Çözüm için bakınız s. 485

Bölüm 44: Kaplan Meier Sağkalım Analizi

Hamit ACEMOĞLU

Amaç: Bu konu sonunda okuyucunun Kaplan Meier sağkalım analizi hakkında bilgi sahibi olması ve bu testleri SPSS ile yapabilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Kaplan-Meier sağkalım analizinin yapılabilecek durumları açıklayabilmeli
- Kaplan-Meier sağkalım analizi varsayımlarını sayabilmeli
- Sansürlü (durdurulmuş) veriyi açıklayabilmeli
- SPSS kullanarak Kaplan-Meier sağkalım analizini yapabilmeli
- Kaplan-Meier sağkalım analizi SPSS çıktısını açıklayabilmeli

*Diyabetli hastalarımızda 60 ay boyunca **retinopati** gelişimini takip ettiğimizi ve cinsiyete göre (erkek/kadın) retinopati görülme durumu (var/yok) açısından fark olup olmadığını araştırdığımızı varsayalım (H_0 : Takip edilen süre içerisinde erkek ve kadınlarda retinopati gelişme durumu açısından fark yoktur). [Sonuç ölçütümüz (olay) retinopati olup olmaması iki kategorisi olan kategorik bir değişkendir. Zaman (nümerik) ve kategorik bir grup değişkeni (cinsiyet) değişkenleri söz konusudur].*

Sağkalım analizi (Survival analysis), belirli bir hastalığa maruz kalan bir bireyin hastalığın tanısından sonra, uygulanan bir girişimden (tıbbi tedavi, operasyon, kemoterapi vb.) sonra daha ne kadar yaşayabileceğini ya da hastalığın ne kadar sürede nüks edebileceğini tahmin etmek, tedavi tiplerinin yaşam süresine etkilerini incelemek amacıyla geliştirilmiş yöntemler ailesidir. Bu yöntemlerin içerisinde Yaşam tabloları (Life tables), Kaplan-Meier ve Cox regresyon analizi bulunmaktadır. Bu bölümde Kaplan Meier sağkalım analizi açıklanacaktır.

Yukarıdaki senaryo için uygun analiz yöntemi Kaplan-Meier sağkalım analizidir. Kaplan-Meier sağkalım analizi zaman-olay değişkenlerinin dağılımını incelemek için tanımlayıcı bir işlemdir. Ayrıca, bir faktör değişkeninin düzeylerine göre dağılımını karşılaştırmak ve her bir düzey için ayrı ayrı analiz yapılabilir. Örneğin, özefagus kanserli hastalara uygulanan

rezeksiyon ve diğer tedavi yönteminin hastaların yaşam süresine etkileri analiz edilebilir.

Kaplan-Meier sağkalım analizi farklı alanlarda kullanılabilmektedir. Tıp araştırmalarında farklı tedavi yöntemleri uygulanan hastalarda yaşam sürelerinin karşılaştırılması veya prognoza etki eden faktörlerin karşılaştırılması, ekonomide insanların işini kaybettikten sonra işsiz kalma sürelerini ölçmek, mühendislikte makina parçalarının arızalanmasına kadar geçen zamanın analizi gibi problemlerin çözümünde bu analiz kullanılabilir.

Tanımlar

Sağkalım analizi konusunda geçen bazı tanımlar aşağıda verilmiştir:

Süre: Yıl, ay, hafta, gün ... olabildiği gibi olgunun ortaya çıktığı andaki yaş da olabilir. Bu süre **sağkalım süresi** ya da **başarısızlık süresi** olarak da adlandırılır. Başlangıç noktasının her denek için aynı tarihte olması gerekmez. Farklı tarihte de olabilir.

Genel Sağkalım (Overall survival): Başlangıçtan **olay**'ın görülmesine kadar geçen süredir.

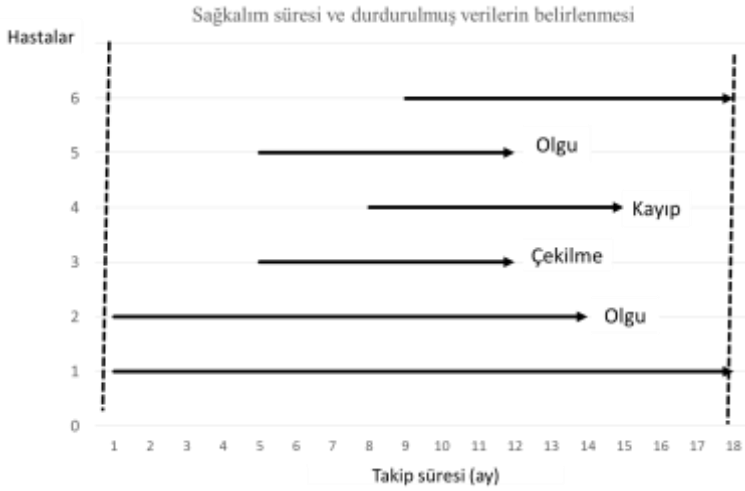
Olgu / olay (event): Ölüm, hastalık, tekrarlama, iyileşme gibi araştırmacının ilgilendiği herhangi bir olay olabilir.

Durdurulmuş / Sansürlü (censored) veri ve türleri:

Bazen verilerimizin kesin durumu hakkında bilgimiz olmayabilir. Bu durumda durdurulmuş (censored, gizli, sansürlü) veriden bahsedebiliriz. Aşağıdaki sayılan nedenlerden dolayı durdurulmuş veriden bahsederiz:

1. Bazen araştırma tamamlanmadan çeşitli nedenlerle (ölüm, yer değiştirme, araştırmaya katılmamaya karar verme...) hastalar ayrılabilirler (**withdrawing**).
2. Araştırmamızın izlem süresi içinde incelediğimiz olay görülmeyen kişilerin daha sonraki durumları bizim için belli değildir (**administrative censoring**).
3. Araştırma bitmeden hastalardan bilgi alınamaması (**lost to follow- up**)

Aşağıdaki şekilde sağkalım süresi ve durdurulmuş verilerin belirlenmesi ile ilgili bir örnek verilmiştir. Bu örnekte birinci hasta, başlangıçtan bitişe kadar izlenene, olgu gözlenmeyen ve sağkalım süresi en az 18 ay olan durdurulmuş veri örneğidir. İkinci kişinin başlangıçtan itibaren 14 ay izlendiğini, 14. ayın sonunda olgu gözlendiğini ve sağkalım süresinin 14 ay olduğunu, üçüncü hastanın 3-12. aylar arasında takip edildiği ancak 12. ayın sonunda araştırmadan ayrıldığını (çekilme), dördüncü hastanın araştırma bitmeden hastadan bilgi alınmaması nedeniyle kayıp olduğunu görmekteyiz.



Varsayımlar

Kaplan-Meier yönteminin karşılanması gereken altı varsayımı vardır. Bu varsayımlar yerine getirilmezse, Kaplan-Meier yöntemi kullanılamaz, ancak bunun yerine başka sağ kalım analizleri kullanılabilir. Bu nedenle Kaplan-Meier yöntemini kullanmadan önce çalışma tasarımının aşağıdaki altı varsayımı karşılayıp karşılamadığını kontrol etmek gerekir:

1. Olay birbirini dışlayan iki durumdan birisi olabilir. Ya durdurulmuş “censored” ya da başarısızlık “event”. Örneğin: olayın (event) “ölüm” olduğu özefagus kanserli hastaların yaşam süreleriyle ilgilendiğimizi düşünelim. Beş yıllık bir takip planlanan böyle bir araştırmanın sonunda katılanların tümü ya durdurulmuş “censored” ya da ölmüş “event” olacaktır. Bu iki durumun birlikte olması mümkün değildir.
2. Olay “event” ya da durdurulmuş durumun zamanı (“sağkalım süresi” olarak da bilinir), açıkça tanımlanmış ve kesin olarak ölçülebilmelidir. Buna ek olarak, sağkalım süresi açık bir şekilde gün, hafta, ay, yıl, ya da başka bir zaman ölçümü şeklinde hesaplanmış olmalıdır.
3. Mümkün olduğu kadar, sol durdurulmuş (left censored) durumdan kaçınılmalı veya minimize edilmelidir. Bir deneyin başlangıç noktası kolayca tanımlanabilir değilse sol-durdurulmuş durum oluşur. Yine, özefagus kanserli hastaların hayatta kalma süreleri ile ilgilendiğimizi düşünelim. “İdeal” başlangıç noktası kanser geliştiği andan itibaren hayatta kalma süresini ölçmek olacaktır.

4. Durdurulma ve olay bağımsız olmalıdır. Vakaların sansürlü olma sebepleri olayla ilgili olmamalıdır. Bunun olabilmesi için bir katılımcının sansürlü olduğunu kaydettiğimizde sansür nedeniyle kişide olay gelişme olasılığı (örneğin ölüm) daha yüksek olmamalıdır.
5. Vakaların dahil edildikleri süre içerisinde hataya (bias) yol açabilecek bir trend söz konusu olmamalıdır. Örneğin 500 deri kanserli vaka toplayacak ve sağkalım bakacak olsak 500 vakayı bulmak zaman alacağından ve bu zaman içerisinde yeni ilaçlar çıkmış olabileceğinden dolayı yakın zamandaki vakaların sağkalım süresi daha uzun olacaktır (sağdan durdurulmuş). Tersine belli bir tarihten sonra ulusal bir kanser tarama programı başlatılmış olsa, tarama programından önceki vakaların sağkalım süreleri -nispeten geç teşhis nedeniyle- daha kısa olacaktır (soldan durdurulmuş).
6. Gruplar arasındaki durdurulmuş veri miktarı ve deseni aynı olmalıdır.

Sağkalım eğrilerinin karşılaştırılması

Kanser hastalarının sağkalım olasılığını belirleyen pek çok prognostik öge vardır. Bunlar arasında yaş, cinsiyet, hastaya uygulanan tedavi yöntemi, kanserin hangi evrede olduğu gibi özellikler sayılabilir. Örneğimizde, cinsiyetin retinopati gelişimi üzerinde etkili olup olmadığını araştırıyoruz.

Araştırmacılar hastaların sağkalımları üzerinde bu prognostik öğelerin etkili olup olmadığını karşılaştırmak istediğinde “*sağkalım eğrilerinin karşılaştırılması*” yöntemleri kullanır.

Sağkalım eğrilerinin karşılaştırılmasında kullanılan Logrank, Breslow ve Tarone-Ware testlerini seçebilirsiniz. Çoğunlukla her üç test de aynı sonucu verir. Hangi testi seçeceğiniz sağkalım dağılımının nasıl olacağına bağlıdır; en iyi test diye bir şey yoktur. Seçeceğiniz test verilerinize bağlı olarak değişebilir.

SPSS’te Kaplan –Meier Sağkalım Analizi

<https://drive.google.com/file/d/0B0ieO-Fy3z-5a3cydjLPQ0YtYk0/view?usp=sharing> bağlantısından diyabet.sav SPSS dosyasını indiriniz ve açınız.

Hipotez testleri için 5 basamaklı genel yaklaşımımızı uygulayacak olursak:

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması

H_0 : Takip edilen süre içerisinde erkek ve kadınlarda retinopati gelişme durumu açısından fark yoktur.

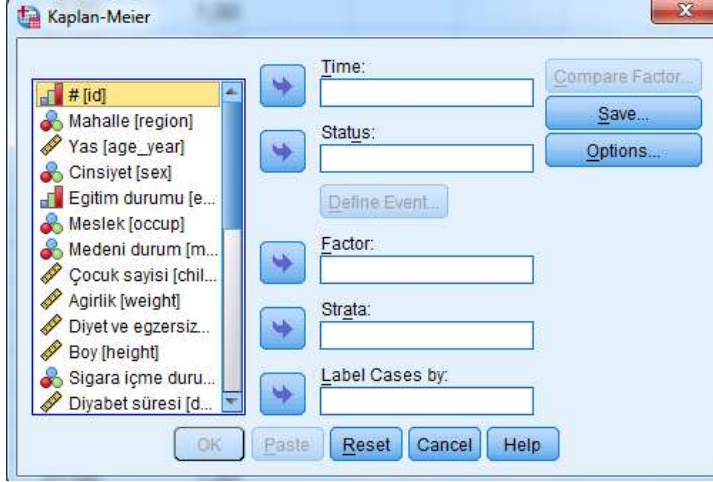
H_1 : Takip edilen süre içerisinde erkek ve kadınlarda retinopati gelişme durumu açısından fark vardır.

2. Verilerin toplanması

Verilerimizi topladık ve SPSS’e girdik. Diyabet.sav verisetinde “Olay”, “izlemsüresi” ve “sex” değişkenlerini oluşturduk.

3. İlgili sıfır hipotezi için *test istatistiğinin* hesaplanması

Analyze > Survival >Kaplan Meier

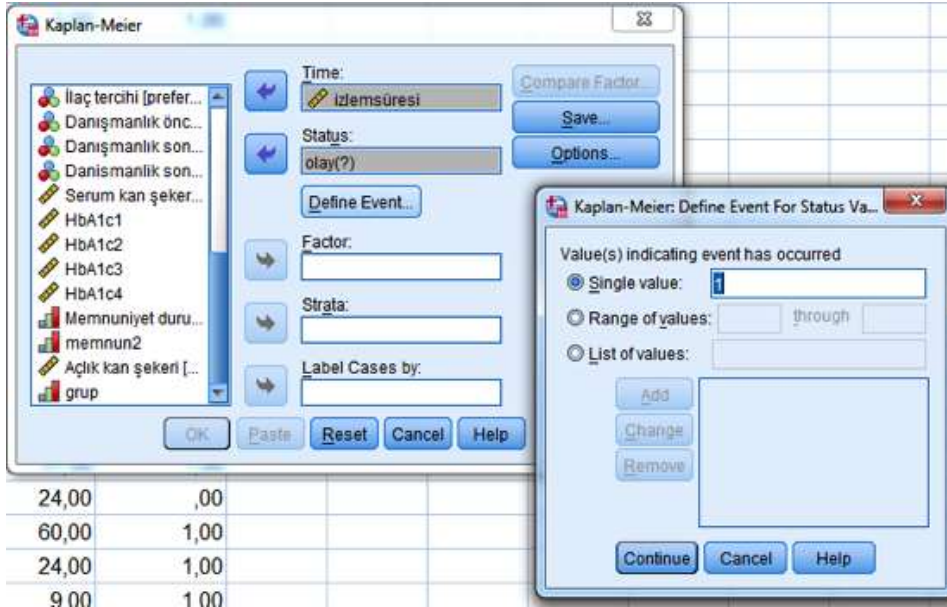


[“Time” alanına “izlemsüresi” değişkenini taşıyın].

[“Status” alanına hastanın son izlemindeki durumunu gösteren “olay” değişkenini taşıyalım].

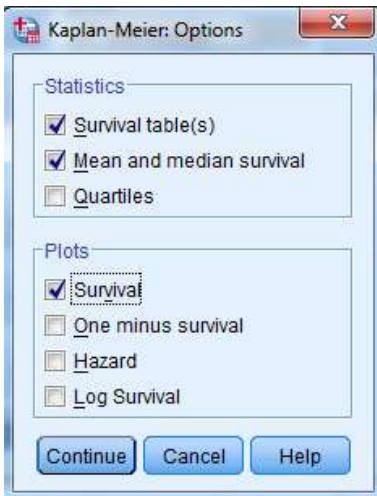
Diyabet.sav veri setinde “olay” adlı değişken hastada retinopati varsa “1”, yoksa “0” olarak kodlanmıştır.

[“Status” penceresine “olay” değişkenini aktardıktan sonra “Define Event”i tıklayarak “retinopati”ye “1”in karşılık geldiğini yazın].



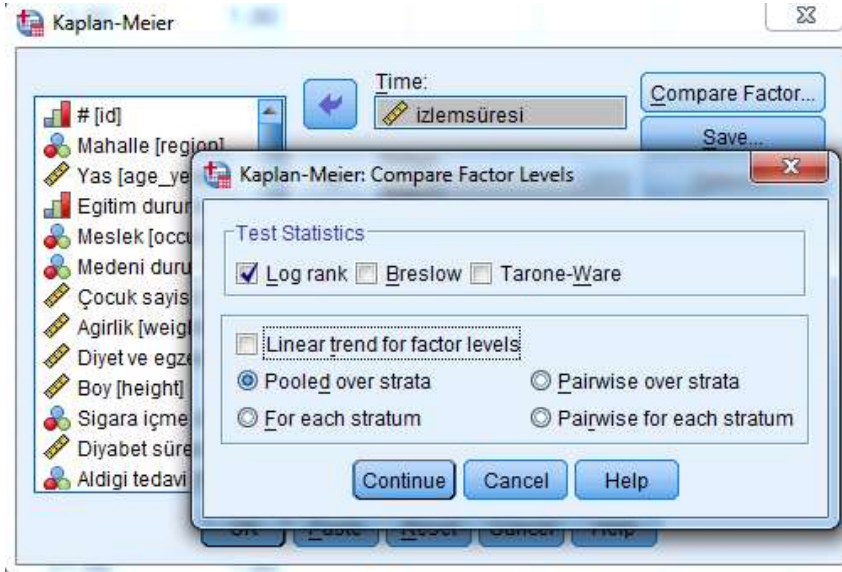
[“Factor” penceresine bir değişken seçilmemişse, veritabanındaki tüm hastaların sağkalım analizi yapılır. Alt grupların ayrı ayrı sağkalım analizlerini elde etmek için “Factor” penceresine alt grupları belirten değişkeni taşıyın. Örneğimizde Sex değişkeni kadın ve erkek olarak kodlanmıştır. “Factor” penceresine Sex değişkenini aktaralım.]

[“Option” seçeneğini tıklayarak “Statistics” bölümünde “Survival table(s)” ve “Mean and median survival”, “Plots” bölümünde ise “Survival” grafiğini işaretleyelim.]



Sağkalım eğrilerinin karşılaştırılmasında en çok kullanılan Logrank testi bu örnekte kullanılacaktır.

Eğer faktör düzeyleri eşit aralıklı ise “Linear trend for factor levels” seçeneği ile faktör alt gruplarının sağkalım eğrilerinin artma ya da azalma eğilimi test edilebilir. Örneğin bir ilacın 10, 20, 30 ve 40 mg/gün gibi eşit aralıklı dört farklı dozu ile sağkalım süresi arasında aynı yönde doğrusal bir ilişki varsa, yani doz arttıkça sağkalım süresi artıyorsa, bu durum “Linear trend for factor levels” ile saptanabilir.



Continue>OK

SPSS çıktısı:

Kaplan-Meier				
Case Processing Summary				
Cinsiyet	Total N	N of Events	Censored	
			N	Percent
Male	236	76	160	67,8%
Female	194	68	126	64,9%
Overall	430	144	286	66,5%

Survival Table

Cinsiyet	Time	Status	Cumulative Proportion Surviving at the Time		N of Cumulative Events	N of Remaining Cases	
			Estimate	Std. Error			
Male	1	1,000	Durdurulmuş		0	235	
	2	1,000	Durdurulmuş		0	234	
	55	12,000	Retinopati	,930	,018	15	181
	114	23,000	Retinopati	,850	,027	28	122
	117	24,000	Retinopati			29	119
	118	24,000	Retinopati			30	118
	119	24,000	Retinopati	,829	,029	31	117
	162	34,000	Retinopati	,710	,038	45	74
	164	36,000	Retinopati			46	72
	165	36,000	Retinopati	,691	,040	47	71
	188	46,000	Retinopati	,576	,047	57	48
	189	46,000	Durdurulmuş			57	47
	193	48,000	Retinopati	,527	,049	61	43
	207	56,000	Retinopati	,484	,051	64	29
	224	60,000	Retinopati	,242	,056	76	12
Female	1	1,000	Durdurulmuş		0	193	
	2	2,000	Retinopati	,995	,005	1	192
	38	12,000	Retinopati	,916	,021	15	156
	80	24,000	Retinopati	,822	,030	29	114
	116	36,000	Retinopati	,715	,038	42	78
	132	48,000	Retinopati	,692	,040	44	62
	165	58,000	Retinopati	,487	,054	59	29
	169	58,000	Durdurulmuş			59	25
	177	60,000	Retinopati			67	17
	178	60,000	Retinopati	,312	,058	68	16

*tablo kısaltılmıştır.

*tablo kısaltılmıştır.

Means and Medians for Survival Time

Cinsiyet	Mean ^a				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Male	45,3	1,5	42,5	48,2	56,0	2,1	51,8	60,2
Female	47,2	1,5	44,3	50,1	58,0	1,2	55,7	60,3
Overall	46,3	1,0	44,3	48,3	58,0	1,0	56,0	60,0

a. Estimation is limited to the largest survival time if it is censored.

Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	,749	1	,387

Test of equality of survival distributions for the different levels of Cinsiyet.

SPSS tablolarında raporlanacak bazı değerler

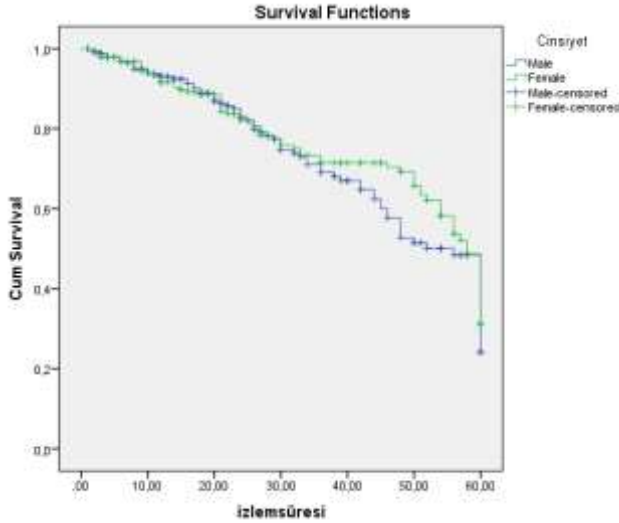
Olay (event) sayısı=144/430=%33,5

Ortalama sağkalım süresi (ay) [%95 CI] = 46,3 [44,3-48,36]

Ortanca sağkalım süresi (ay) [%95 CI] =58 [56-60]

1. yıl kümülatif sağkalım hızı $\pm$ SEKadın=% 93,0 $\pm$ % 1,8Erkek= % 91,63 $\pm$ % 2,15. yıl kümülatif sağkalım hızı $\pm$ SEKadın=% 24,2 $\pm$ % 5,6Erkek= % 31,2 $\pm$ % 5,8

Kaplan-Meier bulgularını görsel olarak sunan aşağıdaki grafik de çok kullanışlı olabilir:



4. Test istatistiğinden elde edilen değerin bilinen bir olasılık dağılımı ile karşılaştırılması

Elde edile test istatistiği Ki-kare dağılımı ile karşılaştırılır.

5. *P* değerinin ve sonuçların yorumlanması ve raporlanması

Overall Comparisons tablosunda Log Rank (Mantel-Cox) *p* değeri 0,387 hesaplanmıştır. Bu değer aşağıdaki şekilde raporlanabilir:

“Log Rank testin kullanıldığı sağkalım analizinde kadın ve erkek gruplarında retinopati gelişmesi durumu arasında istatistiksel olarak anlamlı bir fark bulunmadı ($\chi^2 = 0,75$, $p = 0,387$).”

Alıştırma 16

Artrit tedavisi için yeni geliştirilen antienflematuvar bir ilacın kullanılan tedaviye göre etki süresini karşılaştırmak istiyoruz. Klinik çalışmanın verileri “agritedavisi.sav” isimli dosyaya (<https://drive.google.com/file/d/0B0ieO-Fy3z-5VWZlNjhQakR5Y3c/view?usp=sharing>) kaydedilmiştir. Kaplan Meier sağkalım analizini kullanarak bu iki tedavinin etkinliğini “etki süreleri” açısından analiz ederek sonucu raporlayınız.

Çözüm için bakınız s. 486

Bölüm 45: Cox Regresyon Analizi

Hamit ACEMOĞLU

Amaç; Cox regresyon analizi hakkında bilgi vermek ve bu analiz yöntemini SPSS programında yapabilmek.

Hedefler: Bu konu sonunda katılımcılar;

- Cox regresyon analizi yapılabilecek durumları açıklayabilmeli
- Zarar Oranı (Hazard Ratio) açıklayabilmeli
- SPSS kullanarak Cox regresyon analizi yapabilmeli
- Cox regresyon analizi SPSS çıktısını açıklayabilmeli

*Diyabetli hastalarımızda 60 ay boyunca retinopati gelişimi açısından takip ettiğimizi ve **cinsiyet ve eğitim durumunun retinopati** görülmesi üzerindeki etkisini araştırdığımızı varsayalım. (H_0 : 60 aylık takip süresinde cinsiyet ve/veya eğitim durumunun retinopati gelişimi üzerinde etkisi yoktur).*

Uygulanan tedavi yönteminin yanı sıra bireyin yaşı, cinsiyeti, kanser evresi gibi pek çok prognostik faktör olguların sağkalım durumunu belirler. Sağkalım “**Bağımlı değişken**” olarak adlandırıldığında olguların sağ kalma durumunu belirleyen prognostik faktörlere “açıklayıcı değişkenler, kafa karıştırıcılar, risk etmenleri ya da **bağımsız değişkenler**” denir. Sağkalım süresi üzerinde iki ya da fazla sayıda bağımsız değişkenin bir arada belirleyici etkisinin olduğu düşünüldüğünde, çok değişkenli analiz yöntemlerinin uygulanması gerekir.

Özellikleri dolayısıyla sağkalım üzerinde iki ya da daha çok bağımsız değişkenin etkisini ortaya koymada Cox oransal risk modeli ya da Cox regresyon analizi olarak adlandırılan yöntem kullanılır. Cox regresyon analizi Cox tarafından geliştirilen, tamamlanmamış (sansürlü) gözlemlerde **zamana bağlı değişimi içeren** bir yöntemdir. Cox regresyon analizini lojistik regresyondan ayıran yönü de **zamana bağlı değişimi** içermesidir.

Bağımlı değişken gruplanmış olan sağ kalma (sansürlü gözlem) ve olay/olgu ikili seçeneğinden oluşur. Bağımsız değişkenler gruplanmış ya da sürekli nitelikte olabilir.

İki ya da fazla sayıdaki değişkenin bir arada değerlendirildiği bir yapı oluşturulması ve yapı içine istenen değişkenlerin alınabilmesi nedeniyle bir “model” den söz edilmektedir.

Çok değişkenli Cox regresyon analizinde bir arada sınanacak değişkenlerin doğru biçimde belirlenmesi ve doğru modelin oluşturulması, elde edilecek sonuçların güvenilirliği açısından önemlidir.

Cox Regresyon yönteminde hangi değişkenlerin analize alınacağı ya da alınmayacağı ve nasıl değerlendirileceği istatistiksel bir yaklaşımdan çok klinik bir karar sürecini tanımlar.

Örneğimizde retinopati üzerinde cinsiyet ve eğitim durumunun etkisi incelenmek istendiğinden bu durumun analizi için Cox regression analizi kullanılmalıdır.

Zarar Oranı (Hazard Ratio, HR)

Logrank karşılaştırmasından elde edilen değerlerden yararlanılarak Zarar Oranı (Hazard Ratio, HR) olarak adlandırılan bir değer elde edilebilir. Bu değer, her bir grubun gözlenen değerinin beklenen değerine bölümü ile elde edilen oranların birbirine oranlanması ile elde edilir ve Odds Oranı gibi yorumlanabilir. Elde edilen değer bir grubun diğerine göre herhangi bir zamanda kaç kat fazla ölüm riskine neden olduğunu gösterir. Daha çok sağaltım yöntemlerinin (ve ilaçların) karşılaştırılmasında yararlanıldığı için adında “zarar” sözcüğü vardır.

SPSS’te Kaplan – Cox regresyon Analizi

<https://drive.google.com/file/d/0B0ieO-Fy3z-5a3cydjlPQ0YtYk0/view?usp=sharing> adresinden diyabet.sav SPSS verisetini indiriniz ve açınız.

Hipotez testleri için 5 basamaklı genel yaklaşımımızı uygulayacak olursak:

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması

H_0 :60 aylık takip süresinde cinsiyet ve/veya eğitim durumunun retinopati gelişimi üzerinde etkisi yoktur.

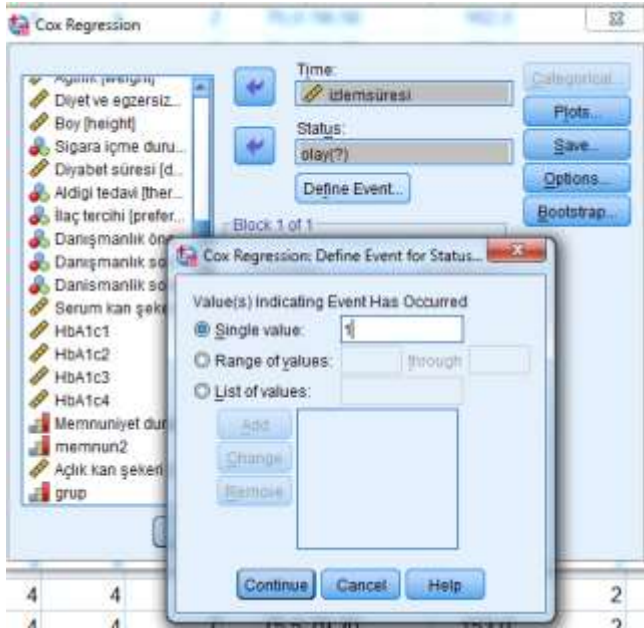
H_1 :60 aylık takip süresinde cinsiyet ve/veya eğitim durumunun retinopati gelişimi üzerinde etkisi vardır.

2. Verilerin toplanması

Verilerimizi topladık ve SPSS’e girdik. Diyabet.sav verisetinde “olay”, “izlemsüresi” ve “sex” değişkenlerini oluşturduk.

3. İlgili sıfır hipotezi için *test istatistiğinin* hesaplanması

Analyze > Survival > Cox Regression

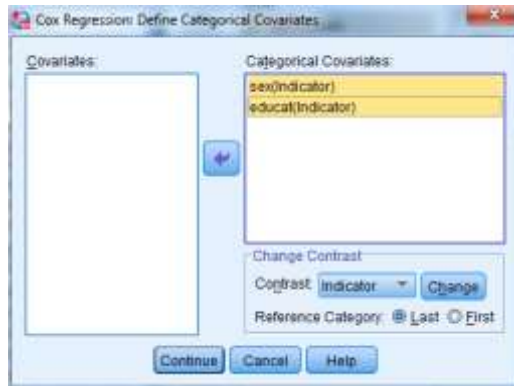


[“Time” penceresine izlem süresini sürükleyin].

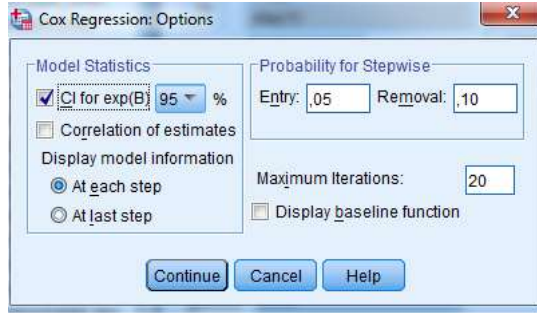
[“Status” penceresine hastanın son izlemindeki durumunu gösteren değişkeni seçin. Örneğimizde “olay” adlı bir değişkene hastada retinopati var ise “1”, yok ise “0” olarak girildiğinden, “Status” penceresine “olay” değişkenini sürükleyin].

>Define Event [“retinopati”ye “1”in karşılık geldiğini belirleyin].

>Covariates [sağkalım süresine etkilerinin inceleneceği kategorik veya numeric bağımsız değişkenleri seçin. Örneğin hastalarda retinopati gelişimi üzerinde etkili olduğu düşünülen sex ve eğitim değişkenlerini Covariate penceresine taşıyalım].



>Categorical [“Define Categorical Covariates” penceresinde “tedavi” değişkenini “Categorical Covariates” listesine taşıyalım].



>Option [Hazard Ratio (HR) için %95 güven aralığını işaretleyelim].

>Continue>OK

SPSS çıktısı

Omnibus Tests of Model Coefficients

-2 Log Likelihood
1457,435

Block 1: Method = Enter

Omnibus Tests of Model Coefficients^a

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
1456,318	1,129	5	,952	1,117	5	,953	1,117	5	,953

a. Beginning Block Number 1. Method = Enter

Tabloda ilk olarak değerlendirilen ölçüt sağkalımı belirlemede kullanılan bağımsız (açıklayıcı) değişken ile oluşturulan analizin uyumudur. Bunun için “omnibus tests of model coefficients overall score” Ki-Kare değeri ve anlamlılık düzeyine bakılır. Tabloya göre modelin uyumu istatistiksel olarak anlamlı bulunmamıştır (p=0.953).

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95,0% CI for Exp(B)	
							Lower	Upper
sex	,116	,192	,364	1	,546	1,123	,770	1,638
educat			,286	4	,991			
educat(1)	-,027	,432	,004	1	,951	,974	,417	2,271
educat(2)	,063	,447	,020	1	,888	1,065	,444	2,557
educat(3)	-,011	,484	,001	1	,981	,989	,383	2,551
educat(4)	,111	,489	,051	1	,821	1,117	,429	2,911

Bulunan Exp(B)-HR değeri cinsiyet için şöyle yorumlanabilir: Kadın ve erkek cinsiyetler arasında retinopati gelişmesi açısından istatistiksel olarak anlamlı bir fark bulunmadı (HR=1,123 [0,77-1,64]; p=0,546).

Alıştırma 17

Artrit tedavisi için yeni geliştirilen antienflematuvar bir ilacın kullanılan tedaviye göre etki süresi karşılaştırmak istiyoruz. Klinik çalışmanın verileri “agrittedavisi.sav” (<https://drive.google.com/open?id=0B0ieO-Fy3z-5VWZlNjhQakR5Y3c>) isimli dosyaya kaydedilmiştir. Cox regresiyon analizini kullanarak tedavinin etkinliği etkileyen iki farklı ilaç tedavisi, cinsiyet, yaş ve sağlık durumunun “etki süreleri” de göz önüne alarak analiz ederek sonucu raporlayın.

Çözüm için bakınız s. 487.

Bölüm 46: Nümerik Veri Uyum Analizi (Sınıf İçi Korelasyon)

Hamit ACEMOĞLU

Amaç: Bu konu sonunda okuyucunun iki veya daha fazla gözlemciden elde edilen nümerik verilerin uyum analizinde kullanılan sınıf içi korelasyon katsayısı (SKK)- (Intraclass Correlation Coefficient - ICC) hakkında bilgi sahibi olması ve bu testi uygulayabilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Testin hangi durumlarda kullanılacağını açıklayabilmek
- Testin mantığını açıklayabilmek
- SPSS'te sınıf içi korelasyon analizini yapabilmek
- SPSS çıktısını yorumlayabilmek

Kliniğimizde takip edilen diyabet hastalarının tedaviye uyumları iki farklı asistan tarafından değerlendirilmiştir. Hastaların tedavi uyumu 0-10 puan arasında puanlanmıştır. Aynı hastaları değerlendiren asistanlar arasında uyumu araştırmak istiyoruz. İncelediğimiz değişken nümerik bir değişkendir. (H_0 : İki asistanın hastaların tedaviye uyum değerlendirmeleri sıfırdan farklı değildir).

Uygulayıcılar arası güvenilirlik tek bir formun iki uygulayıcı tarafından uygulanması ve aralarındaki korelasyona bakılması ile ölçülür. Ölçek nümerik bir ölçüm yapıyorsa iki araştırmacının uygulamasının ne kadar korelasyon gösterdiğine (Intraclass Correlation Coefficient- ICC) bakılır (1).

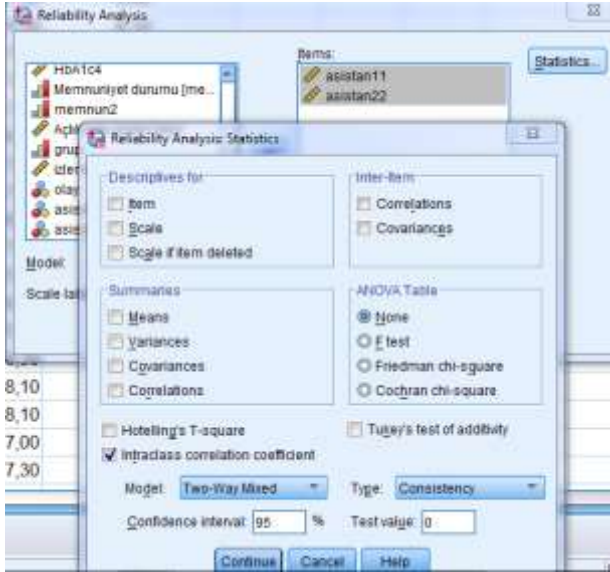
Sağlık bilimlerinde, gözlemci uyumunun incelendiği güvenilirlik çalışmalarına sıklıkla rastlanmaktadır. Bu tür çalışmalarda en temel amaç, aynı denekten elde edilen çoklu ölçümlerin uyumunun değerlendirilmesidir. Bir bireyden elde edilen çoklu ölçüm, aynı gözlemcinin tekrarlı ölçümleri olabileceği gibi, iki ya da daha fazla sayıda gözlemcinin ölçümleri de olabilir. Ölçümler nümerik olduğunda ilk durumda, gözlemci-içi (intra-rater), ikinci durumda ise gözlemciler-arası (inter-rater) uyum, sınıf içi korelasyon katsayısı (SKK) kullanılarak değerlendirilir (2) .

Yukarıdaki senaryodaki iki asistanın hastaların tedaviye uyum değerlendirmelerinin güvenilirliğini ölçmek için sınıf içi korelasyon katsayısı hesaplanabilir.

Örneğin SPSS ile yapılması

1. diyabet.sav veri setine asistan11 ve asistan22 değişkenleri oluşturularak iki asistanın gözlem değerleri girildi.

2. *Analyze>Scale> Reliability Analysis> Açılan pencerede Items bölümüne [“asistan11” ve “asistan22” değişkenlerini taşıyalım.]>Statistics [Intraclass correlation coefficient işaretleyelim]*



Model bölümünde üç seçenek bulunmaktadır. SKK hesaplamasında kullanılan tek yönlü modellerde ölçüm/gözlemci etkisi her zaman rastgele alınır ve kullanılan model “One-Way Random ” olarak adlandırılır. İki yönlü modellerde ölçüm/gözlemci etkisinin rastgele ya da sabit olmasına bağlı olarak, kullanılan modeller sırasıyla “Two-Way Random” ve “Two-Way Mixed” olarak adlandırılır. One-Way Random modeli hangi ölçüm

birinci ölçümle elde edildiği hangisinin ikinci ölçümle elde edildiğinin belirlenemediği durumlarda kullanılır. n sayıda kişinin değerleri büyük bir kitleden seçilen k sayıda gözlemci tarafından değerlendirilirse Two-Way Random model kullanılmalıdır. n sayıda kişinin değerleri k sayıda gözlemci tarafından değerlendirilirse Two-Way Mixed model kullanılmalıdır ve örneğimiz de buna uyduğundan bu model seçilmelidir.

3. Araştırmanın amacına göre Type bölümünde ya “tutarlılık-consistensy” ya da “mutlak uyum- absolute agreement” seçilir. Mutlak uyumda gözlemci değerlerinin aynı olup olmadığı incelenirken, Tutarlılıkta SKK tahminleri toplanabilirlik indeksidir ve bu indeks bir değişkenin (y) bir diğer değişkene (x) bir sabit ile eşitlenmesinin derecesini ölçer ($y=x+b$) (2). Örneğimizde “consistency” seçilmiştir.

Continue>ok.

Aşağıdaki çıktıyı elde ederiz:

Item Statistics

	Mean	Std. Deviation	N
asistan11	6,9644	,86170	430
asistan22	8,8879	,86423	430

Intraclass Correlation Coefficient							
	Intraclass Correlation	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	,903 ^a	,884	,919	19,694	429	429	,000
Average Measures	,949 ^c	,939	,958	19,694	429	429	,000
Two-way mixed effects model where people effects are random and measures effects are fixed.							

Gözlem değerleri tek bir ölçüme ait ise “Single Measures” satırındaki değerler ortalamaya ait ise “Average Measures” satırındaki değerler kullanılır. Örneğimizde SKK- Intraclass Correlation değeri 0,903 olarak oldukça yüksek bulunmuş ve %95 güven aralığı da [0,884-0,919] oldukça dar olarak hesaplanmıştır. Buna göre iki asistanın hasta değerlendirmelerinin yüksek düzeyde benzer olduğu yorumunu yapabiliriz.

Kaynaklar

1. Aktürk Z, Acemoğlu H. Tıbbi araştırmalarda güvenilirlik ve geçerlilik. Dicle Tıp Dergisi 2012; 39 (2): 316-319.
2. Can A, Öztuna D, Yasemin G. Sağlık araştırmalarında sınıf içi korelasyon katsayısının kullanımı. Türkiye Klinikleri Journal of Biostatistics. 2009;1(2):59-64.

Bölüm 47: Örneklem Hesabı

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun çeşitli araştırma tasarımları için örneklem ve güç hesabı yapma konusunda bilgi sahibi olmaları ve bunları uygulayabilmeleri amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Örneklem sayısına etki eden faktörleri sayabilmek
- Effect size, Effect of interest, standardized difference, hata marjı ve önemlilik derecesini açıklayabilmek
- Çeşitli istatistik analizler için gerekli örneklem hesabını yapabilmek
 - Independent samples t test
 - Ki kare
 - Tek yönlü ANOVA

Örneklem sayısı ve güç hesabında birbirine etki eden faktörlerden daha önce bahsedilmişti (Bakınız s. 188). Bu bölümde önceki bilgiler kısaca tekrar edilecek ve bazı senaryolar için örneklem hesabı yapılacaktır.

Örneklemimizin sayısını hesaplamak için kullanacağımız formül kullanacağımız istatistik analiz yöntemine göre değişiklik gösterir. Genelde formülde girmemiz gereken bilgiler daha önce de bahsedilen, alfa yanılma payı, hipotezin yönü, farkların önemlilik derecesi, standart sapma ve istenen güç düzeyidir.

Araştırmamızın planlama aşamasında gücünü hesaplamamız çok önemlidir. Böylece -eğer varsa- gruplar arasında saptamayı hedeflediğimiz farkın miktarını (effect of interest) ve araştırmamızın *yeterli* bir güce sahip olup olmadığını önceden belirlemiş oluruz (*yeterli güç*'ten genelde %80 anlaşılır). Sadece %40'lık bir güce sahip bir araştırmayı yürütmek etik açıdan kabul edilemez olduğu gibi ekonomik açıdan da bir israftır.

Araştırmanın gücüne etki eden başlıca faktörleri tekrar edecek olursak:

1. **Örneklem büyüklüğü:** örneklem arttıkça araştırmanın gücü de artar. Dolayısıyla büyük bir örneklem – eğer varsa – klinik olarak daha küçük bir farkı saptayabilir.
2. **Gözlemlerin değişkenliği:** gözlemlerin standart sapması düştükçe araştırmanın gücü de artar.
3. **Effect of interest (=etki genişliği):** gruplarımız arasındaki küçük bir farkı yakalamak istediğimizde araştırmanın gücü düşük olur. Büyük farkları saptama konusunda güç daha fazladır.

4. **Anlamlılık düzeyi:** anlamlılık düzeyini yüksek tuttuğumuzda (yani tip I hata payını artırdığımızda) araştırmamızın gücü de artar. Bu durumda tip II hata da azalmış olur.
5. **Hipotezin Yönü:** Hipotezin tek yönlü veya iki yönlü olarak kurulması örneklem sayısını etkileyecektir. İki yönlü hipotez kurulması durumunda daha fazla örneklem gerekecektir.
6. **Hipotez Testi:** Yapılacak hipotez testine göre örneklem hesabı değişecektir. Araştırmanın gücünü hesaplariken öncelikle hipotez testine karar vermek gerekecektir.
7. **Araştırma Tasarımı:** Araştırma tasarımına göre bazı araştırmalarda veri kayıpları beklenebilir. Örneğin, araştırmadan ayrılanların %10 civarında olması beklenirse, bu durumda örneklem sayısının üzerine %10 ilave edilmelidir. Diğer taraftan, örneklemin alt kümelerden oluşması halinde kümelerin içindeki bireyler birbirine benzeyeceğinden alınan örneklem sayısı küme katsayısı (desen etkisi) dikkate alınarak artırılmalıdır.

Desen Etkisi: Bu noktada desen etkisinin nasıl hesaplanacağını belirtmemizde fayda var. Desen etkisi küme içi korelasyon katsayısına (Intraclass Correlation Coefficient; ICC) dayanır. ICC şu şekilde hesaplanır:

$$KIK = \frac{SS_{ka}^2}{(SS_{ka}^2 + SS_{ki}^2)}$$

KIK=Küme içi korelasyon katsayısı

SS=Standart sapma

ka=kümeler arası

ki=küme içi

Desen etkisi (DE) ise aşağıdaki formüle göre hesaplanır:

$$DE = 1 + (n - 1) * KIK$$

n=kümelerdeki ortalama birey sayısı

Örnek verecek olursak: Araştırmamızda 10 farklı aile hekiminin gördükleri hastalardan veri toplayacağız. Her bir aile hekiminden 10 hasta dahil etmek suretiyle 100 hastalık bir örneklem hedefliyoruz. Örneklemimizdeki bireylerin sistolik kan basınçlarını (mmHg) ölçeceğiz. Sistolik kan basıncının standart sapması küme içi 5, kümeler arasında ise 6 olsun. Bu durumda:

$$KIK = \frac{36}{36 + 25} = 0,59$$

$$DE = 1 + (10 - 1) * 0,59 = 6,3$$

Etkili örneklem sayısı=gerçek örneklem/desen etkisi formülünden $100/6,3=15,8$

Örneklem hesabımızı formülleri kullanarak yapabileceğimiz gibi bu amaçla geliştirilmiş yazılımlardan da yararlanabiliriz. İstatistik paket programlarının çoğunda örneklem hesabıyla ilgili modüller de vardır. SPSS'te örneklem hesabı için bir araç sunulmamıştır. Bu amaçla aşağıdaki yazılımları inceleyebilirsiniz:

1. PASS (Power Analysis and Sample Size Calculation Software). Örneklem hesabı için geliştirilmiş bir yazılımdır. <http://www.ncss.com/pass.html> .
2. MINITAB™. Profesyonel bir paket istatistik yazılımıdır. <http://www.minitab.com/> .
3. MedCalc™. Profesyonel bir paket istatistik yazılımıdır. <http://www.medcalc.com/> .
4. Russ Lenth'in Java modülü. Web üzerinden veya Java modülünü kullanarak hesap yapabileceğiniz ücretsiz bir uygulamadır. <http://www.stat.uiowa.edu/~rlenth/Power/> .
5. G*Power. Düsseldorf Heinrich Heine Üniversitesi'nin geliştirdiği ücretsiz bir örneklem hesabı yazılımıdır. Bu yazılımla ilgili örnek çözümleri ve videoları <http://www.mormonsandscience.com/gpower-guide.html> adresinde bulabilirsiniz. <http://www.gpower.hhu.de/>.

Tanımlayıcı istatistikler için örneklem büyüklüğü formülleri

Sonuç ölçütümüz kategorik ise:

Hedef toplumdaki birey sayısını bilip bilmemize göre farklı formüller kullanırız.

$$\text{Hedef kitledeki birey sayısı bilinmiyorsa: } n = \frac{t^2 * p * q}{d^2}$$

$$\text{Hedef kitledeki birey sayısı biliniyorsa: } n = \frac{N * t^2 * p * q}{d^2 * (N - 1) + t^2 * p * q}$$

N : Hedef kitledeki birey sayısı

n : Örneklem alınacak birey sayısı

p : İncelenen olayın görülüş sıklığı (gerçekleşme olasılığı)

q : İncelenen olayın görülmeyiş sıklığı (gerçekleşmeme olasılığı)

t : Belirli bir anlamlılık düzeyinde, t tablosuna göre bulunan teorik değer
d : Olayın görülüş sıklığına göre kabul edilen örnekleme hatasıdır (=hata marjı;
İng.: margin of error).

Hata marjı

Hata marjı, güven aralığı ile ilgilidir. Örneğin, hesapladığımız prevalans %19,6, %95 güven aralığı ise [18,0-21,2] olsun. Güven aralığını şu şekilde göstersek belki görsel olarak daha anlaşılır olabilir: [18,0_(alt sınır) - 19,6_(ortalama) - 21,2_(üst sınır)]. Bu durumda hata marjımız; 19,6-18,0 (veya 21,2-19,6), yani 1,6 olacaktır:



Görülme sıklığı yüzde 10'un altında olan olaylar için d prevalansın yarısı olarak alınabilir. Örn. Prevalans %1 ise d=0,005 olabilir. d'nin küçük tutulması daha ince ayrıntıları saptama eğiliminde olduğumuzu gösterir.

Bir bölgedeki kadınların doğum kontrol yöntemlerini kullanma durumlarını incelemek istiyoruz [Doğum kontrol yöntemi kullanma durumu (evet/hayır) kategorik]. Tahminlerimize göre doğum kontrol yöntemini kullanma durumu %20 olsun. Sonuçları %95 güvenilirlik aralığında (alfa = 0.05), d = 0.05 hata marjı içerebileceğini kabul ediyorsak, kaç kadına anket uygulamamız gerekir?

Verilerimizi formüle uygulayalım:

$$p = 0.20$$

$$q = 0.80 (1-0.20)$$

$$t = 1.96 \text{ (alfa} = 0.05 \text{ de serbestlik derecesine göre tablodan bulunmuştur.)}$$

$$d = 0.05 \text{ (alfa \%5 örnekleme hatasını kabul ettiğimiz için)}$$

$$n = (1.96^2) * (0.20 * 0.80) / (0.05^2) = \mathbf{246 \text{ kişi}}$$

Hedef toplumdaki evli kadın sayısının 500 olduğunu varsayarsak aynı değerler için örneklem büyüklüğü:

$$n = [(500) * (1.96^2) * (0.20 * 0.80)] / [(0.05^2 * 499) + (1.96^2) * (0.20 * 0.80)] = \mathbf{165 \text{ kişi}}$$

olacaktır.

Sonuç ölçütümüz nümerik ise:

$$n = (t^2 * s^2) / d^2$$

n : Örneklem alınacak birey sayısı

s : İncelenen nümerik değişkenin tahmini standart sapması

t : Belirli bir anlamlılık düzeyinde, t tablosuna göre bulunan teorik değer

d : Hata marjı (Bakınız Hata marjı, S. 385).

Bir bölgedeki bireylerin hemoglobin değerlerini araştırmak istiyoruz. Daha önce yapılan bir pilot çalışmada hemoglobin ortalaması 11, standart sapması (s) 4 olarak bulunmuştur. Bilinmeyen toplum ortalamasını %95 güvenirlilikle ± 0.5 'lik bir farkla (d) bulabilmek için kaç kişiden ölçüm yapmalıyız?

$$n = (1,96^2 * 4^2) / 0,5^2 = 246 \text{ kişi.}$$

Yukarıdaki örneklem hesaplarını yapmanız ve Alıştırma 18'teki örnek problemleri çözmeniz için hazırladığımız Excel dosyasını buradan indirebilirsiniz: <https://drive.google.com/file/d/0B0ieO-Fy3z-5VnVWWFI0RkNXYXc/view?usp=sharing>

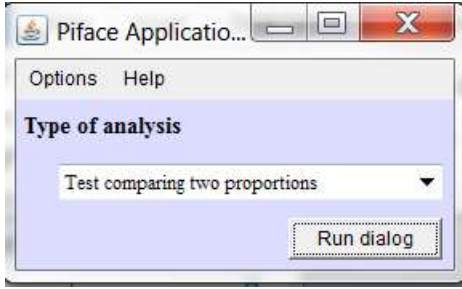
Hipotez testleri için örneklem büyüklüğünün hesaplanması

Diyabeti olan ve olmayan bireylerin sigara içme yüzdeleri arasında fark olup olmadığını araştırmak istiyoruz. Hipotez testi olarak Ki Kare analizi kullanacağız. Diyabetik bireylerde sigara içme yüzdesinin %30, diyabeti olmayan bireylerde ise %40 olduğunu tahmin ediyoruz. İki grup arasındaki %10 'luk bir farkı %5 yanılma payıyla ve %80 güçle saptayabilmek için kaç kişilik bir örneklem gerektiğini hesaplamak istiyoruz.

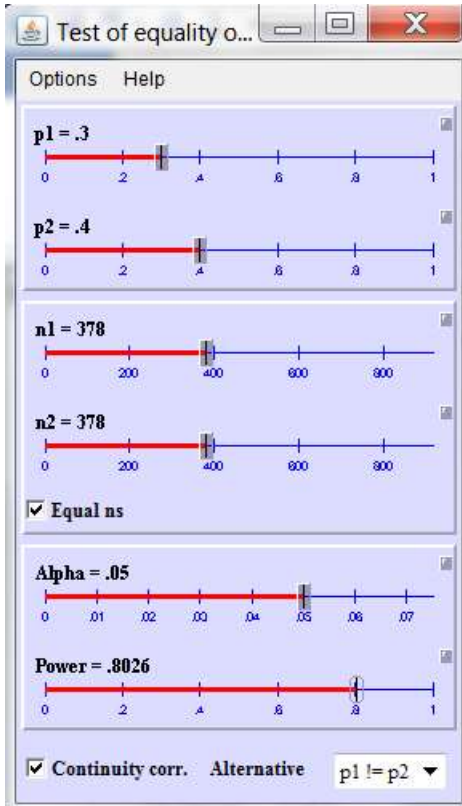
[Araştırmamızın çıktısı sigara içme durumu (evet/hayır) kategorik bir değişken. Diyabeti olan ve olmayan olmak üzere iki grubumuz var]

Russ Lenth'in Java modülünü kullanarak:

“Piface” Java modülünü açınız. “Type of analysis” den “Test comparing two proportions” seçiniz.



>Run dialog > Açılan pencereden p_1 'i 0.3'e, p_2 'yi 0.4'e ayarlayınız. n_1 ve n_2 'yi power=0.80 olana kadar artırarak ayarlayınız.



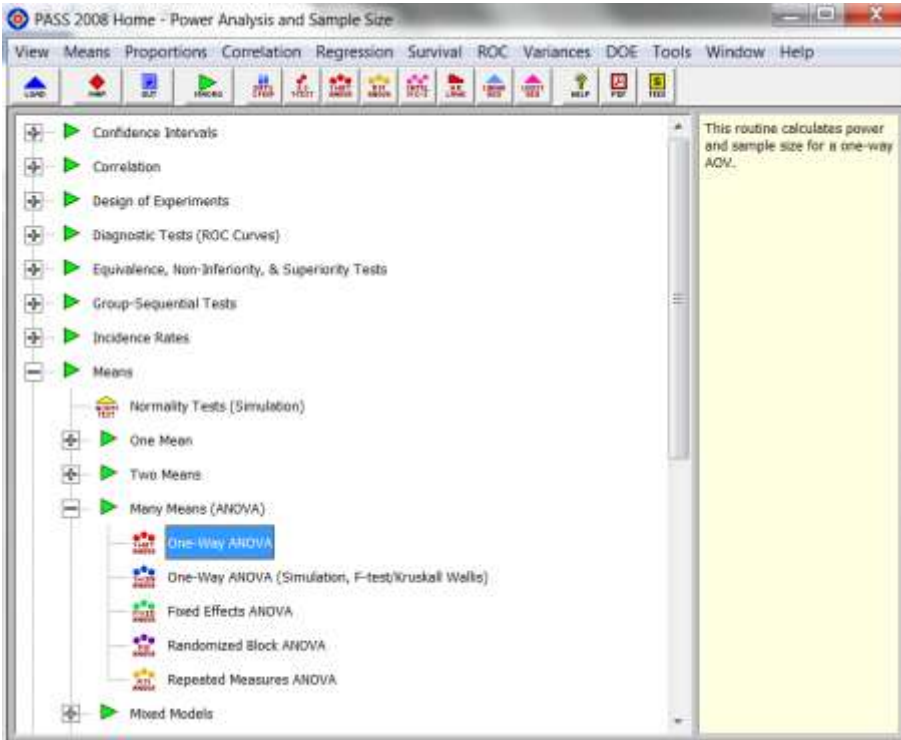
Görüldüğü gibi alfa = %5, iki yönlü p (alternatif hipotez = p_1 ile p_2 eşittir) ve her iki grupta eşit sayıda birey olacak şekilde gruplar arasındaki %10'lu bir farkı saptamak için %80'lik güç düzeyinde her iki grupta da 378'er kişilik bir örneklem gerekmektedir.

Diyabet hastalarımızın eğitim durumlarına göre kan şekeri ortalamalarının farklılık gösterip göstermediğini araştırmak istiyoruz.. Hipotez testi olarak tek yönlü varyans analizi (ANOVA) kullanacağız. Gruplar arasındaki farkı %5 yanlışma payı ve %80 güç ile gösterebilmek için araştırmanın her bir kolunda kaç bireyin olması gerektiğini hesaplamak istiyoruz.

[Araştırmanın çıktısı kan şekeri değerleri (nümerik). Eğitim durumlarına göre 4 grubumuz var (okuryazar / ilköğretim / üniversite / doktora)]

PASS yazılımını¹ kullanarak:

Means > Many Means (ANOVA) > One-Way ANOVA



[Açılan pencerede örnekleme etki edecek faktörleri belirtiniz (alfa yanlışma payı %5, grup sayısı 4, grupların tahmin edilen ortalamaları 90, 100, 110 ve 115 mg/dl, tahmini standart sapma 15)] > Run

¹ <http://www.ncss.com/pass.html>

Aşağıdaki çıktıyı elde ederiz:

One Way ANOVA Power Analysis								
Page/Date/Time 1 16.05.2010 10:22:05								
Numeric Results								
Power	Average	n	k	Total	N	Alpha	Beta	Effect Size
0.00000	1.00	4	4	4	4	0.05000	1.00000	9.60
0.14689	2.00	4	8	8	8	0.05000	0.85311	9.60
0.28306	3.00	4	12	12	12	0.05000	0.71694	9.60
0.42217	4.00	4	16	16	16	0.05000	0.57783	9.60
0.54999	5.00	4	20	20	20	0.05000	0.45001	9.60
0.65962	6.00	4	24	24	24	0.05000	0.34038	9.60
0.74895	7.00	4	28	28	28	0.05000	0.25105	9.60
0.81888	8.00	4	32	32	32	0.05000	0.18112	9.60
0.87185	9.00	4	36	36	36	0.05000	0.12815	9.60
0.91089	10.00	4	40	40	40	0.05000	0.08911	9.60
0.93899	11.00	4	44	44	44	0.05000	0.06101	9.60
0.95882	12.00	4	48	48	48	0.05000	0.04118	9.60
0.97255	13.00	4	52	52	52	0.05000	0.02745	9.60

Görüldüğü üzere %80'lik bir güce ulaşmak için araştırmanın her bir kolunda yaklaşık 8 kişi olması (toplam 32 kişi) yeterlidir.

Alıştırma 18

Örneklem ve araştırma tasarımı problemleri. Örneklem hesabı için Russ Lenth'in modülünü kullanabilirsiniz: <http://www.stat.uiowa.edu/~rlenth/Power/>.

1

Bulunduğunuz mahallede erişkin bireylerin kan basınçlarını araştırmak istiyorsunuz.

Toplum genelinin sistolik kan basıncı daha önce araştırılmış ve 135 ± 15 mmHg bulunmuş. Mahallenizdeki bireylerin sistolik kan basıncı ortalamalarının toplumunkile karşılaştırmak istiyorsunuz.

% 5 yanılma payıyla ortalamalar arasındaki 5 mmHg'lik bir farkı yakalayacak, %80 güçte bir araştırma planlamak için;

- H_0 Hipotezini belirtiniz.
- Hangi araştırma tasarımını uygularsınız?
- Hangi hipotez testini uygularsınız?
- Kaç kişilik örneklem gerekir?

2

Bulunduğunuz mahallede erişkin bireylerin kan basınçlarını araştırmak ve erkeklerle kadınların kan basınçlarını karşılaştırmak istiyorsunuz. Daha önce benzer bir çalışma yapılmış ve sistolik kan basıncı erkekler için 138 ± 12 mmHg, kadınlar için 129 ± 9 mmHg bulunmuş (varyansların homojen olduğunu varsayınız).

% 5 yanılma payıyla ortalamalar arasındaki 4 mmHg'lik bir farkı yakalayacak, %80 güçte bir araştırma planlamak için;

- H_0 Hipotezini belirtiniz.
- Hangi araştırma tasarımını uygularsınız?
- Hangi hipotez testini uygularsınız?
- Kaç kişilik örneklem gerekir?

3

Toplum genelinde şeker hastalığının %10 oranında görüldüğünü biliyorsunuz. Mahallenizdeki şeker hastalığı yaygınlığının toplumdakinden farklı olup olmadığını karşılaştırmak istiyorsunuz. % 5 yanılma payıyla ve toplum yüzdesinden %10'luk bir sapmayı yakalayacak, %80 güçte bir araştırma planlamak için;

- H_0 Hipotezini belirtiniz.
- Hangi araştırma tasarımını uygularsınız?
- Hangi hipotez testini uygularsınız?

d. Kaç kişilik örneklem gerekir?

4

Bulunduğunuz mahallede şeker hastalığı yaygınlığını araştırmak ve erkeklerle kadınlardaki durumu karşılaştırmak istiyorsunuz. Toplum genelinde şeker hastalığının erkeklerde %11 oranında, kadınlarda %8 oranında görüldüğünü biliyorsunuz. % 5 yanılma payıyla %80 güçte bir araştırma planlamak için;

- H_0 Hipotezini belirtiniz.
- Hangi araştırma tasarımını uygularsınız?
- Hangi hipotez testini uygularsınız?
- Kaç kişilik örneklem gerekir?

5

Hastanenizdeki dosya kayıtlarını inceleyerek hipertansiyon nedeniyle inme geçiren bireyleri araştırmak ve geçmişte iki farklı hipertansiyon ilacı kullananlardaki durumu karşılaştırmak istiyorsunuz. İlaçlar arasında inme oranını %1,5'ten %0,9'a düşürecek bir farkın olup olmadığıyla ilgileniyorsunuz.

- H_0 Hipotezini belirtiniz.
- Hangi araştırma tasarımını uygularsınız?
- Hangi hipotez testini uygularsınız?
- Kaç kişilik örneklem gerekir?

6

Kan şekerini en iyi kontrol altına alacak bir tedavi araştırıyorsunuz. Hastalarınızı rastgele gruplara ayırmak ve 4 farklı tedaviden birini vermek istiyorsunuz. Üç ay sonra gruplar arasında açlık kan şekeri ortalamalarını karşılaştırmak istiyorsunuz. Gruplar arasında 5 mg/dl düzeyinde bir ortalama kan şekeri farkı olmasını önemli kabul ediyorsunuz. Grupların kan şekeri ortalamalarının standart sapmalarını 10 olarak bekliyorsunuz.

İki yönlü % 5 yanılma payıyla %80 güçte bir araştırma planlamak için;

- H_0 Hipotezini belirtiniz.
- Hangi araştırma tasarımını uygularsınız?
- Hangi hipotez testini uygularsınız?
- Kaç kişilik örneklem gerekir?

Çözümler için bakınız s. 489

Bölüm 48: Duyarlılık ve Özgüllük (ROC) Analizi

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun duyarlılık ve özgüllük kavramları hakkında bilgi sahibi olması ve SPSS ile ROC analizi yapıp yorumlayabilmesi amaçlanmıştır

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Duyarlılık kavramını açıklayabilmek
- Özgüllük kavramını açıklayabilmek
- Olabilirlik oranı (likelihood ratio) kavramını tanımlayabilmek
- SPSS ile ROC analizi yapabilmek
- SPSS’te ROC analizi çıktısını yorumlayabilmek

Hastalarımızın kan şekeri kontrollerinin ne kadar kontrol altında olduğunu takip etmek için hemoglobin A1c (HbA1c) değerlerini ölçüyoruz. HbA1c değerlerine göre bireyleri 1-iyi kontrollü ve 0-KontROLSÜZ olarak sınıflandırdık (HbA1cGrup değişkeni). Diğer taraftan vücut kitle indeksi (VKİ) arttıkça bireylerin şeker kontrollerinin zorlaştığını biliyoruz.

Acaba VKİ değerlerine bakarak bireylerin kan şekerlerinin kontrol altında olup olmadığını tahmin edebilir miyiz? Böyle bir tahmin yapmak istesek hangi HbA1c değerini (kesme değeri) sınır olarak almalıyız?

Yukarıdaki senaryoyu “Öğrencilerin sınıfı geçip geçemeyeceklerini belirlemek için zeka seviyeleri sınırı kaç puan alınmalıdır?” veya “Yeni bir kan testinin bilinen bir altın standarda göre bir hastalığı saptamada kesme değeri, duyarlılık ve özgüllük değerleri kaçtır?” gibi farklı senaryolara da uyarlayabiliriz.

Tahmin etmede kullanacağımız bir nümerik değişken (=sonuç ölçütü=bağımlı değişken) (“VKİ”) ve olayın oluş durumunu bildiren ikili bir değişken (“HbA1cGrup”) olmalıdır.

Duyarlılık ve Özgüllük Kavramları

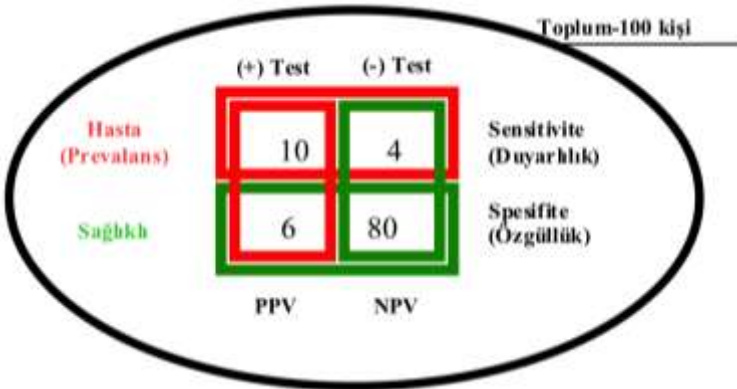
Aşağıda verilen şekli inceleyiniz. 100 kişilik bir toplumda 14 kişinin hasta, 86 kişinin ise sağlıklı olduğunu kabul edelim. Bir test uygulayarak bu bireylerin hasta mı sağlıklı mı olduklarını tahmin etmek istiyoruz. (Bizim örneğimizde HbA1c değerini kullanarak bir sınır değerin üzerinde ve altında olanları “test +” ve “test –“ kabul edeceğiz).

Hasta olduklarını altın standart bir yöntemle bir şekilde bildiğimiz bireylerden bazılarında testimiz pozitif sonuç verecek, bazılarında ise negatif olacaktır. İşte hasta olduğu bilinen bir grupta testin gerçek hastaları gösterebilme özelliğine testin “duyarlılığı” (=sensitivity) denir. Buradaki örnekte duyarlılık $10/(10+4)=0,71$ yani %71’dir.

Diğer taraftan, sağlam olduğu bilinen bireyler arasında testimizin gerçek sağlamları gösterebilme (yani negatif çıkma) özelliğine de testin “özellikliği” (=specifity) denir. Buradaki örnekte özgülük $80/(80+6)=0,93$ yani %93’tür.

PPV ve NPV Kavramları

Testlerin duyarlılık (gerçek pozitiflik) ve özgülük (gerçek negatiflik) özellikleri ilgili testin teknik özellikleridir ve her zaman aynıdır. Bununla birlikte, testin kullanılacağı klinik ortama göre değerini belirten pozitif tahmin ettirici değer (positive predictive value=PPV) ve negatif tahmin ettirici değer (negative predictive value=NPV) testin uygulandığı toplumdaki hastalığın yaygınlığına göre değişebilir. Verilen örnekte $PPV=10/(10+6)=\%63$ ve $NPV=80/(80+4)=\%95$ ’tir. Dikkat edilirse, aynı test duyarlılık ve özgülük değerleri aynı kalacak şekilde hastalığın 14 kişide değil de daha fazla bireyde görüldüğü bir topluma uygulanırsa PPV değeri artacaktır. Bu konu ROC analizi ile direkt ilişkili olmamakla birlikte, örn. Beyin tomografinin hastanenin nöroloji kliniğindeki değeri çok yüksek iken (nöroloji kliniğine başvuranlar içerisinde beyin tümörlü hasta sayısı nispeten daha fazla olacaktır) aynı testin bir mahalledeki aile hekimliği polikliniğinde neredeyse hiç değerinin olmamasını açıklamak açısından önemlidir.

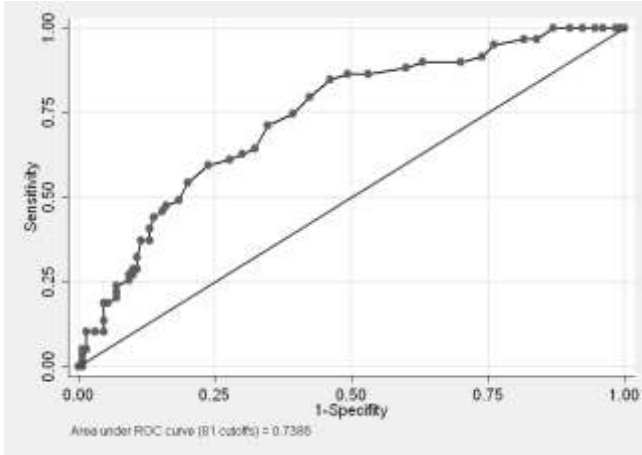


Şekil 1: Duyarlılık ve özgülük kavramlarının şematik olarak gösterilmesi.

ROC Analizi

Receiver Operating Characteristics (ROC) eğrisi testin duyarlılık değerini 1-özgüllük değerinin (yalancı pozitiflik) grafiksel olarak gösterilmesidir. ROC eğrisi Relative Operating Characteristics eğrisi olarak ta bilinir. Bu ismin verilmesinin nedeni, gerçek pozitiflik ile yalancı pozitiflik değerlerini rölatif (nispi) olarak karşılaştırmasındandır. ROC eğrisinin değerlendirilmesinde önemli olan eğrinin altında kalan alandır (Area Under the Curve; AUC). AUC, testimizin rastgele seçilen bir pozitif durumda bireyi rastgele seçilen bir negatif duruma göre doğru sınıflandırabilme ihtimalidir.

Dolayısıyla, ROC eğrisinin altında kalan alan 1'e ne kadar yakınsa testimizin doğru tahmin değeri o kadar yüksek olacaktır. Tersine %50'ye yakın olması ise değerinin düşük olduğunu gösterecektir. Aşağıdaki ROC eğrisinde %73'lük bir alan bulunmaktadır.



Hangi Kesme Değeri?

Görüldüğü gibi, ROC eğrisinde duyarlılık ile 1-özgüllük değerlerinin çeşitli kombinasyonlarına gelen noktalar söz konusudur. Testimizin hem duyarlılık, hem de özgüllüğünün yüksek olmasını istediğimizden seçeceğimiz kesme değeri en yüksek duyarlılık değerine karşılık en düşük 1- özgüllük noktası olmalıdır. Bu noktayı bulmanın bir yolu **olabilirlik oranına** (likelihood ratio; LR+) bakmaktır.

Olabilirlik oranı en yüksek çıkan koordinatları kesme noktası olarak alabiliriz.

Olabilirlik oranları için aşağıdaki sınıflama önerilmektedir¹ :

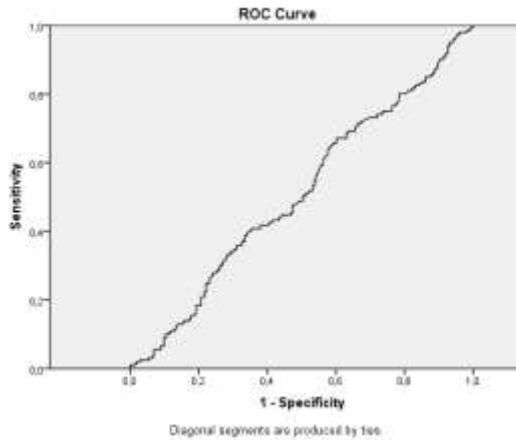
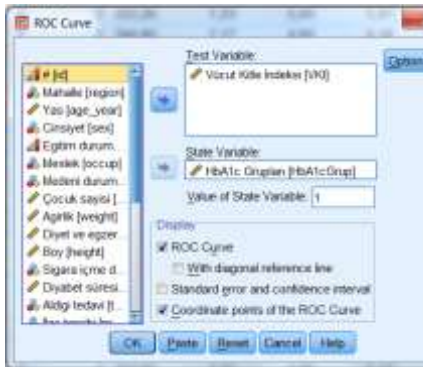
LR+ = 10	→ Mükemmel
LR+ = 6	→ Çok iyi
LR+ = 2	→ İyi
LR+ = 1	→ İşe yaramaz

Örnek uygulama

Diyabet.sav SPSS verisetini kullanarak konunun başında verdiğimiz senaryoyu çözelim. Verisetimizde “VKİ” değişkenini kullanarak “HbA1cGrup” değişkeninde bireylerin durumunu belirlemeye çalışacağız. Sonuçta VKİ için en uygun kesme değerini bulmayı hedefliyoruz.

Analyze>ROC Curve [“VKİ” değişkenini “Test Variable” alanına, “HbA1cGrup” değişkenini “State Variable” alanına geçirelim; “Value of State Variable” kutusuna 1 yazalım (1=HbA1c kontrol durumu bozuk olanlar olarak 0= HbA1c kontrol durumu iyi olanlar olarak kodlandı). “Coordinate Points of the ROC Curve” kutucuğunu işaretleyelim]>ok

Aşağıdaki çıktıları elde ederiz:



¹ Mayer D. Essential Evidence-based Medicine. Cambridge University Press. 2004.

Area Under the Curve	
Test Result Variable (s):VKI	
Area	
	,510

The test result variable(s): VKI has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

Coordinates of the Curve			
Test Result Variable(s):VKI			
Positive if Greater Than or Equal To ^a	Sensitivity	1 - Specificity	
13,8729	1,000	1,000	
14,9693	,995	1,000	
16,0955	,995	,995	
17,5756	,990	,995	
18,9549	,990	,990	
19,8901	,985	,990	
20,0492	,985	,985	
20,3857	,985	,980	
20,6045	,980	,980	
20,7303	,980	,975	
20,8255	,980	,970	
20,9228	,980	,966	
21,0327	,980	,961	
21,0718	,975	,961	
21,0950	,975	,956	
21,1620	,970	,956	
21,3033	,970	,951	
21,4239	,965	,951	
21,4768	,960	,946	
21,4944	,960	,941	
21,5586	,955	,941	
21,6462	,950	,941	
21,7673	,950	,936	
21,8563	,950	,931	
21,9447	,945	,931	
22,1105	,940	,931	
22,2664	,940	,926	
22,3718	,935	,926	
22,4162	,930	,926	

(sığmadığından tablonun tamamı alınmamıştır)

Şimdi elde ettiğimiz koordinatlar tablosundan LR+ için en büyük değerin hangi kesme sınırında olduğunu bulabiliriz. Bunun için tabloyu kopyalayıp Microsoft Excel sayfasına yapıştırabilir ve $LR+ = \text{duyarlılık} / (1 - \text{özellik})$ formülünden hesaplama yapabilir veya LR+ değerini elle hesaplayabiliriz.

Test Result Variable(s):VKI			
Positive if Greater Than or Equal To ^a	Sensitivity	1 - Specificity	LR+
13,8729	1,000	1,000	1,00
14,9693	,995	1,000	1,00
16,0955	,995	,995	1,00
17,5756	,990	,995	0,99

...

44,5264	,015	,020	0,76
44,8964	,015	,015	1,01
47,1311	,010	,015	0,67
49,3186	,010	,010	1,01
51,2502	,010	,005	2,02
53,2005	,010	,000	#SAYI/0!
55,4939	,005	,000	#SAYI/0!
58,5279	,000	,000	#SAYI/0!

Hesaplamaya baktığımızda en yüksek LR+ değerinin 51.25’e denk gelen VKİ sınırında olduğunu görüyoruz. Bulgularımızı yorumlayacak olursak:

1. ROC eğrisinin altında kalan alan 0,51 olup bu değer oldukça düşüktür.
2. En yüksek LR+ değerini veren VKİ kesme noktası 51,25’tir.
3. En yüksek LR+ değeri 2,02’dir.
4. Kesme noktası olarak 51,25 kullanıldığında HbA1c kontrol durumunu %1 duyarlılık ve %99,5 özgüllükle saptayabiliriz.
5. Sonuç olarak, ROC eğrisi altında kalan alanın düşük olması, LR+ değerinin “zayıf” olması, testimizin duyarlılığının çok düşük olması gibi nedenlerle kan şekerinin kontrol altında olma durumunu saptamada vücut kitle endeksinin kullanılabilir bir değişken olmadığı yorumunu yapabiliriz.

Bölüm 49: Geçerlilik ve Güvenilirlik Analizi

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun bir geçerlilik ve güvenilirlik analizinin hangi durumlarda yapılacağı ve hangi aşamalardan oluştuğu hakkında bilgi sahibi olması amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Geçerlilik ve güvenilirlik analizinin öneminin kavranması
- Güvenilirlik çeşitlerini açıklayabilme
 - Paralel formlar
 - Uygulayıcılar arası (Interrater)
 - Test/tekrar test (Test/re test)
 - Yarıya bölme (Split half)
 - Kuder-Richardson
 - Cronbach alfa
- Geçerlilik çeşitlerini açıklayabilme
 - İçerik geçerliliği (Content validity)
 - Tahmin ettirici geçerlilik (Predictive validity)
 - Yapı geçerliliği (Construct validity)
 - Eş zamanlı geçerlilik (Concurrent validity)
 - Görünüm geçerliliği (Face validity)

Senaryo 1: Temel İstanbul'a gitmiş. Otobüste karşılaştıkları kişi kendisine "Pardon beyefendi, losyonunuz ithal mi?" demiş. Temel onu vurmuş. Sorulunca da "Ne dedüğünü anlamadım; her ihtimale karşı furdum oni" demiş.

Senaryo 2: Temel İstanbul'a gitmiş. Otobüste karşılaştıkları kişi kendisine "Af edersiniz beyefendi, hangi marka kolonya kullanıyorsunuz?" demiş. Temel "Rize çay kolonyası" demiş.

Bir testin söz konusu bir durumu ölçebilmesi için (a) bahse konu edilen durum var olmalıdır ve (b) ölçümü hedeflenen durumdaki değişimler ölçüm sonuçlarını da değiştirmelidir. Tıbbi araştırmalarda sıklıkla araştırma ölçeklerine başvururuz. Yaşam kalitesi için SF-36, ağrı şiddeti için WOMAC, depresyon için Beck depresyon ölçeği, benlik saygısı için Rosenberg kullanıldığı gibi, daha birçok sağlık durumunu değerlendirmek için ölçekleri kullanırız.

Aslında sadece psikometrik ölçümler için değil, her çeşit ölçümde geçerlilik ve güvenilirlik kavramlarını gündeme getirebiliriz. Örneğin biyokimyasal analiz yapan cihazların da doğru ölçüp ölçmediklerinden ve her defasında aynı sonucu vermelerinden emin olmak gerekir.

Bu bölümde geçerlilik ve güvenilirlik kavramlarının örneklendirilerek açıklanması amaçlanmıştır.

Geçerlilik ve Güvenilirlik Kavramları

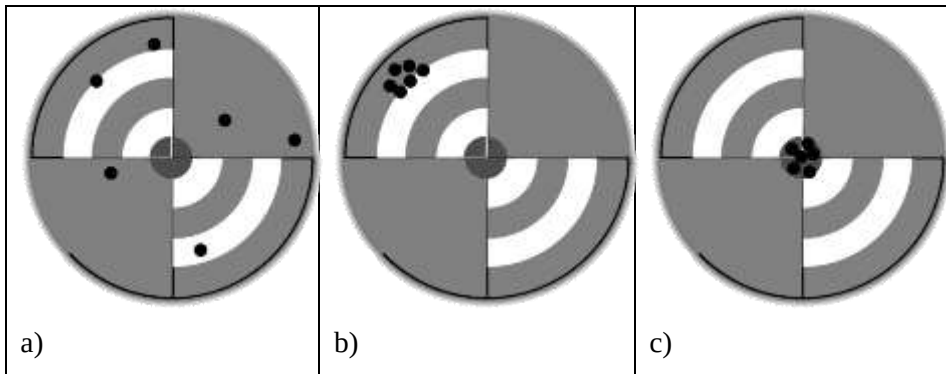
Özellikle sözlü sınavlarda şikâyet edilen bir durum vardır: eğitici iyi bir sınav yaptığını düşünür ama öğrenciler aynı görüşte olmayabilir. Sınavın eğitimin içeriğini ne kadar ölçtüğü sorgulanabileceği gibi, aynı sınava tekrar girilse geçme/kalma durumunun çok farklı olacağı da iddia edilebilir.

Bu durumda alt ekstremita kemiklerinin anlatıldığı anatomi dersinin sınavında birkaç çeşit eğitici den bahsedebiliriz. Birinci eğitici tipi sınavda hem konuyla alakasız sorular sorar, hem de her öğrenciye farklı içerikte sorular sorar. Hedefi tuturmaktan uzak olan ve hem de her ölçümde farklı sonuçlar veren bu ölçüme “Hem geçersiz hem de güvenilmez” ölçüm denilebilir (Şekil 1a).

İkinci eğitici tipi ise yine anlatılan içerikle ilgili sorular sormaz ama tüm öğrencilere örneğin üst ekstremita kemikleri ile ilgili sorular sorar. Bu eğitici için ölçümü için “Güvenilir ama geçersiz” bir ölçüm diyebiliriz (Şekil 1b).

Üçüncü eğitici tipi ise sorularını alt ekstremita kemikleri hakkında hazırlamıştır ve tüm öğrencilere de aynı içerikte sorular sormaktadır. Bu eğitici için ölçümü için ise “Hem güvenilir, hem de ama geçerli” bir ölçüm diyebiliriz (Şekil 1c).

Şüphesiz her ölçüm aracından hem geçerli, hem de güvenilir olması beklenir. Ancak, tamamen tutarsız bir eğiticiyense, en azından yanlış yerden de sınav yapsa hep aynı ters köşeye yatıran eğiticiyi tercih edileceği gibi, güvenilir olmayan bir ölçek için geçerliliğin bir değerinin olmayacağını söyleyebiliriz.



Şekil 1: a) Geçersiz ve güvenilmez ölçüm, b) Geçersiz ama güvenilir ölçüm, c) Geçerli ve güvenilir ölçüm.

Güvenilirlik Çeşitleri

Bir ölçeğin güvenilirliğinden bahsedince akla stabilliği (Aynı örnekleme yapılan tekrarlayan ölçümlerden aynı sonucun alınması halinde ölçeğin stabil olduğu söylenir.), eşdeğerliliği ve homojenliği gelir. Homojenlik açısından güvenilirlik ölçeğin iç özelliğiyle ilgilidir (iç tutarlılık da “internal consistency” denir). Aynı yapıyı ölçen maddelerin ne kadar benzer sonuçlar verdiğinin ölçülmesidir. Yani aynı özelliği ölçmede farklı maddelerin ne kadar tutarlı olduğunun incelenmesidir. Bu kavramları açacak olursak aşağıdaki güvenilirlik çeşitlerinden bahsedebiliriz:

1. Paralel formlar

Paralel form güvenilirliğine bakmak için aynı kişilere iki farklı ölçek uygulanıp aralarındaki korelasyona bakılabilir. Korelasyon ne kadar yüksekse ölçeklerin o kadar eşdeğer olduğunu söyleriz. Ölçülen aynı şeydir. Sadece ifade tarzları ve/veya soruların tasarımı farklıdır. Bunu yapmanın bir yolu, soruları hazırlamak ve rastgele ikiye ayırarak uygulamaktır. Paralel formlar uygulaması ölçeğin STABİLLİĞİNİ de ölçer.

2. Test/tekrar test (Test/re test)

Aynı araştırma aracı aynı deneklere benzer şartlar altında 2 veya daha fazla kez uygulanır. Burada iki ölçüm arasında fazla bir farklılık olmamalıdır. Tabii ki, ölçüm yapılan zaman aralığının ne kadar olduğu da önemlidir. Uygulama ne kadar erken tekrarlanırsa o kadar benzer sonuçlar elde edilir. Diğer taraftan çok geciktirilmesi ölçülen durumun değişmesine yol açabilir. 2-4 hafta gibi zaman aralıkları genel anlamda uygun kabul edilse de hafıza faktörü, bireyin konuya duyarlaşması ve zaman içerisinde oluşabilecek değişiklikler de dikkate alınarak tekrar test yapılmalıdır.

3. Gözlemciler arası güvenilirlik

Uygulayıcılar arası (Interrater) güvenilirliği tek bir formun iki uygulayıcı tarafından uygulanması ve aralarındaki korelasyona bakılması ile ölçülür.

Ölçek kategorik bir ölçüm yapıyorsa (evet/hayır gibi) iki araştırmacının uygulamasında ne kadar uyum olduğuna bakılır. Aralarındaki uyum (örn. %82) rapor edilir. Ölçek nümerik bir ölçüm yapıyorsa iki araştırmacının uygulamasının ne kadar korelasyon gösterdiğine bakılır. SPSS ile Phi (basit korelasyon), Kappa (rastlantı açısından düzeltme yapılmış) ve Kendall's tau (sıralı veriler için) katsayıları hesaplanabilir.

4. Yarıya bölme (Split half)

Bir özelliği ölçmek için kullanılan tüm maddeler rastgele ikiye ayrılır. Ölçek bir grup bireye uygulanır ve her iki yarımın puanları hesaplanır. Bu iki yarımın karşılaştırılmasıyla güvenilirliğin derecesi belirlenir.

5. Cronbach alfa

Cronbach alfa yarıya bölmenin (matematiksel anlamda) eşdeğeridir. Güvenilirlik hesaplarında sıkça kullanılan bir katsayıdır. Maddeler arası korelasyon ortalamasını da dikkate alarak iç güvenilirliği hesaplar.

Cronbach alfa hesaplamasında ölçek maddeleri rastgele ikiye ayrılarak karşılaştırılır. Bu rastgele ikiye ayırma işlemi tüm ihtimaller için tekrarlanır. Benzer bir ölçüm de Kuder-Richardson'dur.

Özetleyecek olursak, güvenilirlik açısından paralel formlar ve uygulayıcılar arası güvenilirlik testin eşdeğerliliğini, yine uygulayıcılar arası güvenilirlik ve test/tekrar test testin stabilliğini, yarıya bölme, Kuder-Richardson ve Cronbach alfa gibi ölçümler ise homojenliğini belirler.

Geçerlilik Çeşitleri

Geçerlilik açısından içerik geçerliliği (Content validity), tahmin ettirici geçerlilik (Predictive validity), yapı geçerliliği (Construct validity), eş zamanlı geçerlilik (Concurrent validity) ve Görünüm geçerliliğinden (Face validity) bahsedilebilir.

1. İçerik geçerliliği

İçerik geçerliğinden anlaşılan ölçeğin içeriğinin gerçekten ölçülmesi hedeflenen fenomenle ilgili olup olmamasıdır. Depresyonu taramak için oluşturduğumuz bir ölçekte keyifsizlik, suçluluk hissi, intihar düşüncesi gibi maddeler bekleriz; ağrıyla ilgili soruların olması içerik açısından geçersiz olduğunu düşündürür.

İçeriğin boyutlarının belirlenmesini belki ölçek geliştirmedeki en zor kısımdır. Bu amaçla bir uzmanlar grubundan yararlanılabilir ve literatür desteği gerekir.

2. Tahmin ettirici (=criterion=ölçüt) geçerlilik

Araştırma aracının gerçek yaşamda durumları ne kadar tahmin ettirici olduğuyla ilgilidir. Depresyon ölçeğinde intihar riski saptananların ne kadarı intihar ediyor? Ya da trafik sınavında yüksek puan alanlar trafikte ne kadar iyi araç kullanıyor?

3. Yapı geçerliliği

Aracın ölçülmeye çalışılan teorik psiko sosyal yapı ile ne kadar korelasyon gösterdiği ile ilgilidir. “Bu ölçek ölçmeye çalıştığımız fenomeni ne kadar ölçüyor?” sorusuna cevap aranmasıdır. Altta yatan fenomenle ilgili farklı konseptleri ölçmeye çalışır. Bu amaçla madde analizi yapılabilir.

4. Eşzamanlı geçerlilik

Eş zamanlı geçerliliği test etmek için ölçek aynı veya ilişkili bir yapıyı inceleyen ve daha önce geçerliliği ispat edilmiş başka bir ölçekle eşzamanlı olarak uygulanır. Bu da tahmin ettirici geçerlilik gibi bir ölçütü tahmin etmeye ne kadar yaradığını gösterir. Yeni geliştirilen depresyon ölçeğinin Beck depresyon ölçeği ile birlikte uygulanmasını örnek olarak verebiliriz.

5. Görünüm geçerliliği

Bir arabanın hızının dış görünüşünden tahmin edilmesi gibidir. Maddelerin görünüşü, okunabilirliği, uygulama kolaylığı gibi konular açısından değerlendirme yapılır. Bu amaçla Kutu 1'deki soruların sorulması faydalı olabilir.

1. Anket hakkındaki genel görüşleriniz nelerdir?
2. Açıklamalarla ilgili görüşleriniz nelerdir?
3. Bu anketi doldurmanız ne kadar sürdü?
4. Soruların sayısı kabul edilebilir mi?
5. Soruların sırası mantıklı mı?
6. Cevap vermede zorlandığınız veya anlayamadığınız sorular oldu mu?
7. Genel olarak anketin anlaşılabilirliği ve sadeliği nasıldır?
8. Soruların derecelendirmesiyle ilgili problem yaşadınız mı?
9. Soruların, yazım şekli, ifadesi veya derecelendirilmesiyle ilgili öneriniz var mı?
10. Anketle ilgili herhangi başka bir öneriniz var mı?
11. Anketi doldurmak için yardıma ihtiyaç duydunuz mu? Kim yardım etti?

Kutu 1: Bir ölçeğin görünüm geçerliliğini değerlendirmek için katılımcılara sorulabilecek sorular.

6. Goodness of fit – Likelihood ratio

Regresyon analizi gibi modelleme yapılan araştırmalarda modelin geçerliliğini incelemek için Goodness of fit veya Likelihood ratio gibi değerler hesaplanabilir. -2 Log likelihood ne kadar büyükse modelin o kadar uyumsuz olduğu düşünülür.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	147,550 ^a	,115	,279

a. Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.

“Bilim ölçüm demektir. Yanlış ölçümler yapmak kanıtlarımızın ve tedavilerimizin güvenilir olmasına yol açar” Bu nedenle ölçüm araçlarımızın geçerli ve güvenilir sonuçlar vermesine azami önem göstermeliyiz.

Nasıl bir tansiyon aletinin veya laboratuvarda kullandığımız bir cihazın kalibre edilmiş olması gerekiyorsa, psikometrik ölçümler için kullandığımız ölçeklerin de geçerli ve güvenilir olması önemlidir. Araştırmalarda kullanacağımız

uluslararası ölçeklerin Türkçe geçerliliğinin olup olmamasının yanında kendi hazırladığımız sınav ve soruların da geçerli ve güvenilir olmasına dikkat etmeliyiz.

Konuyu pekiştirmek için aşağıdaki örnek makalelere bakmanızı öneriyoruz:

1. Arifoğlu B, Richard NŞ, Razi G, Öz F. Çocuklar İçin Boşanmaya Uyum Envanteri Geçerlik ve Güvenirlik Çalışması. Çocuk ve Gençlik Ruh Sağlığı Dergisi 2010;17(3):121-127.
<http://www.scopemed.org/?ift=46&ft=46-1359639224>
2. Yıldırım F, İlhan İÖ. Genel Özyeterlilik Ölçeği Türkçe Formunun Geçerlilik ve Güvenirlik Çalışması. Türk Psikiyatri Dergisi 2010;21(4):301-8.
<http://www.turkpsikiyatri.com/PDF/C21S4/301-308.pdf>
3. Montero J, López JF, Vicente MP ve ark. Comparative validity of the OIDP and OHIP-14 in describing the impact of oral health on quality of life in a cross-sectional study performed in Spanish adults. Med Oral Patol Oral Cir Bucal. 2011;16(6):816-21.
http://www.medicinaoral.com/pubmed/medoralv16_i6_p816.pdf
4. Zhu L, Yu S, Xu T ve ark. Chinese validation of the Pelvic Floor Impact Questionnaire Short Form. Menopause. 2011;18(9):1030-3.

Bölüm 50: Tanı Testleri ve Sağlık Hizmetlerindeki Değeri

Zekeriya AKTÜRK

Amaç: Bu konu sonunda okuyucunun tanı testlerinin sağlık hizmetlerindeki yeri hakkında özellikle tanı testlerinin teknik ve klinik hassasiyetleri, birden çok tanı testi istenmesinin riskleri ve Bayes teoremi kapsamında karar verme stratejimiz olmak üzere bilgi sahibi olması amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Tanı testlerinin teknik ve klinik özelliklerini açıklayabilme
- Duyarlılık, özgüllük, pozitif tahmin ettirme değeri, negatif tahmin ettirme değeri ve olabilirlik oranlarını açıklayabilme
- Birden fazla tanı testi istendiğinde yanlış sonuç olasılığının nasıl artacağını açıklayabilme
- Gereksiz tanı testi istememe konusunda tutum kazanma
- Bayes yaklaşımının tıpta karar vermedeki etkisini açıklayabilme
- Ön test olasılığı ve olabilirlik oranını kullanarak post test olasılığı hesaplayabilme

Senaryo 1: Hastanenizde doğum yapan Yasemin Hanım'ın Göktuğ isminde bir bebeği dünyaya geldi. Doğumhanedeki asistan, bebekten rutin tarama testleri isterken musküler distrofiyi de taramak için CPK (Kreatinin Fosfo Kinaz) bakılmasını da istedi. CPK sonucu pozitif gelmesi durumunda aileye ne diyeceksiniz?

Senaryo 2: Çoğu hastanede rutin uygulama olarak yeni yatan hastalardan bazı tetkikler istenmektedir. Aile hekimliği asistanı Pınar Hanım da iç hastalıkları rotasyonu sırasında yeni yatan Hasan Bey'den serumda şu testlerin yapılmasını istiyor: Tam kan sayımı, AKŞ, T. Bil, D. Bil, AST, ALT, ALP, LDH, Na, K, Cl. Bu testlerden en az birinin yalancı pozitif veya yalancı negatif olma ihtimali yaklaşık ne kadardır?

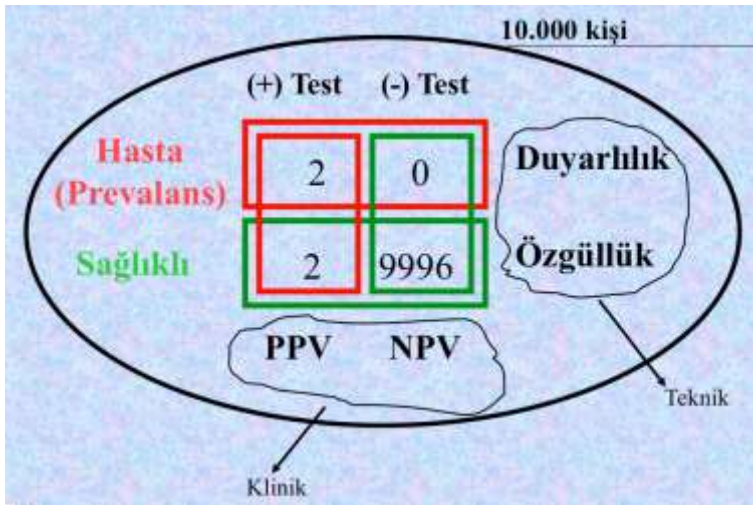
Senaryo 3: Dokuz yaşındaki Merve ateş ve burun akıntısı şikayeti ile polikliniğinize getirildi. Muayenede timpanik membranı göremediniz. Ancak, anne baba çocuğun kulağını çekiştirdiğini ve daha önceki çocuklarında da benzer şikayetler olduğunda otit teşhis edildiğini belirtiyorlar. Aksiller ateşi 38.5 °C olarak ölçtünüz. Otit olasılığı nedir?

Yukarıdaki her iki senaryo da günlük sağlık hizmeti sunumunda sıkça karşılaşılan durumlardır. Dolayısıyla tanı testlerinin kullanım yerlerine göre değerlerinin iyi bilinmesi, tanı testlerinin kişilere zarar verebileceğinin anlaşılması ve gereksiz test istenmemesi önem arz etmektedir. Bu bölümde tanı testlerinin günlük pratikteki kullanımlarıyla ilgili bilgiler verilecektir.

Tanı Testlerinin Özellikleri

Genelde tanı testleri denince akla testlerin duyarlılık (sensitivity) ve özgüllükleri (specificity) gelir (Tablo 1). Duyarlılık, “Hasta olduğu kesin bir yöntemle (altın standart) gösterilmiş olan kişilerde test uygulandığında pozitif sonuç verme ihtimali”, özgüllük ise “Sağlıklı olduğu kesin bir yöntemle (altın standart) gösterilmiş olan kişilerde test uygulandığında negatif sonuç verme ihtimali” olarak tanımlanabilir.

Testin “teknik özellikleri” olarak da bilinen duyarlılık ve özgüllük (Bakınız S. 393) bir test için hep aynıdır; toplumdan topluma değişmez. Oysa tanı testlerinin değeri, uygulanan popülasyona göre değişebilmektedir. Nöroloji kliniğine başvuran baş ağrılı hastalara tomografi istenmesi doğal karşılanırken aile hekimliği polikliniğine başvuran hastalara tomografi istenmesi sıradan bir durum değildir. Aynı şekilde, endokrinoloji polikliniğine başvuran hastaların kan şekere bakılması rutin bir uygulama olabileceği halde aile hekimliğinde günümüzde böyle bir uygulama yoktur.



Şekil 1: Testlerin klinik ve teknik özellikleri.

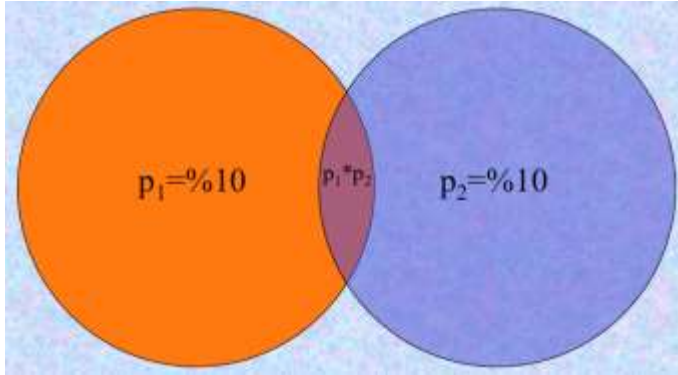
Pozitif tahmin ettirici değer, “Hasta veya sağlıklı olduğu bilinmeyen bir toplulukta uygulandığında testin hastaları gösterme ihtimali” olarak tanımlanabilir (Bakınız S. 394). PPV ve NPV değerleri testin “klinik özellikleri” olarak bilinmektedir. Testlerin farklı popülasyonlardaki değerlerinin değişmesi hastalığın o toplumdaki yaygınlığı ile ilgilidir. Tablo 1 incelendiğinde hastalığın toplumdaki yaygınlığı arttıkça testin PPV değerinin de artacağı görülecektir.

Bu farklı uygulamaların bilinmesi, farklı sağlık uygulayıcılarının birbirini daha iyi anlamasını da sağlayacaktır.

Bu bölümün başında verilen Senaryo 1'e dönecek olursak; CPK için duyarlılık %100, özgüllük ise %99,98 olarak bildirilmiştir. Musküler distrofinin toplumdaki prevalansı 10 binde 2'dir. Şekil 1'den görüleceği üzere CPK testi toplumdaki 2 hastanın ikisini de doğru olarak saptayacak, ancak, hasta olmayan 2 kişiye de yanlışlıkla hasta etiketi yapıştıracaktır. Testin PPV değeri hesaplandığında $2/4 = \%50$ olduğu görülmektedir. Dolayısıyla Göktuğ bebeğin hasta olma olasılığı %50'dir. Şimdi aileye ne diyeceğiz? Muhtemelen ek muayene ve anamnez aracılığıyla bilgiler toplayacağız; belki de kas biyopsisi yapılması gerekecek. Tabii ki, bu süreç boyunca aile madden ve manen stres altında olacaktır.

Fazla Test İstenmesinin Sonuçları

Birden fazla test istendiğinde testlerin birinden birinin hatalı olma olasılığı artacaktır. Araştırmalar laboratuvar testlerinde %0,1 ile %10 arasında hata paylarının olduğunu göstermektedir.¹ Örneğin, bir testin hatalı çıkma ihtimali $p_1=\%10$, ikinci bir testinki ise $p_2=\%10$ olsa p_1 **veya** p_2 testinin hatalı olma olasılığı $= (p_1+p_2)-(p_1*p_2) = (0,10+0,10)-(0,10*0,10) = \%19$ olacaktır (Şekil 2).



Şekil 2: Basit olasılık kuralları.

¹ http://www.medscape.com/viewarticle/758467_3.

<http://www.biochemia-medica.com/content/quality-laboratory-diagnostics-theory-practice>

<http://www.biochemia-medica.com/system/files/Sonntag%20Q.-%20Quality%20in%20analytical%20phase.pdf>.

Gereksiz test istemek hastalar için zarar verici olabilmektedir. Örneğin, gereksiz yere istenen bir Prostat Spesifik Antijen (PSA) testi sonrasında prostat biyopsisine kadar giden, hastalar için stres dolu ve zorlu bir süreç başlayabilir. Görüldüğü üzere test sayısı arttıkça hata olasılığı ciddi bir şekilde artmaktadır. Bu nedenle sağlık çalışanları gereksiz test istememe konusunda azami dikkatli olmalıdırlar.

Bayes Teoremi ve Tanı Testleri

Günlük pratiğimizde karar verirken sürekli ipuçlarını ve olasılıkları üst üste koyarız. Bu karar verme süreci çoğu kez bilinç dışımızda cereyan eder. Bir otomobil almak istesek öncelikle dikkat edeceğimiz özellikler olacaktır. Belki fiyatı en önemli kriter olacaktır. Fiyatı istediğimiz aralıkta olan bir model belirledikten sonra sıra diğer özelliklerine gelecek. Örneğin, motor gücüne, bagaj hacmine, koltuk sayısına bakacağız. Karar vermemizi etkileyen faktörler arasında aracın cant kalınlığı ve lastik markası belki hiç önemli olmayacak. Her bir faktörü gözden geçirip zihnimizdeki aracı alma olasılığını buna göre değiştireceğiz. Ta ki, kabul edebileceğimiz bir orana ulaşınca kadar. O zaman karar verip ortomobilimizi alırız... İşte sağlık hizmetlerinde de karar süreci benzer şekilde işlemektedir. Hastayla görüşmesi sırasında hekimin anamnez alması, laboratuvar bulguları ve muayene sonuçları hep yeni ipuçları sağlamaktadır. Bazı ipuçları ötekilerden daha değerlidir; hekimi tanıya yaklaştırır. Bazıları da tanıdan uzaklaştırıcı etki gösterir. Sonuçta hekim yeterli veriye ulaştığına kanaat getirdiğinde kararını verir. “Yeterli” tanımı kişiden kişiye değişebilmektedir. Belirsizliğe toleransı fazla olan hekimler daha düşük yüzdelerle tanı koyarken, bazı hekimler %90-100 emin olmadan tanı koymazlar. 1701-1761 yılları arasında yaşamış bir presbyterian kilisesi görevlisi görevlisi olan Thomas Bayes, insanlardaki bu karar verme sürecini formülize etmiştir. Bayes’in teoremi ölümünden sonra yayınlanmıştır¹.

Bayes yaklaşımından hareketle iki olasılığı birbirine eklemek için şu formül geliştirilmiştir:

$$Posttest Odds = Pretest Odds \times Olabilirlik Oranı$$

Olabilirlik oranı (LR), bir testin duyarlılık ve özgüllüğünü birarada değerlendiren bir ölçüt olup şu şekilde hesaplanmaktadır:

¹ https://en.wikipedia.org/wiki/Thomas_Bayes

$$\text{Olabilirlik Oranı} = \frac{\text{Duyarlılık}}{(1 - \text{Özgüllük})}$$

Bir tanı testinin olabilirlik oranı ne kadar büyükse test o kadar değerlidir. Olabilirlik oranı için aşağıdaki tabloda verilen değerler bir rehber oluşturabilir¹:

Tablo 1: Olabilirlik oranı sınıflandırması.

Niteliksel güç	LR(+)	LR(-)
Mükemmel	10	0.1
Çok iyi	6	0.2
Orta	2	0.5
Kullanışsız	1	1

Şimdi Bayes yaklaşımını kullanarak bir örnek yapalım; bu bölümün başındaki Senaryo 3 üzerinden otit tanısı koymaya çalışalım. Ramakrishnan ve arkadaşlarının makalesinde² otit için tanı testlerinin değeri Tablo 2’deki gibi verilmiştir.

Tablo 2: Otit tanısı için tanısal testlerin değeri.

	Duyarlılık (%)	Özgüllük (%)	LR+	LR-
Bulgu				
Timpan zarı bombeliği	61	97	20,3	0,4
Timpan zarı bulanık	81	95	16,2	0,2
Zar hareketi azalmış	98	79	4,7	0,03
Belirti				
Ailenin şüphesi	70	80	3,4	0,4
Kulak çekiştirme	42	87	3,3	0,7

¹ Mayer D. Essential Evidence-based Medicine. Cambridge University Press. 2004.

² <http://www.aafp.org/afp/2007/1201/p1650.html>

	Duyarlılık (%)	Özgüllük (%)	LR+	LR-
Kulak ağrısı	54	82	3,0	0,6
Aşırı ağlama	55	69	1,8	0,7
Burun akıntısı	75	43	1,3	0,6
İştah azlığı	36	66	1,1	1,0
Öksürük	47	45	0,9	1,2
Ateş	40	48	0,8	1,2

Adım adım, her bir ipucunu değerlendirerek hareket edeceğiz. Öncelikle elimizde temel bir ön olasılık var: bulunduğumuz toplumda ateşli bir çocukta otit olma olasılığının (yaygınlığının) %15 olduğunu biliyoruz.

İpucu 1: Ateş (Duyarlılık=%40, Özgüllük=%48)
LR+ = 0,80 (çok düşük)
İpucu 2: Rinit (Duyarlılık=%75, Özgüllük=%43)
LR+ = 1,3 (çok düşük)
İpucu 3: Anne şüphesi (Duyarlılık=%70, Özgüllük=%80)
Şimdi ön olasılığımız olan %15'i kullanarak pretest oddsu hesaplayacağız:
Pretest odds = $0,15 / (1 - 0,15) = 0,176$
LR+ = 3,4
Posttest odds = $0,176 \times 3,4 = 0,598$
Posttest oddsu olasılığa çevirecek olursak:
Posttest olasılık = $0,598 / (1 + 0,598) = \%37,4$ (karar vermek için yetersiz)
İpucu 4: Kulağını sürtme/çekme (Duyarlılık=%42, Özgüllük=%87)
Pretest odds = $0,374 / (1 - 0,374) = 0,598$

$LR+ = 3,3$
$Posttest\ odds = 0,598 \times 3,3 = 1,97$
$Posttest\ olasılık = 1,97 / (1 + 1,97) = \%66,32$

Hekim, bu bulgu ve belirtilerle ancak %66 ihtimalle otit tanısı koyabiliyor. Bu durumda ya mevcut olasılıkla yetinerek kararını verecek, veya başka ipuçları toplamaya çalışacaktır.

Bölüm 51: Biyoistatistik Paket Programları

Memet IŞIK

İstatistik programları verilerin elde edilmesi, toplanması, organize edilmesi, yorumlanması ve sunulması için kullanılan bilgisayar programlarıdır. Bu kitapta SPSS programı üzerinden anlatım yapılmış olmakla birlikte, çeşitli avantaj ve dezavantajlara sahip çok sayıda paket program bulunmaktadır.

Bu bölümde en çok kullanılan istatistik programları hakkında kısa bilgiler verilecek, programların birbirine karşı avantaj ve dezavantajlarına değinilecektir. Programların karşılaştırılması konusunda ayrıntılı bilgileri https://en.wikipedia.org/wiki/Comparison_of_statistical_packages web adresinden alınabilir.

Birçoğunun kullanımı ücretsiz olan sayıları yüzlerle ifade edilebilecek çok sayıda istatistik programı vardır. Ücretsiz kullanılabilen programlara örnek olarak;

- GNU Octave (<http://www.gnu.org/software/octave/>),
- ADaMSoft (<http://adamsoft.sourceforge.net/>)
- SAS University Edition (<http://www.sas.com/>),
- GNU PSPP (<http://www.gnu.org/software/pspp/>),
- The R Project (<https://www.r-project.org/>),
- pbdR (<http://r-pbd.org/>),
- Shogun (<http://www.shogun-toolbox.org/>),
- CSPro (www.census.gov/population/international/software/cspro),
- CumFreq (<http://www.waterlog.info/cumfreq.htm>),
- OpenMx (<http://openmx.psyc.virginia.edu/>),
- OpenStat (<http://statpages.info/>),
- Sage Stats (<https://us.sagepub.com/en-us/nam/sage-stats>),
- DAP (<http://www.gnu.org/software/dap/>)

gibi programları örnek olarak verebiliriz.

Ücretli istatistik programlarına örnek olarak ise;

- IBM SPSS (<http://www-01.ibm.com/software/analytics/spss/>),
- Minitab (<http://www.minitab.com/en-us/>),
- SAS (http://www.sas.com/en_us/software/analytics/stat.html),
- JMP (http://www.jmp.com/en_us/home.html),

- Genstat (<http://www.vsni.co.uk/products/genstat/>),
- STATISTICA (<http://www.statsoft.com/>),
- BMDP (<http://www.statsols.com/>),
- LIMDEP (<http://www.limdep.com/>),
- LISREL (<http://www.ssicentral.com/lisrel/>),
- NCSS (<http://www.ncss.com/>),
- SigmaStat (<http://www.sigmaplot.com/>),
- Stata (<http://www.stata.com/>),
- StatsDirect (<http://www.statsdirect.com/>),
- MedCalc (<https://www.medcalc.org/>),
- Unistats (<https://unistats.direct.gov.uk/>)

gibi programları verebiliriz.

İstatistikte veri analizinde kullanılan iki ana metodolojik yaklaşım vardır: tanımlayıcı (descriptive) istatistik ve çıkarımsal (inferential) istatistik. Tanımlayıcı istatistikte bir örneklemden elde edilen verilerin ortalama, standart sapma, minimum, maksimum, dağılım yüzdesi gibi veriyi özetleyecek tarzda verilmesidir. Çıkarımsal istatistikte ise elde edilmiş olan verilerden bir gözlem hataları örneklem farklılıkları gibi rast gele değişkenler de göz önüne alarak çıkarımlar yapmayı sağlar. İstatistikte kullanılan programları seçerken de ihtiyacımıza en iyi cevap veren programı seçmemiz uygun olur.

SPSS

IBM® SPSS® (Statistical Package for the Social Sciences): İlk defa 1968 yılında kullanıma sunulmuştur. 2009 yılında IBM tarafından 1,2 milyar Amerikan Dolarına satın alınmıştır. Program her ne kadar başlangıçta sosyal bilimler için geliştirilmiş olsa da günümüzde sağlık bilimlerinde çok yaygın olarak kullanılmaktadır.

IBM SPSS programı verilerin elde edilmesinde, işlenmesinde, anket analizlerinde, tahmine dayalı analizlerde (predictive) sosyal bilimler ve sağlık bilimlerinde yaygın kullanılan bir istatistik programıdır. SPSS'in ihtiyaca göre satın alınabilen 10'dan fazla modülü vardır.

Baz modelinde:

Tanımlayıcı istatistikler: çapraz tablolar, frekanslar, tanımlama bilgileri, tanımlayıcı oran istatistikleri

İki değişkenli istatistikler (Bivariate statistics): ortalama, t-testi, ANOVA, korelasyonlar, nanparametrik testler

Nümerik çıktılar için öngörüler: Linear regresyon

Grupları belirlemek için öngörüler: Faktör analizleri, küme analizler (iki-basamak, k-ortalamalar, hiyerarşik analizler) ve ayrıştırma analizleri yapılabilir.

IBM® SPSS® Amos değişkenler arasındaki hipotez kurulan ilişkileri belirginleştirmek, bu değişkenler arasındaki ilişkiyi tahmin etmek, değerlendirmek ve bütün bu ilişkileri standart multivaryasyon istatistik tekniklerinden çok daha hassas bir şekilde model olarak sunmak için kullanılan bir ilave modüldür. Grafik veya grafik olmayan ara yüzler şeklinde kullanılabilir. Sosyal bilimlerde ve sağlık bilimlerinde tutum ve davranışların kompleks ilişkilerini araştırmada kullanılabilir. SPSS'in Temmuz 2015 tarihindeki versiyonu SPSS 23'tür.

Minitab

Pennsylvania Devlet Üniversitesi tarafından 1972 yılında geliştirilmiştir. Mülkiyeti Minitab Inc.'e aittir. 1986 yılından beri de programda önemli bir değişiklik yapılmamıştır. Son sürümü Minitab 17 dir (Temmuz 2015 tarihinde) ve 8 dilde yayınlanmıştır (Türkçe versiyonu yoktur). Programı tamamlayan iki modülü vardır.

İstatistik Programlarının Karşılaştırılması

1. Minitab'ın kullanım ve öğrenilmesi kolay, ara yüzleri kullanıcı dostudur.
2. Kalite kontrol programları ile ilgili istatistik analizlerinde Minitab açık ara öndedir. Özellikle çok kullanışlı grafik ve tablolar hazırlayabilme fonksiyonları ile birçok dünya devi şirketlerin kullanmayı tercih ettiği istatistik programıdır.
3. Çok sofistike istatistik işlemleri ve çok yoğun veri içeren çalışmalar için Minitab STATA, SPSS ve R kadar güçlü bir program değildir. Bu tür çalışmalarda SPSS tercih edilmelidir. Fakat sağlık bilimlerinde kullandığımız hemen hemen bütün araştırmalar için Minitab bütün ihtiyaçlarımızı karşılayabilecek niteliktedir.
4. SPSS'in en önemli avantajı syntax fonksiyonu sayesinde kompleks işlemlerin kopyalanarak farklı verilerde uygulanmasına olanak sağlamasıdır. Ayrıca SPSS'de ihtiyaç duyulan ilave programlamar da yapılabilir.
5. Minitab'da da SPSS gibi, ihtiyaç duyulduğunda ilave programlar yazılabilmektedir. Programlama Minitab'da daha kolaydır.
6. ANOVA gibi bazı istatistik işlemlerinde SPSS'in daha güçlü bir istatistik programı olduğu, ilave modüller ile de hemen her ihtiyacı karşılayabilecek bir esneklik sağlar.

7. Çok hassasiyet gerektiren analizler için ise SPSS'in STATA ve R kadar hassas olmadığı kabul edilir.
8. SPSS çok büyük verilerle çalışılıyor, çok aşırı ölçüm hassasiyeti gerektirmeyen orta ve kısmi olarak da ileri istatistikler yapılıyor ise SPSS tercih edilmelidir.

Çok kullanılan istatistik programlarının mukayese edilmesi					
Kullanılan İstatistik Testi	R	MATLAB	SAS	STATA	SPSS
Nonparametrik testler	Evet	Evet	Evet	Evet	Evet
t-testi	Evet	Evet	Evet	Evet	Evet
ANOVA, MANOVA	Evet	Evet	Evet	Evet	Evet
ANCOVA, MANCOVA	Evet	Evet	Evet	Evet	Evet
Lineer Regresyon	Evet	Evet	Evet	Evet	Evet
Generalize Lineer Model	Evet	Evet	Evet	Evet	Evet
Lojistik regresyon	Evet	Evet	Evet	Evet	Evet
Nonlinear regresyon	Evet	Evet	Evet		
Discriminant analiz	Evet	Evet	Evet	Evet	Evet
Faktör analizi	Evet	Evet	Evet	Evet	Evet
Bayesian istatistik	Evet	Evet	Kısmi		
Monte Carlo	Evet	Evet	Evet	Evet	Kısmi
Eksik veri işleme	Evet	Evet	Evet	Evet	Evet
Uç değer tanıma	Evet	Evet	Evet	Evet	Evet
Sağkalım analizi	Evet	Evet	Evet	Evet	Evet
Tabakalı örneklemeler	Evet	Evet	Evet	Evet	Evet
Deneyisel tasarım	Evet	Evet			
Kalite kontrol	Evet	Evet		Evet	Evet
Güvenilirlik teorisi	Evet	Evet	Evet	Evet	Evet
Tek değişkenli zaman serileri	Evet	Evet	Evet	Evet	Kısmi

Çok değişkenli zaman serileri	Evet	Evet	Evet	Evet	
Süreçleri sayma	Evet	Evet	Evet		
Filtreleme	Evet	Evet	Kısmi	Kısmi	
İnstrumental değişkenler	Evet	Evet	Evet	Evet	
Varyans eşitleme	Evet	Evet			
Küme analizi	Evet	Evet	Evet	Evet	Evet
Nöral ağlar	Evet	Evet	Evet		Kısmi
Sınıflandırma ve regresyon ağacı	Evet	Evet	Evet		Kısmi
ROC eğrileri	Evet	Evet	Evet	Evet	Evet

* <http://www.datasciencecentral.com/profiles/blogs/comparison-of-statistical-software> (kısaltılarak alınmıştır)

Bölüm 52: Rapor Yazma

Memet IŞIK

Bilimsel araştırmalar tamamlandıktan sonra en önemli aşama yapılan araştırmanın bir rapor olarak bilimsel camiaya sunulmasıdır.

Bu bölümde bir bilimsel çalışmanın raporlanması ve makaleye dönüştürülmesi hakkında bilgi verilmesi amaçlanmıştır.

Araştırma raporu yazımında uyulması gereken temel iki kural vardır:

1. Biçimin olabildiğince basit ve sade tutulması
2. Raporun tamamında yazım kurallarında belirlenen raporlama biçiminin kurallarına uyulması

Bilimsel çalışmaların raporlanmasında uluslararası ölçekte kabul gören iki bilimsel rapor yazım sistemi vardır: Geleneksel Anglo-Sakson Sistemi ve Kıta Avrupa'sı Sistemi.

Ülkemizde genellikle 'Geleneksel Anglo-Sakson' bilimsel rapor yazım sisteminin kullanılması öngörülmektedir. Özellikle Sağlık ve Sosyal Bilimler ile ilgili alanlarda yazılan raporlarda Amerikan Psikoloji Derneği tarafından belirlenen standartların kullanımı, yaygın olarak benimsenmektedir.

Kullanılan yazı tipi

Bilimsel araştırmalar genellikle Times New Roman yazı fontu ile yazılır. Farklı bilimsel dergiler farklı fontların kullanılmış olmasını isteyebilir. Rapor elektronik ortamda gönderilecekse derginin istediği yazı fontunun kullanılması gerekir. Times New Roman dışında Arial, Bookman, Palatino, Verdana vb. yazı karakterleri de kullanılabilir.

Araştırma raporu satırları iki yana yaslı olarak yazılmalıdır. Raporun farklı kısımlarında farklı yazı karakteri büyüklükleri kullanılmalıdır. Bu büyüklükler Times New Roman yazı karakterine göre şu şekildedir:

- Raporun genel yazı karakteri büyüklüğü 12 puntodur.
- Tablo ve şekillerin başlıkları ve içerikleri için 11 punto yazı karakteri kullanılmalıdır.
- Tablo ve şekillerin altına yazılacak alıntı veya kısaltma ek bilgileri için 10 punto yazı karakteri kullanılmalıdır.
- Dipnot ve son notlar için 10 punto yazı karakteri kullanılmalıdır.
- Alt ve üst indislerin yazımında, düz yazı büyüklüğünden 2 punto daha küçük yazı karakteri kullanılmalıdır.

Araştırma raporunun tamamında yazı tipi ve büyüklüğü tutarlı olmalıdır. Gerekli görülen yerlerde sözcük, cümle, başlık vb. kısımlar koyu veya eğik karakterde yazılabilir, semboller kullanılabilir.

Sayfa Kenar Boşlukları

Rapor kâğıda basılacak şekilde kaydedilecekse sayfa kenar boşlukları, sol kenardan 3 cm, üst, alt ve sağ kenarlardan 2,5 cm olmalıdır Başlıklar, sayfa numaraları, rapor metni, çizim ve çizelgeler vb. belirlenen alana yazılmalı, sayfa kenar boşluklarında hiçbir yazı bulunmamalıdır. Dergiler farklı sayfa boşluğu isteyebilir. Raporun veya makalenin yayınlanacağı derginin yazım kurallarına göre düzenleme yapılmalıdır.

Satır Aralıkları

Satır aralığı, yazımda kullanılan yazı karakterinin büyüklüğü kadar boşluğu ifade etmektedir. Araştırma raporu ‘1,5 satır aralığı’ ile yazılmalı; paragraflar, satır başından itibaren 1,25 cm içeriden başlamalı ve paragraflar arasında bir ‘ara boşluk’ bırakılmalıdır. Çizim, çizelge, tablo, şekil ve grafiklerde, uzun alıntı, dipnot ve son not kısımlarında -raporun tamamında aynı kalmak koşulu ile- ‘tek satır aralığı’ kullanılabilir. Ana başlıklar paragraftan iki satır, alt başlıklar ise bir önceki paragraftan bir satır aralık kalacak şekilde yazılmalıdır. Alt başlık ile bir sonraki paragraf arasında satır aralığı verilmemelidir. Makale şeklinde yazı bir dergiye gönderilecek ise; derginin istediği satır aralıklarına göre yazı düzenlenmelidir. Genellikle dergiler 2 satır aralıklı yazı düzeni istemektedir.

Başlık Düzeni

Her bir bölüm için, içeriği en iyi şekilde ifade eden ve araştırma sınırlılıklarına Uygun bir ana başlık belirlenmelidir. Alt başlıklar, bölümün ana başlıkları ile tutarlı olmalıdır. Her bölüm ayrı bir sayfadan başlamalıdır. Dolayısıyla her bir ana başlık, yeni bir sayfaya yazılmalıdır.

Makale formatında bilimsel bir dergiye gönderilecek araştırma yazılarında derginin istemiş olduğu yazım kurallarına göre başlıklar düzenlenmelidir.

Dipnot ve Son Not

Araştırma raporunda dipnot kullanımı mümkün olduğunca sınırlı tutulmalıdır. Metin içerisinde yer alabilecek ifadelerin dipnot olarak belirtilmesinden kaçınılmalıdır. Metin içerisinde alıntı yapılan kaynak ve göndermeler, dipnot veya son not olarak belirtilmez. Araştırmacı, konu ile ilgili fakat konu akışına uygun olmayan bilgileri gerekli görürse dipnot olarak verebilir. Dipnot, sayfanın alt kısmında yer almalı ve 10 punto Times New Roman yazı karakterinde olmalıdır. Diğer yazı karakterlerinden biri kullanılacaksa, tercih

edilen bu yazı karakterinin büyüklüğü, burada Times New Roman yazı karakteri için verilen ölçülere uyumlaştırılmalıdır. Dipnotlarda ilk cümle paragraf başından (1,25 cm) başlamalıdır. Son not, bölüm sonlarında yer almalıdır. Araştırmada yararlanılan ve gönderme yapılan kaynaklar son not olarak belirtilmez. Bunlar 'KAYNAKLAR' bölümünde yer almalıdır.

Dipnot ve son notlar ana metin/gövdeden yatay satır çizgisi ile ayrılmalıdır. Birden fazla dipnot/son not verilecekse, her not arasında bir satır aralığı bırakılmalıdır. Dipnot ve son notlar normal sayılar (1, 2, 3,...) kullanılarak verilmelidir.

Sayfa Numaraları

Farklı dergiler farklı şekilde numaralandırma isteyebilir veya sayfaların numaralandırılmamasını isteyebilir. Eğer araştırma bir makale tarzında değil de bir araştırma raporu şeklinde yazılacaksa tüm sayfa numaraları, sayfanın sağ-üst kısmına yazılmalıdır. Sayfa numaraları, rapor metninden en az iki satır yukarıda ve sayfanın yazılabilir alanı içerisinde (üst kenardan 2,5 cm aşağıda) yer almalıdır. Rapor metninde 'Giriş Bölümleri' küçük Romen rakamları ile (i, ii, iii,...), diğer kısımlar normal sayı formatı ile (1, 2, 3,...) numaralandırılmalıdır. 'Kapak' ve 'İç Kapak' sayfaları numaralandırılmaz.

Verilerin Tablolar ve Şekillerle Gösterilmesi

Tabloların özellikleriyle ilgili sayfa 76'te "Bölüm 7: Verilerin Tablolarla Gösterilmesi" başlığına bakınız.

İstatistiksel Gösterimler

Özellikle nicel araştırmalarda istatistiklerin (t testi, F testi, ki-kare testi vb.) yazımında ortak bir tarzın belirlenmesi önem taşımaktadır. Yapılan analizlere ilişkin tanımlayıcı istatistikler tablo ya da şekil olarak verilmişse, bu bulgu ve sonuçların metin içerisinde tekrar edilmesine gerek yoktur. Ancak kullanılan test istatistiğine ilişkin değerler (istatistik değeri, serbestlik derecesi, örneklem büyüklüğü ve olasılık değeri) metin içerisindeki anlatımı kolaylaştırmak için verilmelidir. Test istatistiklerinin raporda nasıl gösterileceğine ilişkin bazı örnekler şu şekilde verilebilir:

$$F = 22.37, p = 0,002$$

$$t = 1.06, p = 0,091$$

$$\chi^2 = 14.58, p = 0,001$$

Tablolarda nümerik değişkenler ortalama ve standart sapma ile birlikte verilmelidir. Eğer veriler normal dağılmıyorsa standart sapma yerine minimum, maksimum ve gerekiyorsa ortanca da verilmelidir. Kategorik veriler ise sayı (n) ve yüzde (%) olarak verilmelidir.

Araştırma Raporu Bölümleri

Araştırma raporunda şu bölümler olmalıdır:

1. Giriş veya arka plan
 - a. Problem Durumu
 - b. Araştırmanın Amacı
 - c. Araştırmanın Önemi
 - d. Tanımlamalar
2. Yöntem
 - a. Araştırma Tasarımı
 - b. Araştırmanın Evreni ve Örneklemi
 - c. Veri Toplama Araçları ve Geliştirilmesi
 - d. Verilerin Toplanması
 - e. Verilerin Analizi
 - f. Sınırlılıklar
 - g. Varsayımlar
4. Bulgular
5. Tartışma
6. Sonuç ve Öneriler

Giriş Bölümü

Çalışmanın ana metin/gövde kısmı, problemin sunulduğu ve araştırma stratejisinin tanıtıldığı ‘GİRİŞ’ bölümü ile başlar. Giriş kısmında araştırmanın alt yapısı oluşturulur ve araştırma ile ilişkili değişkenler tanıtılır.

Bilimsel bir araştırma, bilimsel bir yöntem çerçevesinde oluşturulur. Kuramsal çerçeve olarak tanımlanabilecek araştırmanın adı, konusu, amacı, önemi, yöntemi, kapsamı, sınırlılıkları vb. belirlenmesi, bir bilimsel araştırmanın ilk basamaklarıdır. Bu aşamalar, bir araştırma raporunda açık ve anlaşılır bir şekilde sırasıyla yer almalıdır.

Nitel ve kuramsal tezlerde bu aşamalar ‘GİRİŞ’ bölümü içerisinde açıklanabilir. Nicel araştırmalarda ise bu aşamaların ayrı alt başlıklar altında açıklanması uygun görülmektedir.

Giriş bölümü çok uzatılmamalı ve araştırma konusunu tanımlayacak bir içerikte olmalıdır. Giriş, araştırma raporunun birinci bölümü olmalıdır. Bu bölümde Problem Durumu kısaca anlatılmalı, araştırmanın amacı açıkça ifade edilmelidir. Daha sonra ise araştırmanın niçin önemli olduğu vurgulanmalı ve gerekiyorsa yeni tanımlamalar yapılmalıdır.

Araştırma raporunun giriş kısmında literatür taramasına da yer verilmelidir. Burada ele alınan konular araştırma konusu ile sınırlı tutulmalı, gereksiz bilgi ve ayrıntılara yer verilmemelidir. Bu bölümde başvuru ve gönderme yapılan tüm eserler ve kaynaklar, bilimsel alıntı gösterim kurallarına uygun biçimde belirtilmelidir.

Yöntem

Araştırmanın yapılmasında izlenen bilimsel yöntem ve bu yöntemin kullanılmasının gerekçeleri, araştırma raporunda açıkça belirtilmelidir. Veri toplama yöntemleri ve verilerin analizine ilişkin kullanılan istatistiki yöntemler açıklanmalıdır. Nicel araştırmalarda evren ve örnekleme ilişkin bilgilerin de ayrıntılı olarak belirtilmesi gerekmektedir.

Araştırma Deseni genel olarak araştırmanın nasıl yapıldığının ifade edildiği bölümdür. Amaçların açık bir şekilde ortaya konması, bu amaçları gerçekleştirebilmek için en uygun yöntemin belirlenmesini de kolaylaştıracaktır. Araştırmanın yöntem ve deseninin belirlenmesinde araştırma için öngörülen süre de dikkate alınmalıdır.

Araştırmanın evreni, araştırma konusu ve amacına uygun olarak belirlenmelidir. Örneklem, evreni en iyi temsil edecek biçimde bilimsel yöntemlerle (rastgele örnekleme, tabakalı örnekleme vb.) belirlenmelidir. Evren ve evrenin seçiminin gerekçeleri, örneklem ve örneklemin belirlenmesinde izlenen bilimsel yöntem, araştırma raporunda açıkça ve detaylı bir şekilde belirtilmelidir.

Araştırmanın uygulama bölümünde, araştırmanın amaçlarına ve araştırma desenine en uygun veri toplama araçları kullanılmalıdır. Daha önceden geliştirilmiş, literatürde yer alan, geçerlik ve güvenirlik çalışmaları yapılmış veri toplama araçları kullanılabilir. Araştırma desenine uygun veri toplama araçları araştırmacı tarafından da geliştirilebilir. Bu durumda araştırma kapsamında kullanılan veri toplama araçlarının geçerlik ve güvenirlik kanıtları, veri toplama süreci açık ve ayrıntılı bir şekilde anlatılmalıdır. Verilerin nasıl, ne zaman, kimler tarafından toplandığı da belirtilmelidir.

Araştırmacı, toplanan verilerin sınıflandırılmasında ve değerlendirilmesinde kullandığı yöntemleri araştırma raporunda ayrıntılı bir şekilde belirtmelidir. Verilerin analizinde kullanılan istatistikler ve kullanılan istatistik programlarına ilişkin ayrıntılı bilgiler verilmelidir.

Bilimsel araştırma, araştırmanın amacı ve desenine uygun olarak bazı sınırlılıklara sahiptir. Bu sınırlılık, araştırma konusunda olabileceği gibi araştırmanın uygulama boyutunda, evren ve örneklem belirlenmesinde de söz konusu olabilmektedir. Araştırmanın sınırlılıkları ve bu sınırlılıkların gerekçeleri, araştırma raporunda açıkça belirtilmelidir.

Bulgular

Bulgular nesnel olmalıdır. Bulgular bölümü, verilerin analizi ile elde edilen bilgiler ve ulaşılan nesnel sonuçların yorumsuz olarak verildiği bölümdür. Bulguların anlaşılır bir biçimde sunulması için araştırmacı uzun anlatımlar yerine tablo ve/veya şekil kullanabilir.

Tartışma, Sonuç ve Öneriler

Sonuç ve öneriler öznel olmalıdır. Araştırmacı, elde edilen bulgular doğrultusunda özgün fikirlerini tartışarak ortaya koymalı, buna bağlı olarak ulaştığı sonuçları ve önerilerini belirtmelidir. Her bir bulgu ile ilgili bir önerinin getirilmesi tavsiye edilmektedir. Bu bölümde ulaşılan sonuçların diğer araştırmaların sonuçları ile benzerlikleri ve farklılıkları, bunun olası nedenleri ayrıntılı bir şekilde tartışılmalıdır. Öneri ve yorumlar genel olmamalı, araştırma bulgularıyla birebir ilişkilendirilmelidir.

Kaynaklar

Rapor metninde atıf yapılan kaynakların, yazım kurallarında istenilen formata uygun bir sırada ve kaynak gösterme kurallarına uygun bir şekilde listelendiği bölümdür. Rapor metninde doğrudan veya dolaylı olarak atıf yapılmamış hiçbir eser, kaynak olarak belirtilmemelidir¹.

Sıkça kullanılan Vancouver stiline göre kaynaklar şu şekilde gösterilmektedir:

Bilimsel dergi makalesi: Aktürk Z, Işık M. Türkiye’de Hekim İhtiyacı: Aile Hekimliği Özelinde 2013-2033 Yılları İçin Bir Projeksiyon Çalışması. Euras J Fam Med 2013; 2(1): 33-36

Kitap: Aktürk Z, Acemoğlu H. Sağlık Çalışanları İçin Araştırma ve Pratik İstatistik. 2. Baskı. Anadolu Matbaası. İstanbul 2011.

Kitapta bölüm: Işık M, Aktürk Z. A. Turgut (Editörler), Sağlıklı Yaşam (Ünite 6: Şişmanlığın Tanı, Tedavi ve Önlenmesi), 98-109 pp., Anadolu Üniversitesi Yayınları No:2866, Açık Öğretim Fakültesi Yayınevi (No: 1823), Eskişehir 2013.

Bilimsel kongre bildirisi: Harnden P, Joffe JK, Jones WG, editors. Germ cell tumours V: Proceedings of the 5th Germ Cell Tumour conference; 2001 Sep 13-15; Leeds, UK. New York: Springer; 2002.

Web sitesi: Işık M. Tıbbi Sülük Tedavisi. Atatürk Üniversitesi Akupunktur ve Tamamlayıcı Tıp Yöntemleri Uygulama ve Araştırma merkezi (ATYAM). Erişim tarihi: 30 Temmuz 2015.

<http://sayam.atauni.edu.tr/Default.aspx?Default=4>

¹ T.C.Millî Eğitim Bakanlığı, Eğitimi Araştırma ve Geliştirme Dairesi Başkanlığı, Araştırma raporu yazım kılavuzu, ANKARA, Mayıs, 2009

Bölüm 53: Araştırma Fonları

Memet IŞIK

Araştırma fonları bilim insanları için son derece önemlidir. Bir araştırmayı analiz aşamasına getirinceye kadar geçmesi gereken aşamalar arasında maddi kaynak bulmak en önemli hususlardandır.

Bu bölümde ulusal ve uluslararası araştırma fonları hakkında bilgi verilmesi amaçlanmıştır.

Üniversitelerin Bilimsel Araştırma Proje (BAP) destekleri

Ülkemizdeki üniversitelerin sağladığı Bilimsel Araştırma Proje Destekleri için her üniversitenin kendi belirlediği başvuru şartı vardır. Üniversitelerin web sayfalarından bu başvurularla ilgili ayrıntılı bilgiler alınabilir. BAP projeleri de kendi içinde farklı kategorilerde değerlendirilebilir ve verilen destek miktarları da proje türüne göre farklılık gösterir. Aşağıdaki örnek tabloda Atatürk Üniversitesi'nin BAP proje türleri ve verilen desteğin üst sınırları belirtilmiştir:

Projenin Türü	Üst Limiti	Açıklama
Kapsamlı Araştırma Projesi	₺50.000	Yolluk + hizmet alımları bütçenin %30'unu geçemez. Eğer bu oran 2000 TL'nin altında olursa, 2000 TL'ye tamamlanabilir. 12. maddede bildirilen hizmet alımlarında bu limit uygulanmaz.
Lisansüstü Öğrenim Araştırma Projesi	Doktora ve muadili: ₺15.000 Yüksek Lisans: ₺10.000	4.Maddedeki hükümler uygulanır.
Destek Projesi	₺25.000	Sadece sarf malzeme ve yedek parça alımında kullanılacaktır.
Altyapı Projesi	₺150.000	Yolluk ve hizmet alımına kapalıdır.
Güdümlü Araştırma Projesi	₺200.000	Üst sınır gerekli görülen hallerde BAP Koordinasyon Komisyonu tarafından artırılabilir.
Katılımlı Araştırma Projesi	₺50.000	Diğer paydaşların sağladığı maddi katkıdan fazla olamaz.

Sanayi İşbirliği Projesi	₺100.000	Diğer paydaşların sağladığı maddi katkıdan fazla olamaz.
İleri Araştırma Projesi	-	Kalkınma Bakanlığı'na sunulur.
Patent, Fikri Mülkiyet ve Tescil Destek Projeleri	₺20.000	-
Kongre, Sempozyum ve Toplantı Düzenleme Destek Projeleri	₺20.000	Üst sınır gerekli görülen hallerde BAP Koordinasyon Komisyonu tarafından artırılabilir.

TÜBİTAK Araştırma Destek Programları

Türkiye Bilimsel ve Teknik Araştırmalar Kurumu (TÜBİTAK), Türkiye'deki Ar-Ge fonlamasından sorumlu başlıca kuruluştur. TÜBİTAK, sanayi ve akademik olmak üzere iki çeşit Ar-Ge proje desteği sağlamaktadır (<http://www.tubitak.gov.tr/tr/destekler/akademik/ulusal-destek-programlari>).

Sanayi Ar-Ge destekleri Teknoloji ve Yenilik Destek Programları Başkanlığı (TEYDEB) tarafından yönetilmekte olup,

- 1501 (Sanayi Ar-Ge Destekleme Programı),
- 1503 (Proje Pazarları Destekleme Programı),
- 1507 (KOBİ Ar-Ge Başlangıç Destek Programı),
- 1508 (TEKNOGİRİŞİM: Teknoloji ve Yenilik Odaklı Girişimleri Destekleme Programı) ve
- 1509 (Uluslararası Sanayi Ar-Ge Destekleme Programı) gibi programlardan oluşmaktadır.

Teknolojinin toplumsal faydaya dönüşmesi için katkıda bulunmayı amaçlayan bu programlardan 1501, öncelikle firma düzeyinde katma değer yaratan kuruluşlardaki Ar-Ge nitelikli proje harcamalarına %60'a varan oranlarda geri ödemesiz hibe desteği sağlamaktadır. 1507, KOBİ'lerin 400.000 TL bütçeye ve 18 aya kadar Ar-Ge harcamalarını desteklemekte; 1508 ve 1509 ise, büyük ölçekli firmaların Ar-Ge harcamalarına %60, KOBİ'lerin ise %75 oranında hibe desteği vermektedir.

Araştırma Destek Programları Başkanlığı (ARDEB) tarafından yürütülen akademik nitelikli Ar-Ge destekleri ise, 1001, 1002, 1007, 1008, 1010, 1011,

1301 ve 3501 gibi sekiz farklı programdan oluşmaktadır

(<http://www.tubitak.gov.tr/tr/destekler/akademik/ulusal-destek-programlari>):

1001-Bilimsel ve Teknolojik Araştırma Projelerini Destekleme Programı

Bilimsel ve Teknolojik Araştırma Projelerini Destekleme Programının amacı, yeni bilgiler üretilmesi, bilimsel yorumların yapılması veya teknolojik problemlerin çözülmesi için bilimsel esaslara uygun olarak yurtiçinde yapılan çalışmaları desteklemek, bu yolla öncelikli alanlarda uluslararası rekabet kapasitesini arttırmaktır.

- Üniversiteler, kamu ve özel sektör kurum ve kuruluşları başvurabilir.
- Karşılanan masraflar arasında personel, burs, seyahat, teçhizat, danışmanlık ve diğer hizmetler, bakım, kargo, ulaşım ve kapasite maliyetleri vardır.
- Azami destek tutarı yıllık 120.000 TL, azami destek süresi 36 aydır.
- Son başvuru tarihleri Mart ve Eylül aylarındadır.

1002-Hızlı Destek Programı

Hızlı Destek Programı, üniversitelerde, araştırma hastanelerinde ve araştırma enstitülerinde yürütülecek acil, kısa süreli, küçük bütçeli araştırma ve geliştirme projelerine destek sağlamaya yönelik bir programdır.

- Azami destek tutarı 30.000 TL, azami destek süresi 12 aydır.
- Başvuru süresi yoktur.

1003-Öncelikli Alanlar AR-GE Projeleri Destekleme Programı

1003 Öncelikli Alanlar AR-GE Projeleri Destekleme Programı'nın amacı, Ulusal Bilim Teknoloji ve Yenilik Stratejisi çerçevesinde belirlenecek öncelikli alanlarda sonuç odaklı, izlenebilir hedefleri olan, ilgili bilim/teknoloji alanlarının dinamiklerini gözeten ve yurt içinde yapılan Ar-Ge projelerini desteklemek ve bu projeler arasında eşgüdüm sağlamaktır.

- Öncelikli alanlar kapsamında çağrılar açılmaktadır.
- İlgili endüstriyel kuruluşlarla işbirliği içinde hazırlanmış ve/veya ilgili endüstriyel kuruluşlardan aynı/nakdi destek almış olan projelere öncelik verilecektir.
- Projelerde farklı kuruluşlardan farklı disiplinlerde proje ekipleri olması beklenmektedir.
- Bu çağrı kapsamında önerilen projeler küçük, orta veya büyük ölçekli projeler olarak hazırlanabilir.

- Projeler küçük, orta ve büyük ölçekli olabilmektedir. Projeler için destek alt ve üst limitleri (Burs, Proje Teşvik İkramiyesi (PTİ) dahil) aşağıdaki gibi belirlenmiştir;
 - Küçük Ölçekli projeler: 500.000 TL'ye kadar
 - Orta Ölçekli projeler: 500.001 - 1.000.000 TL
 - Büyük Ölçekli projeler: 1.000.001 - 2.500.000 TL
- Orta ve büyük ölçekli projelerin birden fazla kurumun yer aldığı alt projelerden oluşması beklenmektedir.
- Proje süresi;
 - Küçük Ölçekli projelerin süresi en fazla 24 ay,
 - Orta ve Büyük Ölçekli projelerin süresi en fazla 36 aydır.
- İki aşamalı başvuru ve değerlendirme sistemi bulunmaktadır.

1007-Kamu Kurumları Araştırma ve Geliştirme Projelerini Destekleme Programı

Bu program, TÜBİTAK'ın misyonu çerçevesinde, kamu kurumlarının Ar-Ge ile giderilebilecek ihtiyaçlarının karşılanması ya da sorunlarının çözüme kavuşturulması amacıyla kurulmuştur. Kamu kurumlarının Ar-Ge ile giderilebilecek ihtiyaçlarının karşılanmasına ya da sorunlarının çözümüne yönelik projeleri desteklemek amacıyla, 1007 Programına sunulan projeler aşağıda belirtilen iki Araştırma Destek Grubu tarafından desteklenmektedir.

1. Kamu Araştırmaları Destek Grubu (KAMAG)
2. Savunma ve Güvenlik Teknolojileri Araştırma Destek Grubu (SAVTAG)

1008-Patent Başvurusu Teşvik ve Destekleme Programı

Patent Başvurusu Teşvik ve Destekleme Programı, Türkiye'nin ulusal ve uluslararası patent başvurularının sayısının arttırılması, kişilerin patent başvurusu yapmaya teşvik edilmesi ve ülkemizde fikri ve sınai hakların tescili yönünde bilinçlenmeye katkıda bulunulması amacıyla hazırlanmış bir programdır.

1011-Uluslararası Bilimsel Araştırma Projelerine Katılma Programı

Uluslararası Bilimsel Araştırma Projelerine Katılma Programı (1011), birçok ulusun araştırmacılarının ve kuruluşlarının ortaklığıyla yürütülen uluslararası projelere, ülkemizde istihdam edilmekte olan araştırmacıların katılmak için gerek duydukları desteğin sağlanması amacıyla oluşturulmuş bir programdır.

- Uluslararası ortak projelerin saygın bir uluslararası üst/şemsiye kuruluşun koordinatörlüğünde olması gerekmektedir.

- Projelerin Türkiye’de yürütülen kısmı sadece TÜBİTAK 1011 programının desteğiyle yürütülebilir.
- Türkiye hariç en az iki ülkenin yürütücü kuruluşlarının da üst/şemsiye kuruluş koordinatörlüğünde bu projede yer alması gerekmektedir.
- Karşılanan masraflar arasında personel, burs, seyahat, teçhizat, danışmanlık ve diğer hizmetler, bakım, kargo, ulaşım ve kapasite maliyetleri vardır.
- Azami destek tutarı yıllık 120.000 TL, azami destek süresi 36 aydır.
- Son başvuru tarihleri Mart ve Eylül aylarındadır.

1301-Bilimsel ve Teknolojik İşbirliği Ağları ve Platformları Kurma Girişimi Projeleri (İŞBAP) Destekleme Programı

Bu program, ulusal ve uluslararası kurum, kuruluş, birim ve gruplar arasında temel bilimler, mühendislik, sağlık bilimleri, sosyal bilimler ve ilgili teknoloji dallarında, özellikle ülkemizin bilim ve teknoloji öngörülerini doğrultusunda gelişmesini sağlamak üzere, ilgili taraflar arasında işbirliğini oluşturmak, artırmak ve bunların somut çıktılara yönelecek şekilde gelişmesini sağlamak üzere önerilecek işbirliği ağlarının ve platformların kurulmasını desteklemeye yönelik bir programdır.

- Sanayi ve iş dünyasından kurum ve kuruluşlar, üniversiteler, kamu kurum ve kuruluşları ile bunların oluşturduğu konsorsiyumlar başvuruda bulunabilir.
- TÜBİTAK toplam bütçenin %50’sini karşılamaktadır. Kalan %50’nin ağ veya platformdaki diğer kuruluşlarca karşılanması gerekir.
- Toplantı organizasyon, seyahat ve konaklama giderleri, büro, makine ve teçhizat, bilgi yayma ile ilgili giderler, web sayfaları, veri bankaları ve sanal laboratuvarlar kurmak için gerekli harcamalar (donanım, yazılım ve diğer), fikri ve sınai hizmetler ile danışmanlık ücretleri destek kapsamındadır.
- Azami destek tutarı yıllık 250.000 TL, azami destek süresi 36 aydır.
- Başvuru süresi yoktur.

3501-Ulusal Genç Araştırmacı Kariyer Geliştirme Programı (Kariyer Programı)

Ulusal Genç Araştırmacı Kariyer Geliştirme Programı’nın amacı kariyerlerine yeni başlayan doktoralı genç bilim insanlarının çalışmalarını proje desteği vererek teşvik etmektir. 21. yüzyılın akademik önderliğini yüklenecek genç araştırmacıların çalışmaları desteklenerek, hem genç bilim insanlarının kariyerlerini araştırmacı ve eğitimci olarak en iyi şekilde sürdürmeleri, hem de

bilimsel düzeyimizin geliştirilmesi ve bilimin ülke kalkınmasındaki rolünün artırılması amacıyla yönelik bir programdır.

- Adayların doktora derecelerini başvuru tarihinden en çok 5 yıl önce almış olması gereklidir.
- Karşılanan masraflar arasında personel, burs, seyahat, teçhizat, danışmanlık ve diğer hizmetler, bakım, kargo, ulaşım ve kapasite maliyetleri vardır.
- Azami destek tutarı yıllık 75.000 TL, azami destek süresi 36 aydır.
- Son başvuru tarihleri Mart ve Eylül aylarıdır.

TÜBİTAK Bilim İnsanı Destekleme Programları

Ar-Ge destek programlarının yanı sıra, TÜBİTAK bünyesindeki Bilim ve Toplum Daire Başkanlığı tarafından, 4004 Doğa Eğitimi ve Bilim Okulları Destekleme Programı ile 4003 Bilim Merkezi Kurulması Destek Programı adları altında iki destek programı yürütülmektedir. 4004 nolu program her yıl periyodik olarak çağrı yapmakta olup, doğa, bilim ve teknoloji konularında farkındalık yaratmak amacıyla bilimsel konu, kavram ve süreçlerin gözlem ve uygulamalarla anlaşılmasını sağlayan etkinlikler, eğitim programları, gözlem atölyesi, saha çalışması ile drama, ölçme-değerlendirme, deneysel, grup çalışmaları ile sanat ve spor etkinliklerini 120.000 TL'ye kadar desteklemektedir. 4003 ise, büyükşehir belediyelerinin bilim merkezi kurma kapsamında yürüttüğü master plan oluşturması, bilim merkezinin tasarımları, sergilerin tasarımı, üretimi, kurumu ve nakliyesi, atolye alanlarının tasarımı ve temini ile personelin eğitimi ve profesyonel destek satın alımı gibi faaliyetlerini desteklemektedir (<https://e-bideb.tubitak.gov.tr/giris.htm>).

Bunlara ilaveten, **Avrupa Birliği Çerçeve Programları**'na başvurmak isteyen bilim insanlarına her türlü teknik yardım ve destek, TÜBİTAK bünyesindeki AB Çerçeve Programları Ulusal Koordinasyon Ofisi (UKO) tarafından verilmektedir. TÜBİTAK Avrupa Birliği Çerçeve Programları Müdürlüğü de 2012'den itibaren üç yeni destek programını yürürlüğe koymuştur. Bu programlardan birincisi konsorsiyum oluşturma amaçlı seyahat desteği sağlarken, diğer ikisi de katılımı özendirme ödül desteği ve proje önerisi ön değerlendirme destekleri sunmaktadır.

TÜBİTAK, ayrıca, Bilim İnsanları Destekleme Daire Başkanlığı (BİDEB) aracılığıyla bilim insanlarının yurt içi ve yurt dışı faaliyetlerini desteklemektedir. Yurt içi burslar; lisans, yüksek lisans ve doktora burs programları ile doktora sonrası araştırma, konuk bilim insanı ve doktora sonrası geri dönüş burs programını kapsamaktadır. Yurt dışı bilimsel faaliyetler için verilen burslar ise, doktora burs ve araştırma bursu programı ile bilimsel

etkinliklere katılım ve bilimsel işbirliğine yönelik etkinlikleri destekleme programlarını kapsamaktadır.

Devlet Planlama Teşkilatı (DPT)

DPT, 2007 yılından itibaren, güdümlü teknoloji geliştirme projeleri ile araştırma altyapısı ve araştırmacı insan gücü yetiştirme programlarını desteklemektedir. DPT'nin desteklediği projelerde bütçenin alt sınırı 1.000.000 TL olup, öncelikli alanlar 2011 itibarıyla; geleceğe yönelik olarak nanoteknoloji, biyoteknoloji, yeni nesil nükleer teknoloji, hidrojen ve yakıt pili teknolojileri, sanayinin ihtiyaç duyduğu ve rekabet gücünü artıracak araştırmaların yanı sıra uzay ve savunma teknolojileri olarak belirlenmiştir.

Üniversite rektörlükleri ya da kurum idarelerinin sahipliğinde yürütülecek DPT Araştırma Altyapısı Destekleri ise iki gruba ayrılmaktadır. Bunlardan ilkinde “uzmanlık ve mükemmeliyet merkezleri” desteği olup Ar-Ge bakımından belirli bir kapasiteye ulaşmış üniversite ve araştırma kurumlarına verilirken, ikinci destek türü araştırma altyapısı eksik olan üniversite ve araştırma kurumlarında ileri düzeydeki araştırma projelerinin yürütülmesini sağlayacak büyük ölçekli merkezlerin oluşturulmasını amaçlamaktadır.

DPT'nin koordinasyonu altında çeşitli kurumlar Ar-Ge projelerine destek vermektedirler. Bunlardan Kalkınma Araştırmaları Merkezi (KAM), hükümet programları ve kalkınma planlarında yer alan stratejik öncelikler doğrultusunda ekonomik, sosyal ve kültürel her türlü araştırma yaparak, komşu ve üçüncü ülkelerle kalkınma tecrübelerini paylaştığı teknik yardım ve eğitimler vermekle yükümlüdür.

Bakanlık Araştırma Fonları

Bilim, Sanayi ve Teknoloji Bakanlığı, üniversite-sanayi işbirliğini geliştirmeyi ve başta KOBİ'ler olmak üzere Ar-Ge çalışmalarına verilen desteği artırıp, sanayinin teknolojik altyapısını güçlendirerek rekabet gücünü artırmaya yönelik çalışmalar yürütmektedir (<http://www.sanayi.gov.tr/>). Bu bakanlık, Teknoloji Geliştirme Bölgeleri ve üniversitelerle işbirliği içerisinde endüstri sektörü için teknolojik çözümler üreten *Sanayi Tezleri Programı*'ndan (SAN-TEZ <http://sagm.sanayi.gov.tr/ServiceDetails.aspx?dataID=108>) sorumlu kuruluşur.

Enerji politikalarına, arz güvenliğine, yerli enerji teknolojileri ve endüstrisine hizmet edecek projeleri destekleme amacıyla **Enerji ve Tabii Kaynaklar Bakanlığı** (<http://www.enerji.gov.tr>) tarafından yürütülen Enerji Sektörü Araştırma Geliştirme Projeleri Destekleme Programı (EN-AR), 36 ay süreli projelere 2014 senesine kadar toplam 50 milyon TL hibe vermektedir. Aynı bakanlığın bünyesindeki Ulusal Bor Araştırma Enstitüsü (BOREN

<http://www.boren.gov.tr>) de bu alanda araştırma yapan kamu ve özel hukuk tüzel kişileri ile işbirliği yaparak bilimsel araştırmaları yapmak, yaptırmak, koordine etmek ve bu araştırmalara katkı sağlamaktadır.

Bütün bunların yanı sıra; Türk Patent Enstitüsü (TPE), telif haklarını onaylamak ve korumakla ilgili politikalar geliştirerek, ulusal inovasyon sistemi içerisinde faal bir telif hakkı ortamının oluşturulmasından; Türk Standartları Enstitüsü (TSE <http://www.tse.org.tr>), yeni ürün ve teknolojilerin geliştirilmesini düzenlemek için ürün ve hizmet standartları konusunda araştırmalar yapmaktan; Türkiye Teknoloji Geliştirme Vakfı (TTGV <http://www.ttgiv.org.tr/tr/ar-ge-proje-destekleri>) ise Ar-Ge projeleri için uzun dönemli kredi temin etmekten sorumludur.

Kalkınma Ajansları Destek Programları

Türkiye’de farklı bölgelerde kurulan kalkınma ajansları da (örn. Kuzey Doğu Anadolu Kalkınma Ajansı – KUDAKA <http://www.kudaka.org.tr>) mali ve teknoloji destekleri vermektedir. Bu konuda ayrıntılı bilgi bulunulan bölgenin kalkına ajansından alınabilir. 2015 yılı mali destek alt limiti 150 bin TL, üst limiti ise 500 bin TL olarak belirlenmiştir.

Uluslararası Fon Kuruluşları

Araştırma projesi desteği sağlayan başlıca uluslararası kuruluşlar, devletlerin desteğiyle hükümetlerarası yönetilen fon kuruluşları ile kâr amacı gütmeyen fon kuruluşları olarak ikiye ayrılmaktadır. Bunlardan Ar-Ge proje desteği bakımından en önemli uluslararası sivil kaynak, günümüzde yedincisi süregelenekte olan **Avrupa Birliği Çerçeve Programları**’dır (http://cordis.europa.eu/fp7/home_en.html). Temel amacı sosyal ve ekonomik kalkınmayı sağlamak üzere bilimsel ve teknolojik araştırma kapasitesini artırmak olan bu programların bütçesi, sırasıyla 1984-1987, 1987-1990, 1991-1994, 1994-1998, 1999-2002 ve 2002-2006 dönemlerini kapsamak üzere düzenli olarak 100 milyar €’dan 50 milyar €’ya kadar yükseltilmiştir.

Çerçeve Programları kapsamındaki **Marie Curie Araştırma Programları ve Bursları** (<http://ec.europa.eu/research/mariecurieactions/>) da araştırmacıların, mühendislerin, sanayi kuruluşlarının, üniversitelerin, araştırma merkezlerinin, laboratuvarların ve mükemmeliyet merkezlerinin yararlanabileceği programlar olup, 2007-2013 yılları arasında 4,7 milyar € destek sunmaktadır. *Marie Curie Programı* altında, yeni araştırma eğitim programlarının kurulması, araştırmacıların Avrupa içinde dolaşımı, tersine beyin göçünün önlenmesi, sanayi-akademi işbirliği ve araştırmacı istihdamı, Amerika, Japonya gibi gelişmiş ülkelerde araştırma imkanı ve araştırma merkezleri arasında personel değişimi gibi 10 farklı faaliyet için destek verilmektedir.

2007-2013 döneminde, Türkiye gibi AB'ye aday ülkelerin AB sürecinin ekonomik ve sosyal kalkımlarına katkıda bulunmayı amaçlayan **Katılım Öncesi Mali Yardım** (IPA: Instrument for Preaccession Assistance - http://europa.eu/legislation_summaries/agriculture/enlargement/e50020_en.htm) adlı yeni bir finansal işbirliği programı geliştirilmiştir. Bunların yanı sıra, AB üye ülkelerinin de, vatandaşlık şartı aramadan sağladıkları devlet ve sivil toplum kuruluşu (STK) fonları bulunmaktadır.

AB ile birlikte Ar-Ge fonlamasında en ön sırada olan Amerika Birleşik Devletleri'nin, uluslararası araştırmacıların da kullanımına açık olan Ar-Ge destek programlarının başında **National Science Foundation** (<http://www.nsf.gov/funding/aboutfunding.jsp>) ve **National Institute of Health** (<http://grants.nih.gov/grants/guide/>) gelmektedir. ABD hükümeti tarafından verilen araştırma, eğitim ve kariyer geliştirme amaçlı fonlara fen, tıp ve mühendislik dallarında faaliyet gösteren ABD'li ve ABD ile işbirliği içinde olan yabancı üniversiteler, yüksek okullar, araştırma merkezleri, özel araştırma şirket ve laboratuvarlarından yapılan başvuruların her yıl 10.000 kadarına fon sağlanmaktadır. ABD kaynaklı diğer önemli fon programları ise, **The Robert S. McNamara Fellowships Program (RSM)**, **International Research and Exchanges Boards (IREX)** ve **Ford Foundation** tarafından yürütülmektedir.

Öte yandan çeşitli uluslararası kuruluşlar da Ar-Ge projelerini desteklemektedir. Bunların başında gelen Birleşmiş Milletler (BM) ile BM bünyesindeki, **UN-Habitat** ve **UNESCO** gibi kuruluşlar, çevre, sağlık, enerji, eğitim, insan hakları ve sürdürülebilir kalkınma konularında çeşitli araştırma ve eğitim desteği vermektedir. Ayrıca, **NATO**'nun savunma ve güvenlik konularında sosyal bilimler alanındaki araştırma fonları ve burslarının dışında, **Science For Peace (SfP)** programı tarafından da temel bilimler ve mühendislik dallarında uygulamalı araştırma ve Ar-Ge projeleri fonlanmaktadır. BM ve NATO'nun yanı sıra, **Dünya Ticaret Örgütü (WTO)**, **Dünya Bankası**, **Uluslararası Para Fonu (IMF)** ve **Avrupa Kalkınma Bankası (EBRD)** gibi bölgesel kalkınma bankaları da Ar-Ge destek fonları sağlamaktadırlar.

Bölüm 54: Yazarlık Hakkı

Fatih AKÇAY

Bu bölümde okuyucuların bilimsel bir eserde yazar olmanın şartları hakkında bilgi sahibi olmaları amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Yazarlık için gereken asgari ölçütleri tartışabilme
- Onursal yazarlık ve sorumluluklarını tartışabilme
- Hayalet yazarlık ve sorumluluklarını tartışabilme
- Yazarlık sırası belirlemedeki ölçütleri tartışabilme
- Bir makalede teşekkür bölümünün özelliklerini tartışabilme

Bilimsel çalışmalarda yazarlık ve yazar sırası bilim dünyasında geçmişte ve günümüzde önem arz eden ve gündemdeki yerini koruyan bir konudur. Yazarlık hakkın nedeniyle ciddi tartışmalar, hatta suçlamalar ve mahkemeler olabilmekte, yerel ve ulusal etik kurullara bu konuda sürekli şikayetler olmaktadır. Özellikle alt kademdeki bilim insanları genellikle bu konuda haklarının yendiğini, çalışmalara katkıları olduğu halde yazılmadıklarını veya ancak son sıralarda yer bulabildiklerini ve bu durumdan memnun olmadıkları gözlenmektedir.

Dikkat çeken diğer bir nokta da 1950’li yıllara kadar makale başına düşen yazar sayısı ortalama 1-2 iken, bu sayı 2000’li yıllarda ortalama 4-5’e çıkmıştır. Günümüzde de eser başına düşen yazar sayısı yüksek olmaya devam etmekte ve bu konuyla ilgili problemler de süregelmektedir.

Sonradan mağdur olmamak için yazar adaylarının çalışmanın başlangıcında sıralarıyla beraber belirlenmesi ve görev tanımlarının yapılması bu tür problemleri önlemede en etkin yoldur. Dergi editörlerinin katkıda bulunan yazarlardan makaleyi tasarlama, veri toplama, analiz ve yazım aşamasında destek olduklarına dair teminat almaları da faydalı olabilir. Yazarların hangi işe ne kadar katkı sağladıkları isimleriyle birlikte açıkça belirtilmelidir. Ayrıca para ve materyal desteği de belirtilmelidir.

Yazarlık Ölçütleri

Yazarlık hakkı konusunda son zamanlara kadar net kriterler tanımlanamamıştır. Yazar olabilmenin ana şartı, ilgili çalışma için toplum nezdinde sorumluluk alabilmektir. “International Committee of Medical Journal Editors” (Uluslararası Tıp Dergileri Editörler Kurulu)’na göre yazarlık hakkı için **şu koşulların tamamının** sağlanması gerekmektedir

(<http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html#two>):

1. Çalışmanın planlanması, veri toplanması, verinin analizi ve yorumlanması konusunda belirgin katkıda bulunmak VE
2. Makalenin taslağını oluşturmada veya gözden geçirmede belirgin entellektüel katkı sağlamak VE
3. Yayın öncesinde makalenin son halini inceleyip onaylamak VE
4. Çalışmanın tüm aşamaları için sorumluluk almayı ve hesap verebilmeyi kabul etmek

Her bir yazar çalışmanın tamamına veya bir kısmına etkin bir şekilde katılmalıdır. Yazarlıkla ilgili kurallar kurumlar arasında farklılıklar arz edebilir. Kurum içerisinde kesin tanımlanmış kurallar olabileceği gibi, yazarlık hakkı kurum kültürü çerçevesinde yazılı olmadan da düzenlenebilir. Türkiye’de yazarlık hakkıyla ilgili ihtilafların olması halinde üniversiteler veya gerekirse Yükseköğretim Kurulu uhdesinde gerekli işlemler yapılmaktadır. Yazarlık kurallarının yanlış uygulanması akademik bir suç olup ciddi yaptırımları olabilir.

Ekonomi, ticaret, finans, parçacık fiziği ve matematik bilimlerinde yazarlar Hardy-Littlewood kuralı olarak bilinen alfabetik sıraya göre sıralanır (Örn. Acemoğlu, Akçay, Aktürk, Işık gibi). Sosyal bilimlerde ise Amerikan Psikoloji Birliği bir rehber yayınlamıştır. Buna göre, sadece makaleyi yazmak yazarlık hakkı için yeterli olmayıp buna ilave olarak hipotezi veya problemi formülize etmek, deneysel düzeneği yapılandırmak, istatistik analizi yapmak, sonuçları yorumlamak gibi kayda değer katkıların yapılması da gerekmektedir. Bir akademik kuruluşta yetkili olmak (dekan, başhekim, bölüm başkanı, anabilim dalı başkanı, müdür gibi) yazarlık hakkı için tek başına yeterli değildir.

Sağlık bilimlerinde ise, yukarıda bahsedilen üç koşulun birden sağlanması gerekmektedir. Araştırmaya maddi destek bulmak veya çalışma grubunun genel denetimini yapmak yazarlık hakkını doğurmaz. Bu nedenle birçok tıbbi dergi terim olarak yazarlık (authorship) ifadesi yerine katkıda bulunanlar (contributors) ifadesini kullanmaya başlamıştır.

Sağlık bilimlerinde yazarlık hakkı ihtilaflarının artmasının bir nedeni de bir makalede çok sayıda yazar olmasıdır. Günümüzde iletişim imkanlarının da artmasıyla çok merkezli araştırmalar fazlaca yapılabilmektedir. 1993 yılında New England Journal of Medicine’de yayınlanan bir makalede 15 farklı ülkeden 1081 hastanede 41021 hastadan veri toplanmış ve 972 yazarlı bir çalışma ortaya çıkmıştır.. Ayrıca öyle tek olguluk vaka sunumları var ki, yazar sayısı bir futbol takımının sayısından dahi fazladır.

Ancak yukarıda bahsedilen çok uzun listeli yazarlık durumu da bir çok eleştiriyi beraberinde getirmiştir. Diğer yandan ise bir çok alanda da alanında uzman kişilere ihtiyaç duyan çalışmalar olduğundan yazar sayısının artabileceği de göz ardı edilmemelidir.

Çok yazarlığa gidiş bilim insanlarının değerlendirilme şeklinin bir sonucu olarak ortaya çıkmıştır. Bu değerlendirme nicelik bakımından yayınlanan makale sayısı ve nitelik bakımından bu makalelere yapılan atıf sayısı ile belirlenmektedir. Bu iki yaklaşım tek yazarlı durumlar için uygundur; ancak çoklu yazarlar için problem oluşturur. Çünkü her bir yazar makaleyi veya atıfı tek başına sahiplenmektedir ki, bu da aslında yayınların ve atıfların sihirli bir şekilde yazar sayısı ile çarpılması demektir. Genellikle ödüllerin çok yazarlı yayınlara verilmesi de bu durumu teşvik etmektedir. Bu durum yayın ve atıf sayısının yazar sayısına bölünmesiyle realize edilebilir.

Onursal (Hediye) Yazarlık

Çalışmaya hiçbir katkısı olmayan veya katkısı olmakla birlikte yazarlar arasında yer almaya yetecek kadar katkısı olmayan kişilerin yazarlar arasında yer almasıdır. Genellikle bölüm veya bilim dalı başkanının isminin yazılması şeklinde görülür. Son zamanlara kadar Almanlarda departman veya bölüm başkanının (katkısı olsun veya olmasın) yazarlar listesine eklenmesi standart bir davranıştı. Bununla birlikte ABD Ulusal Bilim Akademisi, bu tür durumların yayının kalitesini sulandıracağını belirtmektedir.

Bu kişiler (özellikle büyük araştırma gruplarını sevk ve idare edenler) saygınlıklarını ve akademik ünvanlarını uzun yayın listelerinden dolayı elde ettikleri için hediye yazarlıktan kolay kolay feragat edememektedirler. Bunun önüne geçebilmek için alanında saygın dergilerin önemli bir kısmı her makale için yazarların katkısı ve sorumluluklarını ayrıntılı olarak belirten belge istemektedirler. Bunun da ne kadar etkin olduğunu kestirmek zor olmakla birlikte kısmi caydırıcılığı olabilir. Bu uygulamanın ayrıca bilimsel sahtekarlık durumlarında failin belirlenmesine de katkısı olabilir.

Yazar Paslaşması

Bir diğer sık rastlanan kıyak yazarlık şekli ise, araştırmacıların yayınlarına birbirlerini karşılıklı olarak yazmalarıdır. Örneğin, üç kişinin üç ayrı çalışma yapıp birbirlerinin isimlerini kendi çalışmalarına yazması, böylece bir çalışma yapıp üç çalışmaya sahip olunması suretiyle yayın sayısı da üç misli artmış olur.

Hayalet Yazarlık

Kıyak yazarlığın tam tersidir. Burada yazar olmayı hak etmiş (bir çalışmanın planlanması, yapılması ve yazılması aşamalarına katılmış) bir kişinin isminin yazarlar arasında yer almaması durumudur. Bu, genellikle alt kademede ki

araştırmacıların başına gelen bir sorundur. Bu durum ilaç endüstrisiyle ilgili olan bazı çalışmalarda da rastlanabilmektedir.

İlgili ilaç firması bir şekilde makale yazma şirketlerine istedikleri makaleleri hazırlamaları için ücret ödüyor, daha sonra da bu makaleleri çalışmaya ciddi katkisi olmayan veya az katkisi olan akademisyenlerin isimlerinin ilave edilmesiyle yayınlanmak için dergilere gönderiliyor..

Bazen de yazarlar izinleri olmadan makalelere dahil edilebilirler. Burada genellikle alanında söz sahibi olan, bilinirliği olan ve saygınlığı olan bilim adamları yazar olarak eklenmektedir. Böylelikle ilgili yazının gönderildiği dergide kabul edilmesi ve yayınlanması ihtimali artmaktadır. Bu uygulama iyi niyetle olsa da yazarın sorumluluk almasına neden olacağından ileride problemlere yol açabilir.

Yazarlık Sırası nasıl olmalıdır?

Yazarlık sırasının belirlenmesinde aşağıdaki prensiplerin göz önünde bulundurulması faydalı olacaktır:

- Yazar isimleri ve sırası olabildiğince erken dönemde belirlenmelidir.
- Yazarlık sırası yazarların kendileri tarafından belirlenmelidir.
- Makaleye en fazla katkıyı veren kişi birinci sırada olmalıdır.
- Her bir yazarın esere olan katkısı belirtilmelidir.
- İletişimi sağlayan yazarın (corresponding author) özel bir sorumluluğu vardır.

Yazarlık Sorumlulukları

Bir yazarın bilmeden bir makalede isminin bulunması başına dert açabilir. Bu konudaki rehberler ortak yazarların da makaleden sorumlu olduklarını ve makalenin ana noktalarını anlayıp onaylamaları gerektiğini belirtmektedir. Konuyla ilgili dikkat çeken vakalardan birinde Amerikalı kök hücre araştırmacısı Gerald Schatten'in Hwang Woo-Suk ile ortak bir makalede adı geçmiştir. Sonraları ilgili makale hakkında intihal işlemi yapılmış, Schatten intihalden suçlu bulunmamıştır. Buna rağmen üniversitesi “adının geçtiği makaleleri takip etmemesi ve ihmalkar davranması” nedeniyle kendisini suçlamıştır. Ortak yazarlardan birisinin makalede yaptığı usulsüzlük nedeniyle diğer yazarların da suçlu bulundukları örnekler de vardır.

Yazarların sorumluluklarından biri de yayınlandıktan sonra dahi tüm araştırma verilerini incelenmek üzere hazır bulundurmaktır. Birçok bilimsel dergi de yazarların makale yayınlanmadan da gerektiğinde verileri incelenmek üzere sunmalarını şart koşar.

Teşekkür (Acknowledgment) kime yapılmalı

Yazarlık ölçütlerini karşılamayan ama makaleye katkıda bulunan herkes teşekkür bölümünde zikredilmelidir. Teşekkür edilecek kişilere örnek olarak, sadece teknik destek sağlayan, daktilo desteği sağlayan, makaleyi dil açısından kontrol eden kişiler veya sadece genel destek sağlayan bir anabilim dalı/bölüm başkanı veya yönetici verilebilir.

Bölüm 55: Dergi Seçimi

Zekeriya AKTÜRK

Bu bölümde okuyucuların bilimsel bir eseri yayınlanmak üzere göndermek için dergi seçimi hakkında bilgi sahibi olmaları amaçlanmıştır.

Dergi önerisi	http://jane.biosemantics.org/suggestions.php
	https://journalsuggester.springer.com/
ISI	http://mjl.clarivate.com/
Index of Open Access Fees	http://eigenfactor.org/projects/openAccess/oa.php?catid=2&pt=0
Ücretsiz dergiler	http://www.journals4free.com/
PubMed	https://www.ncbi.nlm.nih.gov/nlmcatalog
Ulakbim	https://trdizin.gov.tr/statistics/listAcceptedJournals.xhtml
Sahte Yayıncılar	https://web.archive.org/web/20170111172306/https://scholarlyoa.com/publishers/
Sahte Dergiler	https://predatoryjournals.com/journals/

Bölüm 56: Kendini Değerlendirme Çalışması

Değerli Okuyucumuz,

Bu sınavınızın amacı öğrendiklerimizi tekrar etmek, bilgilerimizi pekiştirmek ve eksiklerimizi fark edip gidermektir.

Lütfen testi kendinize zaman ayırarak yalnız başınıza, yardım almadan çözünüz. Takıldığınız yerlerde kitabınıza bakabilirsiniz. Arkadaşınızdan veya başka birisinden yardım almayınız.

Bu test için 2 saatlik bir sürenin yeteceğini düşünüyoruz. Takıldığınız sorular olursa onları boş bırakarak testi 2 saatte bitirmeye çalışınız.

Testteki sorular için <https://drive.google.com/open?id=0B0ieO-Fy3z-5WUMxX2ZnWUZTTnM> adresindeki SPSS veri setini kullanınız.

Sınav süresi sonunda sayfa 444'daki "Kendini Değerlendirme Çalışması Çözümleri" bölümüne bakarak sınavınızı puanlandırınız.

Başarılar...

Tarih:

Sınav puanınız:

Adınız Soyadınız:

Yaşınız:

Mesleğiniz:

Cinsiyetiniz:

Sorular:

1. Verileri nasıl sınıflandırırsınız? Ordinal veriye örnek veriniz.
3 puan
2. ‘Oran’ ve ‘hız’ ne demektir? Örnek verebilir misiniz?
2 puan
3. SPSS veri setinizdeki yaş (age) değişkenini veri hatası ve uç değerler açısından inceleyiniz. Gerekli düzeltmeleri yapınız. Bulduğunuz hataları ve yaptığınız değişiklikleri yazınız.
4 puan
4. Saplı kutu grafikleri (asansör grafikleri veya box plot) değişkenin hangi özellikleri hakkında bilgi verir? SPSS veri setinizi kullanarak yaş (age) değişkeninin saplı kutu grafiğini çıkarınız ve grafikteki bulguları açıklayınız.
5 puan
5. Değişkenin yaygınlığı (dispersion, spread) hakkında bilgi veren ölçütler nelerdir?
5 puan
6. Bildiğiniz teorik dağılımları yazınız.
4 puan
7. Başarı/başarısızlık ihtimaline dayanan, bir demir para atıldığında yazı gelme ihtimali veya IVF tedavisi gören bir çiftte başarılı gebelik oluşması ihtimali gibi veriler nasıl bir dağılım gösterir (hangi teorik dağılıma uyar)?
5 puan
8. Standart normal dağılım eğrisinde;
 - a. Ortalama kaçtır?
 - b. Verilerin %95’inin bulunduğu standart sapma aralığı kaçtır?
 - c. Verilerin %99’unun bulunduğu standart sapma aralığı kaçtır?3 puan

9. Verilere neden dönüşüm (transformation) uygularız? Logaritmik dönüşüm verinizin dağılımında nasıl bir etki yapar?
4 puan
10. Aşağıda ilimizdeki çeşitli hastanelerin günlük hasta taburcu sayıları ve yatak kapasiteleri verilmiştir. Hastanelerin kapasitesine göre ağırlıklandırılmış günlük taburcu sayılarını hesaplayınız.
5 puan
11. Verilen SPSS veri setinde ‘glucose’ değişkeninin normal dağılıp dağılmadığını test ediniz? Nasıl test ettiğinizi ve sonucunuzu yazınız.
5 puan
12. On birinci soruda değişkeninizin normal dağılmaması halinde normal dağılıma uymasını sağlayacak uygun bir dönüşüm uygulayınız ve dağılımınızı tekrar test ediniz. Yaptığınız işlemleri açıklayınız ve sonucunuzu yazınız.
8 puan
13. Bir örnekleme bireylerin yaş ortalamalarını ve %95 güven aralığını hesapladık. Ortalamanın 27,01 yıl, güven aralığının da 25,54-28,48 yıl olduğunu bulduk. Yakın zamanda yapılan bir nüfus sayımında bölgedeki toplumun nüfusunun yaş ortalamasının 26,8 olduğunu biliyoruz. Örnekleminizin yaş ortalaması ile toplumunkini karşılaştırınız?
5 puan
14. Bir araştırmada iki ilacın etkinliği karşılaştırılmıştır. A ilacı verilen 100 hastadan 25’i iyileşirken B ilacı verilen 200 hastadan 20’si iyileşmiştir. İki ilacın etkinliği arasındaki farkı %95 güven aralığını da hesaplayarak yorumlayınız.
8 puan
15. Bir tıp fakültesi nöroloji polikliniğinde baş ağrısı polikliniği bulunmaktadır. Doktorlar hipotiroidisi olan hastaların daha fazla baş ağrısı çektiğinden şüpheleniyorlar. Poliklinikte kayıtlı 500 hastayı TSH’sı yüksek ve normal olanlar olarak ayırıyorlar. TSH’sı yüksek olan 60 hastayı randomize olarak 30’ar kişilik iki gruba ayırıyorlar. Bu gruplardan birisine tiroid hormonu veriyorlar diğer gruba ise plasebo

veriyorlar. (Hastaların baş ağrısı tedavisinde herhangi bir değişiklik yapılmıyor.) Hastalar haftalık baş ağrısı günlüğü ile takip ediliyor. 6 ay sonra her iki gruptaki baş ağrısı durumları karşılaştırılıyor.

Bu araştırma, hangi araştırma tasarımına uymaktadır?

3 puan

- a. Randomize kontrollü deney
- b. Kohort
- c. Vaka kontrol çalışması
- d. Klinik kohort
- e. Kesitsel tanımlayıcı çalışma

16. Bir serviste 6 kızamık ve 4 astım hastası yatmaktadır. Bu servisten sırayla (rasgele) iki hasta seçsek ikisinin de astımlı olma olasılığı nedir?
3 puan

17. Verilen SPSS veri setini kullanarak aşağıdaki hipotezi test ediniz. Hangi testi kullandığınızı ve sonuçlarınızı yazınız.

H₀: Erkek ve kadınların (sex) diyabet süreleri (duration) arasında anlamlı bir fark yoktur.

7 puan

18. Verilen SPSS veri setini kullanarak aşağıdaki hipotezi test ediniz. Hangi testi kullandığınızı ve sonuçlarınızı yazınız.

H₀: Çeşitli medeni durum grupları (marital) arasında vücut kitle endeksi (BMI) açısından istatistiksel olarak anlamlı bir fark yoktur (BMI'nin normal dağılmadığına dikkat ediniz).

7 puan

19. Verilen SPSS veri setini kullanarak aşağıdaki hipotezi test ediniz. Hangi testi kullandığınızı ve sonuçlarınızı yazınız.

H₀: Örneklemimizdeki diyabetik hastaların sigara içicilik oranı (smoking) toplumunkinden farklı değildir (Toplumda sigara içme yaygınlığının %25 olduğu bilinmektedir).

7 puan

20. Verilen SPSS veri setini kullanarak aşağıdaki hipotezi test ediniz. Hangi testi kullandığınızı ve sonuçlarınızı yazınız.

H₀: Cinsiyet (sex) açısından hastalara verilen tedavi çeşitleri (therapy) arasında istatistiksel olarak anlamlı bir fark yoktur.

7 puan

Bölüm 57: Kendini Değerlendirme Çalışması Çözümleri

1. Verileri nasıl sınıflandırırsınız? Ordinal veriye örnek veriniz.
3 puan

Nümerik (sürekli ve kesikli) ve kategorik (ordinal ve nominal) veriler.

(2 puan)

Ordinal veriler belli bir sıra izler. Eğitim durumu, sosyoekonomik durum gibi. **(1 puan)**

2. ‘Oran’ ve ‘hız’ ne demektir? Örnek verebilir misiniz?
2 puan

Oran ve hızın farkı hızın zamanı içermesidir. Erkek ve kızların sigara içme oranı (sigara içen erkekler/sigara içen kızlar). **(1 puan)** Bebek ölüm hızı (Bir yılda 0-28 gün arası ölen bebekler/o yıldaki tüm canlı doğumlar) **(1 puan)**

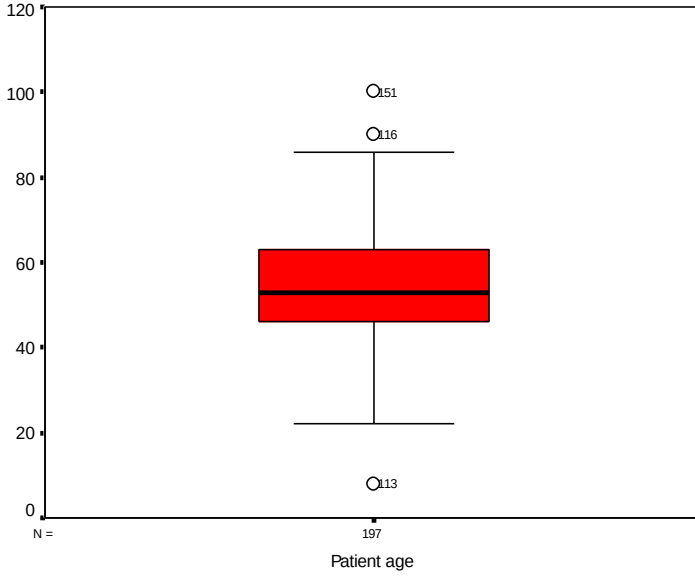
3. SPSS veri setinizdeki yaş (age) değişkenini veri hatası ve uç değerler açısından inceleyiniz. Gerekli düzeltmeleri yapınız. Bulduğunuz hataları ve yaptığınız değişiklikleri yazınız.
4 puan

Veri age değişkenine göre sıralanabilir veya frekansı alınabilir **(2 puan)**
8 yaşında (113 numaralı birey), 90 (116 numaralı birey) , 100 (151 numaralı birey) ve 165 (110 numaralı birey) yaşlarında bireyler dikkat çekicidir. 165 kontrol edilip yeniden girilmelidir. Ham veriler elimizde olmadığından 165 boş bırakıldı. **(2 puan)**

(Saplı kutu (Boxplot) çizene fazladan 2 puan)

4. Saplı kutu grafikleri (asansör grafikleri veya box plot) değişkenin hangi özellikleri hakkında bilgi verir? SPSS veri setinizi kullanarak yaş (age) değişkeninin saplı kutu grafiğini çıkarınız ve grafikteki bulguları açıklayınız.
5 puan

Medyan, %95 güven aralığı, interquartile range (%25-75 arası), minimum ve maksimum değerler ve uç değerler (outliers) **(herhangi üçünü belirtene 2 puan).**



1. Medyan 50 yaş civarında,
 2. %95 aralık 25-85 yaş arasında,
 3. Interquartile range 45-65 arasında,
 4. 113, 116 ve 151. bireyler uç değer oluşturuyor.
- (Şekli çizdikten sonra herhangi ikisini belirtene 3 puan)**

5. Değişkenin yaygınlığı (dispersion, spread) hakkında bilgi veren ölçütler nelerdir?
5 puan

Range, varyans, standart sapma, persantil. **(Her birisi 2 puan. Üç ve fazla belirtene 5 puan)**

6. Bildiğiniz teorik dağılımları yazınız.
4 puan

Normal dağılım (Gauss dağılımı), t dağılımı, ki kare dağılımı, F dağılımı, binomiyal dağılım, Poisson dağılımı **(Her birisi 1 p. Dört veya fazla belirtene 4 puan)** (istatistikte yüzlerce teorik dağılım olduğunu belirtme ek 1 p.)

7. Başarı/başarısızlık ihtimaline dayanan, bir demir para atıldığında yazı gelme ihtimali veya IVF tedavisi gören bir çiftte başarılı gebelik oluşması ihtimali gibi veriler nasıl bir dağılım gösterir (hangi teorik dağılıma uyar)?

5 puan

Binomiyal dağılım (5 puan)

8. Standart normal dağılım eğrisinde;

- Ortalama kaçtır?
 - Verilerin %95'inin bulunduğu standart sapma aralığı kaçtır?
 - Verilerin %99'unun bulunduğu standart sapma aralığı kaçtır?
- 3 puan

Ortalama=0 (1 puan)

%95 güven aralığı $\pm 1,96$ x standart sapma (1 puan)

%99 güven aralığı $\pm 2,58$ x standart sapma (1 puan)

9. Verilere neden dönüşüm (transformation) uygularsınız? Logaritmik dönüşüm verinizin dağılımında nasıl bir etki yapar?

4 puan

Teorik dağılımla ilgili varsayımları karşılamak için (1 puan)

Varyansı sabitlemek (azaltmak) için (1 puan)

İki değişken arasındaki ilişkiyi lineer hale getirmek için (1 puan)

Logaritmik dönüşüm sağa eğimli bir dağılımı normal dağılıma yaklaştırır. Değişkenler arasında üstlü bir ilişki varsa ilişkiyi lineer hale getirir. Varyansları eşitler. (herhangi birisini belirtene 1 puan)

10. Aşağıda ilimizdeki çeşitli hastanelerin günlük hasta taburcu sayıları ve yatak kapasiteleri verilmiştir. Hastanelerin kapasitesine göre ağırlıklandırılmış günlük taburcu sayılarını hesaplayınız.

5 puan

Hastane	Yatak kapasitesi	Günlük taburcu	Ağırlıklı taburcu
Devlet hastanesi	200	40	
Göğüs hastanesi	100	10	
Tıp fakültesi	600	80	
Özel hastane	50	10	
Toplam			

Hastane	Yatak kapasitesi	Günlük taburcu	Ağırlıklı taburcu
Devlet hastanesi	200	40	29 (200*140/950)
Göğüs hastanesi	100	10	14 (100*140/950)
Tıp fakültesi	600	80	88 (600*140/950)
Özel hastane	50	10	7 (50*140/950)
Toplam	950	140	

Ağırlıklı taburcu ortalaması:

$$(200 \times 40 + 100 \times 10 + 600 \times 80 + 50 \times 10) / (200 + 100 + 600 + 50) = 60,5$$

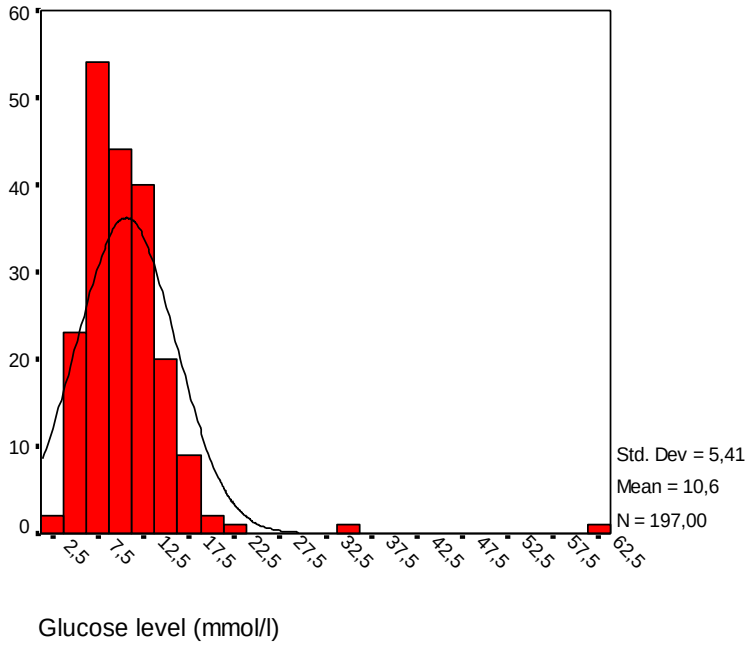
Genel ağırlıklı ortalamayı (60,5) bulana 3 puan

Tek tek hastanelerin ağırlıklı günlük taburcularını yazmak 3 puan

Tabloda toplam satırını gösterenlere ek 2 puan.

11. Verilen SPSS veri setinde 'glucose' değişkeninin normal dağılıp dağılmadığını test ediniz? Nasıl test ettiğinizi ve sonucunuzu yazınız.
5 puan

Normal dağılım grafiği çizilerek gösterilebilir. (2 puan)



Descriptive statistics'ten skewness bakılabilir. (2 puan)

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Dev.	Skewness		
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	
Glucose level (mmol/l)	197	2,00	61,70	10,6278	5,4098	4,831	,173	
Valid N (listwise)	197							

Normal dağılım eğrisinde kuyruğu sağa bakan bir eğim olduğundan sağa eğimlidir (right skewed). Skewness değeri (+4,831) standart hatasının (0,173) iki katından büyük olduğundan veri normal dağılmamaktadır. Değeri + olduğundan sağa eğimli olduğunu anlıyoruz.

(yukarıdakilerden herhangi birisini belirtene 3 puan)

(Hem skewness, hem de normal dağılım grafiği kullanana ek 2 puan)

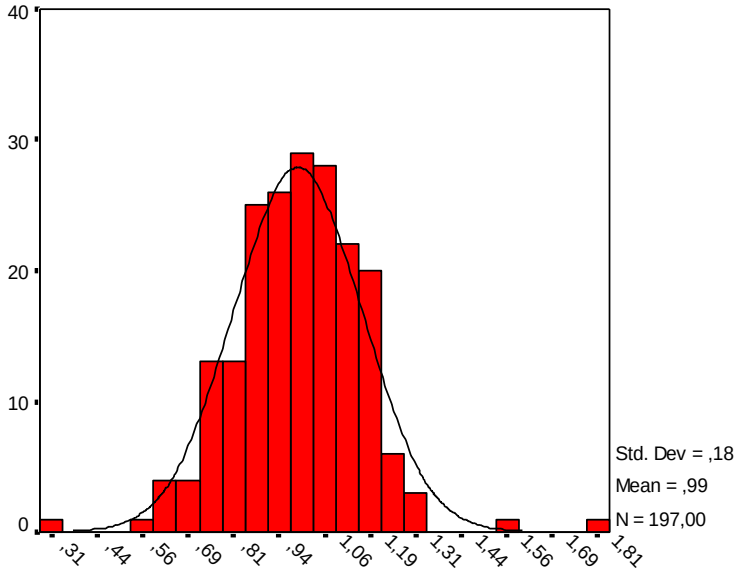
12. On birinci soruda değişkeninizin normal dağılmaması halinde normal dağılıma uymasını sağlayacak uygun bir dönüşüm uygulayınız ve

dağılımınızı tekrar test ediniz. Yaptığınız işlemleri açıklayınız ve sonucunuzu yazınız.

8 puan

Transform – Compute kullanılarak glucose değişkeninin logaritması alınarak [LG10(glucose)] yeni bir değişken (örn. GLUCLOG) oluşturulabilir.

Bu değişkenin histogram grafiğine bakıldığında normal dağıldığı görülmektedir.



GLUCLOG

Skewness değeri de standart sapmasının iki katından düşüktür:

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Dev.	Skewness		
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	
GLUCLOG	197	,30	1,79	,9890	,1757	,136	,173	
Valid N (listwise)	197							

Logaritmik dönüşüm uygulanacağını belirtilmesi 3 puan

Logaritmik dönüşümün yapılarak histogram veya skewness değerinin verilmesi 5 puan.

13. Bir örnekleimde bireylerin yaş ortalamalarını ve %95 güven aralığını hesapladık. Ortalamanın 27,01 yıl, güven aralığının da 25,54-28,48 yıl olduğunu bulduk. Yakın zamanda yapılan bir nüfus sayımında bölgedeki toplumun nüfusunun yaş ortalamasının 26,8 olduğunu biliyoruz. Örneklemimizin yaş ortalaması ile toplumunkini karşılaştırınız?

5 puan

Toplunun yaş ortalaması örneklemimizin yaş ortalamasının %95 güven aralığının içerisinde olduğundan örneklemimizin yaş ortalamasıyla toplumun yaş ortalaması arasında istatistiksel olarak anlamlı bir farkın olmadığı sonucuna varırız.

Toplunun yaş ortalamasının örneklemimizin %95 güven aralığının içerisinde olduğunun belirtilmesi 3 puan.

Aradaki farkın anlamsız olduğunun belirtilmesi 2 puan.

14. Bir araştırmada iki ilacın etkinliği karşılaştırılmıştır. A ilacı verilen 100 hastadan 25'i iyileşirken B ilacı verilen 200 hastadan 20'si iyileşmiştir. İki ilacın etkinliği arasındaki farkı %95 güven aralığını da hesaplayarak yorumlayınız.

8 puan

Confidence interval for the difference in proportions formülü kullanılmalı (pooled standart error kullanılarak): %95 GA = $p_1 - p_2 \pm 1.96\sqrt{((p_1 \times q_1/n_1) + (p_2 \times q_2/n_2))}$

Confidence interval for the difference in proportions (pooled standart error kullanılarak) formülü kullanılmalı (Kitapta sayfa 61):

$$\%95 \text{ GA} = p_1 - p_2 \pm 1.96\sqrt{((p_1 \times q_1/n_1) + (p_2 \times q_2/n_2))}$$

Oranlardaki farkın %95 güven aralığı

$$= 0,25 - 0,10 \pm 1.96\sqrt{((0,25 \times 0,75/100) + (0,10 \times 0,90/200))}$$

$$= 0,15 \pm 0,0945 = \%15 \pm \%9,45$$

$$= [5,55 - 24,45]$$

Yorum: iki ilacın etkinliği arasında %15'lik bir fark vardır. Bu fark %5 gibi düşük olabileceği gibi %24 gibi yüksek te olabilir. Her hâlükârda

aradaki fark sıfırdan büyüktür. Dolayısıyla istatistiksel olarak A ilacı B ilacından daha etkindir.

Güven aralığının doğru hesaplanması 6 puan

Yorumun doğru yapılması 2 puan.

15. Bir hastanede baş ağrısı polikliniği bulunmaktadır. Doktorlar hipotiroidisi olan hastaların daha fazla baş ağrısı çektiğinden şüpheleniyorlar. TSH'sı yüksek olan 60 hastayı randomize olarak 30'ar kişilik iki gruba ayırıyorlar. Bu gruplardan birisine tiroid hormonu veriyorlar diğer gruba ise plasebo veriyorlar. (Hastaların baş ağrısı tedavisinde herhangi bir değişiklik yapılmıyor.) Hastalar haftalık baş ağrısı günlüğü ile takip ediliyor. 6 ay sonra her iki gruptaki baş ağrısı durumları karşılaştırılıyor.
Bu araştırma, hangi araştırma tasarımına uymaktadır?
3 puan

- a. **Randomize kontrollü deney**
- b. Kohort
- c. Vaka kontrol çalışması
- d. Klinik kohort
- e. Kesitsel tanımlayıcı çalışma

16. Bir serviste 6 kızamık ve 4 astım hastası yatmaktadır. Bu servisten sırayla (rastgele) iki hasta seçsek ikisinin de astımlı olma olasılığı nedir?
3 puan

İlk hastanın astımlı olma olasılığı: 4/10
İkinci hastanın astımlı olma olasılığı 3/9

Birinci hastanın astımlı olması veya olmaması ikincinin astımlı olmasını engellemez, yani bu iki olasılık mutually exclusive değildir. Dolayısıyla çarpma kuralını uygulayacağız: $4/10 \times 3/9 = 12/90 = \%13,33$ **(3 puan).**

(Mutually exclusive kuralını hatırlayınız: Bir sepette 4 elma, 6 armut olsa elimizi daldırdığımızda elma **veya** armut alma olasılığımız kaç

olur? Elma alma olasılığı 4/10, armut alma olasılığı 6/10. Birisini aldığımızda ötekini alamayız. Dolayısıyla bunlar mutually exclusivedir. Yani ikisinden birini alma olasılığı $4/10 + 6/10 = 10/10$ olur.)

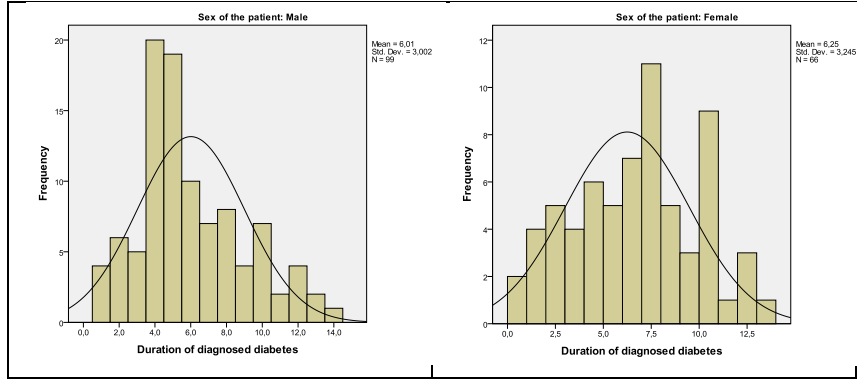
Mutually exclusive kuralını açıklayanlara ek 2 puan.

17. Verilen SPSS veri setini kullanarak aşağıdaki hipotezi test ediniz. Hangi testi kullandığınızı ve sonuçlarınızı yazınız.

H_0 : Erkek ve kadınların (sex) diyabet süreleri (duration) arasında anlamlı bir fark yoktur.

7 puan

Örneklem sayısı yeterlidir ve nümerik değişkenimiz histogram grafiğinde her iki grupta da normal dağılıma yakındır:



Bu durumda nümerik veri, iki grup, bağımsız: Independent samples t test uygulayabiliriz. Veya Skewness değerine göre normal dağılmadığına karar verip veri dönüştürme uygulayabilir veya t testinin alternatifi olan Mann Whitney U testini yapabiliriz.

Bağımsız örneklemelerde t testi sonucu:

Group Statistics

	Sex	N	Mean	Std. Dev	Std. Error Mean	
Duration of diabetes	Male	99	6,010	3,002	,302	
	Female	66	6,245	3,245	,399	

Independent Samples Test

	Levene		t-test						
	F	Sig.	t	Sig. (2-t)	Mean Dif	Std. Er. Dif	95% CI		
							Lower	Upper	
Equal variances	,869	,353	-,477	,634	-,235	,493	-1,208	,738	
Equal var not ass.			-,469	,640	-,235	,501	-1,225	,755	

Diyabetin süresi erkeklerde 6,01 yıl, kadınlarda 6,24 yıldır. Levene testinde $p=0,353$ olduğundan varyansların eşitliği ilkesi de karşılanmış oluyor.

t değeri - 0,477, p ise 0,634 olarak bulundu. H_0 kabul edilir.

Independent samples t test için varsayımların kontrol edilmesi ve belirtilmesi 3 puan.

Bağımsız örneklemelerde t testi veya Mann Whitney U testi yaparak H_0 'ın kabul edileceğinin belirtilmesi 4 puan.

18. Verilen SPSS veri setini kullanarak aşağıdaki hipotezi test ediniz. Hangi testi kullandığınızı ve sonuçlarınızı yazınız.

H_0 : Çeşitli medeni durum grupları (marital) arasında vücut kitle endeksi (BMI) açısından istatistiksel olarak anlamlı bir fark yoktur (BMI'nin normal dağılımadığına dikkat ediniz).

7 puan

Nümerik veri, ikiden fazla bağımsız grup. ANOVA için varsayımları karşılayamadığımızdan Kruskal Wallis varyans analizi yapmalıyız. Nonparametrik testlerden Kruskal Wallis Ki kare 0,840, $p=0,840$ olarak bulunmakta. Dolayısıyla H_0 kabul edilir.

Testin belirtilmesi 3 puan

Testin yapılması ve H_0 'ın kabul edileceğinin belirtilmesi 4 puan

19. Verilen SPSS veri setini kullanarak aşağıdaki hipotezi test ediniz. Hangi testi kullandığınızı ve sonuçlarınızı yazınız.

H_0 : Örneklemimizdeki diyabetik hastaların sigara içicilik oranı (smoking) %50'den farklı değildir.

7 puan

Kategorik veri, 1 grup: z testi yapılabilir. Bu veri binomiyal dağılım

göstermektedir. SPSS’te binomiyal testi kullanabiliriz. Bilgisayar grup 1 olarak sigara içmeyenleri aldığından test proportion olarak %50 girmeliyiz. $P < 0,001$ olduğundan H_0 reddedilir.

Testin belirtilmesi veya formülün yazılması 3 puan

Testin yapılması ve H_0 ’ın reddedileceğinin belirtilmesi 4 puan

20. Verilen SPSS veri setini kullanarak aşağıdaki hipotezi test ediniz. Hangi testi kullandığınızı ve sonuçlarınızı yazınız.

H_0 : Cinsiyet (sex) açısından hastalara verilen tedavi çeşitleri (therapy) arasında istatistiksel olarak anlamlı bir fark yoktur.

7 puan

Kategorik veri, iki kategori, ikiden fazla grup. Ki kare testi yapılmalı.

(Testin belirtilmesi 3 puan)

Therapy * Sex of the patient Crosstabulation

		Sex		Total	
		Male	Female		
No pharmacologic therapy	Count	6	4	10	
	Expected Count	6,0	4,0	10,0	
Sulphonylurea only	Count	56	42	98	
	Expected Count	59,1	38,9	98,0	
Biguanides only	Count	2	0	2	
	Expected Count	1,2	,8	2,0	
Sulphonylurea + Biguanides	Count	46	24	70	
	Expected Count	42,2	27,8	70,0	
Insulin	Count	7	9	16	
	Expected Count	9,6	6,4	16,0	
Insulin + oral antidiabetics	Count	3	0	3	
	Expected Count	1,8	1,2	3,0	
Total	Count	120	79	199	
	Expected Count	120,0	79,0	199,0	

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	
Pearson Chi-Square	6,388	5	,270	
Likelihood Ratio	8,117	5	,150	
Linear-by-Linear Association	,379	1	,538	
N of Valid Cases	199			

a 5 cells (41,7%) have expected count less than 5. The minimum expected count is ,79.

Hücrelerin %41,7'sinde beklenen değer 5'in altındadır. Bazı hücreleri birleştirmek gerekecek.

(Birleştirme gerektiğinin belirtilmesi 2 puan)

Gruplar farmakolojik tedavi almayanlar, oral tedavi alanlar ve insülin alanlar olarak birleştirilebilir (SPSS'te transform-recode).

(Herhangi bir şekilde birleştirme yapabilenlere ek 4 puan)

Bu durumda hücrelerin %16,7'sinde beklenen sayı 5'in altına düşer ve Ki kare yapılabilir.

Therapy * Sex of the patient Crosstabulation

			Sex		Total	
			Male	Female		
Therapy	No pharmacologic therapy	Count	6	4	10	
		Expected Count	6,0	4,0	10,0	
	Oral	Count	104	66	170	
		Expected Count	102,5	67,5	170,0	
	Insulin	Count	10	9	19	
		Expected Count	11,5	7,5	19,0	
Total		Count	120	79	199	
		Expected Count	120,0	79,0	199,0	

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	
Pearson Chi-Square	,522	2	,770	
Likelihood Ratio	,514	2	,774	
Linear-by-Linear Association	,296	1	,586	
N of Valid Cases	199			

a 1 cells (16,7%) have expected count less than 5. The minimum expected count is 3,97.

(H_0 'ın kabul edileceğinin belirtilmesi 2 puan)

Ek 1: İstatistik tablolar***Z tablosu (Standart normal dağılım)***

z	2-yönlü P-değeri
0.0	1.000
0.1	0.920
0.2	0.841
0.3	0.764
0.4	0.689
0.5	0.617
0.6	0.549
0.7	0.484
0.8	0.424
0.9	0.368
1.0	0.317
1.1	0.271
1.2	0.230
1.3	0.194
1.4	0.162
1.5	0.134
1.6	0.110
1.7	0.089
1.8	0.072
1.9	0.057
2.0	0.046
2.1	0.036
2.2	0.028
2.3	0.021
2.4	0.016
2.5	0.012
2.6	0.009
2.7	0.007
2.8	0.005
2.9	0.004
3.0	0.003
3.1	0.002
3.2	0.001
3.3	0.001
3.4	0.001
3.5	0.000

***t* tablosu**

<i>sd</i>	İki yönlü P değeri			
	0.10	0.05	0.01	0.001
1	6.314	12.706	63.656	636.58
2	2.920	4.303	9.925	31.600
3	2.353	3.182	5.841	12.924
4	2.132	2.776	4.604	8.610
5	2.015	2.571	4.032	6.869
6	1.943	2.447	3.707	5.959
7	1.895	2.365	3.499	5.408
8	1.860	2.306	3.355	5.041
9	1.833	2.262	3.250	4.781
10	1.812	2.228	3.169	4.587
11	1.796	2.201	3.106	4.437
12	1.782	2.179	3.055	4.318
13	1.771	2.160	3.012	4.221
14	1.761	2.145	2.977	4.140
15	1.753	2.131	2.947	4.073
16	1.746	2.120	2.921	4.015
17	1.740	2.110	2.898	3.965
18	1.734	2.101	2.878	3.922
19	1.729	2.093	2.861	3.883
20	1.725	2.086	2.845	3.850
21	1.721	2.080	2.831	3.819
22	1.717	2.074	2.819	3.792
23	1.714	2.069	2.807	3.768
24	1.711	2.064	2.797	3.745
25	1.708	2.060	2.787	3.725
26	1.706	2.056	2.779	3.707
27	1.703	2.052	2.771	3.689
28	1.701	2.048	2.763	3.674
29	1.699	2.045	2.756	3.660
30	1.697	2.042	2.750	3.646
40	1.684	2.021	2.704	3.551
50	1.676	2.009	2.678	3.496
100	1.660	1.984	2.626	3.390
200	1.653	1.972	2.601	3.340
5000	1.645	1.960	2.577	3.293

Ki Kare tablosu

<i>sd</i>	İki yönlü P değeri			
	0.10	0.05	0.01	0.001
1	2.706	3.841	6.635	10.827
2	4.605	5.991	9.210	13.815
3	6.251	7.815	11.345	16.266
4	7.779	9.488	13.277	18.466
5	9.236	11.070	15.086	20.515
6	10.645	12.592	16.812	22.457
7	12.017	14.067	18.475	24.321
8	13.362	15.507	20.090	26.124
9	14.684	16.919	21.666	27.877
10	15.987	18.307	23.209	29.588
11	17.275	19.675	24.725	31.264
12	18.549	21.026	26.217	32.909
13	19.812	22.362	27.688	34.527
14	21.064	23.685	29.141	36.124
15	22.307	24.996	30.578	37.698
16	23.542	26.296	32.000	39.252
17	24.769	27.587	33.409	40.791
18	25.989	28.869	34.805	42.312
19	27.204	30.144	36.191	43.819
20	28.412	31.410	37.566	45.314
21	29.615	32.671	38.932	46.796
22	30.813	33.924	40.289	48.268
23	32.007	35.172	41.638	49.728
24	33.196	36.415	42.980	51.179
25	34.382	37.652	44.314	52.619
26	35.563	38.885	45.642	54.051
27	36.741	40.113	46.963	55.475
28	37.916	41.337	48.278	56.892
29	39.087	42.557	49.588	58.301
30	40.256	43.773	50.892	59.702
40	51.805	55.758	63.691	73.403
50	63.167	67.505	76.154	86.660
60	74.397	79.082	88.379	99.608
70	85.527	90.531	100.43	112.32
80	96.578	101.88	112.33	124.84
90	107.57	113.15	124.12	137.21
100	118.50	124.34	135.81	149.45

F tablosu

PaydaınınSD	2-yönlü P-değeri	1-yönlü P-değeri	Payın serbestlik derecesi													
			1	2	3	4	5	6	7	8	9	10	15	25	500	
1	0.05	0.025	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.6	963.3	968.6	984.9	998.1	1017.0	
1	0.10	0.05	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	245.9	249.3	254.1	
2	0.05	0.025	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.43	39.46	39.50	
2	0.10	0.05	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.43	19.46	19.49	
3	0.05	0.025	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.25	14.12	13.91	
3	0.10	0.05	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.70	8.63	8.53	
4	0.05	0.025	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.66	8.50	8.27	
4	0.10	0.05	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.86	5.77	5.64	
5	0.05	0.025	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.43	6.27	6.03	
5	0.10	0.05	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.62	4.52	4.37	
6	0.05	0.025	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.27	5.11	4.86	
6	0.10	0.05	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	3.94	3.83	3.68	
7	0.05	0.025	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.57	4.40	4.16	
7	0.10	0.05	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.51	3.40	3.24	
8	0.05	0.025	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.10	3.94	3.68	
8	0.10	0.05	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.22	3.11	2.94	
9	0.05	0.025	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.77	3.60	3.35	
9	0.10	0.05	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.01	2.89	2.72	
10	0.05	0.025	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.52	3.35	3.09	
10	0.10	0.05	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.85	2.73	2.55	
15	0.05	0.025	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.86	2.69	2.41	
15	0.10	0.05	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.40	2.28	2.08	
20	0.05	0.025	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.57	2.40	2.10	
20	0.10	0.05	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.20	2.07	1.86	
30	0.05	0.025	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.31	2.12	1.81	
30	0.10	0.05	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.01	1.88	1.64	
50	0.05	0.025	5.34	3.97	3.39	3.05	2.83	2.67	2.55	2.46	2.38	2.32	2.11	1.92	1.57	
50	0.10	0.05	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03	1.87	1.73	1.46	
100	0.05	0.025	5.18	3.83	3.25	2.92	2.70	2.54	2.42	2.32	2.24	2.18	1.97	1.77	1.38	
100	0.10	0.05	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97	1.93	1.77	1.62	1.31	
1000	0.05	0.025	5.04	3.70	3.13	2.80	2.58	2.42	2.30	2.20	2.13	2.06	1.85	1.64	1.16	
1000	0.10	0.05	3.85	3.00	2.61	2.38	2.22	2.11	2.02	1.95	1.89	1.84	1.68	1.52	1.13	

Pearson korelasyon katsayısı

Ornekl. sayısı	İki yönlü P değeri		
	0.05	0.01	0.001
5	0.878	0.959	0.991
6	0.811	0.917	0.974
7	0.755	0.875	0.951
8	0.707	0.834	0.925
9	0.666	0.798	0.898
10	0.632	0.765	0.872
11	0.602	0.735	0.847
12	0.576	0.708	0.823
13	0.553	0.684	0.801
14	0.532	0.661	0.780
15	0.514	0.641	0.760
16	0.497	0.623	0.742
17	0.482	0.606	0.725
18	0.468	0.590	0.708
19	0.456	0.575	0.693
20	0.444	0.561	0.679
21	0.433	0.549	0.665
22	0.423	0.537	0.652
23	0.413	0.526	0.640
24	0.404	0.515	0.629
25	0.396	0.505	0.618
26	0.388	0.496	0.607
27	0.381	0.487	0.597
28	0.374	0.479	0.588
29	0.367	0.471	0.579
30	0.361	0.463	0.570
35	0.334	0.430	0.532
40	0.312	0.403	0.501
45	0.294	0.380	0.474
50	0.279	0.361	0.451
55	0.266	0.345	0.432
60	0.254	0.330	0.414
70	0.235	0.306	0.385
80	0.220	0.286	0.361
90	0.207	0.270	0.341
100	0.217	0.283	0.357
150	0.160	0.210	0.266

Spearman korelasyon katsayısı

Ornekl. sayısı	İki yönlü P değeri		
	0.05	0.01	0.001
5	1.000		
6	0.886	1.000	
7	0.786	0.929	1.000
8	0.738	0.881	0.976
9	0.700	0.833	0.933
10	0.648	0.794	0.903

Ek 2: Alıştırmaların Çözümleri

Not: alıştırmaların çözümleri istatistik kurslarında katılımcıların verdikleri cevaplardan derlenmiştir. Bu cevaplar düzeltme ve eleştiriye açıktır.

Çözüm 1 (Sayfa 58)

Araştırma anketinin aşağıdaki şekilde değiştirilmesi önerilebilir:

Diyabet Araştırma Anketi

Hasta no:

S1. Doğum tarihiniz:.....

S2. Cinsiyetiniz:

1 () Kadın

2 () Erkek

S3. Kaç yıldır diyabet hastasıdır?:.....

S4. En son HbA1c değerinizi kaçtı?..... 99 () Bilmiyorum

S5. Son iki hafta içerisinde hastalığınız nedeniyle günlük işlerinizi yapmakta ne kadar zorlandınız?

1 () Hiç zorlanmadım

2 () Biraz zorlandım

3 () Epey zorlandım

4 () Çok zorlandım

5 () Hiçbir iş yapamadım

S6. Diyabet için hangi ilaçları kullanıyorsunuz? Yazınız.

1.....

2.....

3.....

4.....

5.....

6.....

S7. Muayenenizde diyabete bağlı herhangi bir komplikasyon saptandı mı?

0 () Hayır

1 () Evet

Yapılabilecek yorumlar:

- 1) Anket numarası / birey numarası her veri setinde mutlaka bulunmalıdır.
- 2) Kişilerin kaç yaşında olduğunu değil doğum tarihlerini sormak daha güvenilir bilgi verebilir
- 3) Kaç yıldır diyabet hastası olduğunu sormak yerine hangi tarihte tanı konulduğunu sormak yine daha güvenilir bilgi verebilir.
- 4) 6. soruda 1-diyet, 2-diamcron, 3-glukofaj, 4-insülin şeklinde kategorize edebileceğimiz gibi bu soruyu 8, 9, 10, 11. sorular şeklinde yapabilir, diyet kullanıyor musunuz (0-Hayır, 1-Evet), Glukofaj kullanılıyor musunuz (0-Hayır, 1-Evet) vb. şeklinde de kodlayabiliriz.

Örnek SPSS verisetini http://www.aile.net/agep/istat/09_10/odev1.sav adresinden indirebilirsiniz.

Bu ödevin değerlendirilmesinde aşağıdaki noktalar dikkate alınabilir:

1. Anket üzerinde değişiklik yapmış ve değişkenlerin kategorilerine harf yerine numara vermiş mi?
2. Veri setini hazırlayabilmiş ve SPSS'te tüm değişkenler için alanlar oluşturabilmiş mi?
3. Anket/birey numarası alanı oluşturmuş mu?
4. Komplikasyon, tedavi gibi kategorik değişkenlerde Decimal alanı sıfır olarak girilmiş mi?
5. Kategorik değişkenlerin etiketleri kodlanabilmiş mi?
6. S4 için “Bilmiyorum” seçeneği “99-Missing” olarak kodlanabilmiş mi?
7. Çoklu seçenek içeren ilaç kullanımı sorusu doğru olarak kodlamış mı?
8. Değişkenlerin sınıflandırılmasını (Measure) SPSS'te doğru yapmış mı?
9. Anket verileri doğru olarak SPSS'e girilebilmiş mi?

Çözüm 2 (Sayfa 66)

SPSS'te Variable View'de

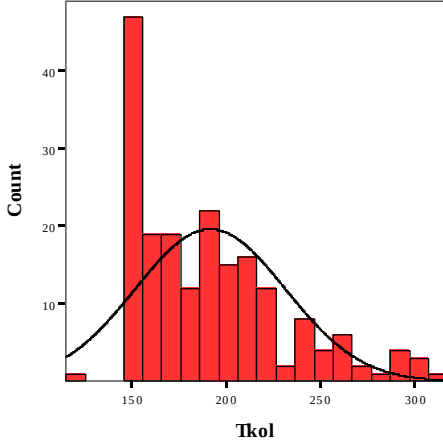
- 1- Anket ve soru numarası verilebilirdi.
- 2- Yaş nümerik bir değişkendir. Kategorize edilmemeliydi.
- 3- Income nümerik bir değişkendir. Kategorize edilmemeliydi.
- 4- Age yerine doğum tarihinin sorulması daha uygun olurdu.
- 5- Decimal değerleri 0 girilmiş. Real age için 1 olabilirdi.
- 6- Real Age için etiket (label) girilmemiş.
- 7- Hiç missing value gösterilmemiş.
- 8- Etnik kökende measure'a Scale yazılı, Nominal olmalıydı
- 9- Life spent in Saudi Arabia (Arabistan'da geçirilen zaman) nümerik bir değişkendir. Kategorize edilmemeliydi.
- 10- Arabistan'da geçirilen zamana ordinal yerine nominal yazılmış
- 11- Din (Religion) seçeneğine "4=Diğer" değeri eklenebilirdi.

SPSS'te Data View'de

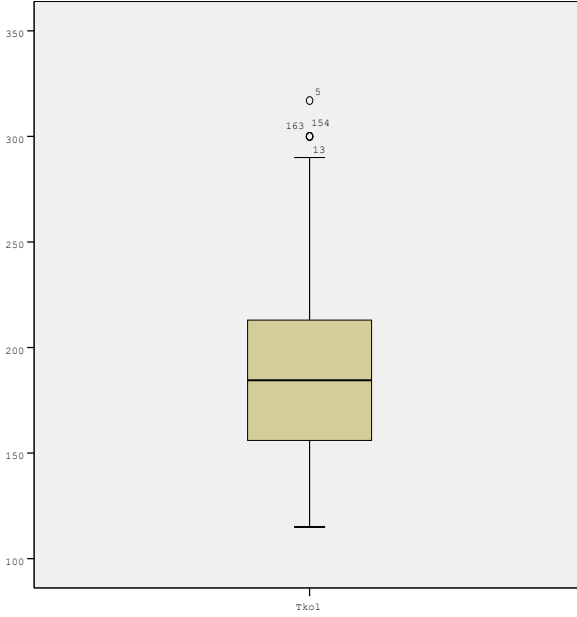
- 1- Education'da bir adet 11 değeri var
- 2- Income'da üç adet 4 değeri girilmiş
- 3- Life'da iki adet 4 değeri girilmiş
- 4- Knowledge'da bir adet 11 değeri girilmiş
- 5- Real age başlığında 16 tane missing değer var. Bunların dışında yaş dağılımı ise uç değer içermiyor.

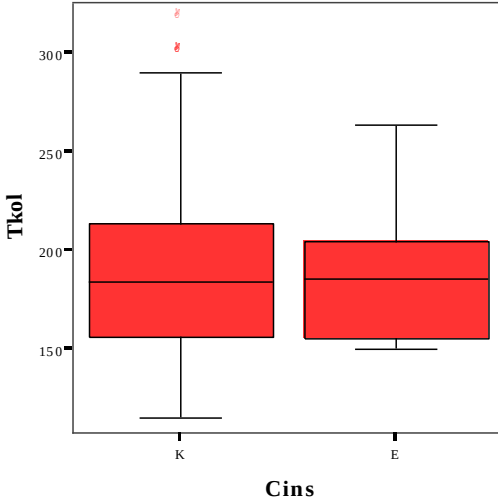
Çözüm 3 (Sayfa 90)

1- Tkol değişkeninin histogram grafiği:



Tkol değişkeninin saplı kutu grafikleri:





2-Tkol histogramı sağa eğilimli görülüyor, boxplotta iki adet uç değer mevcut, ikisi de kadınlarda.

Ortanca 180 civarında (K ve E için neredeyse aynı).

%25 ve 75'lik aralık 155 ile 215 arası.

Aşırı değer yok.

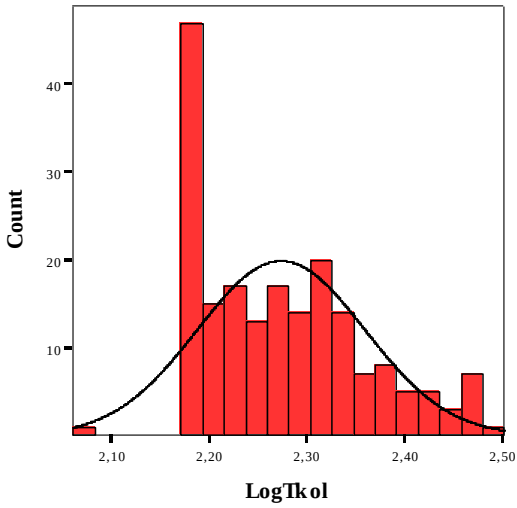
En küçük değer 115, en büyük değer 290 civarında

Erkeklerde en küçük değer 150, en büyük değer 220 civarında

Kadınlarda en küçük değer 115, en büyük değer 290 civarında

3- SPSS'te değişken x yani logTkol oluşturduk

4- logTkol değişkeninin histogram grafiği:



Şimdi dağılım daha normale yakın oldu.

5-tkol AO=190,92

6-tkol GO=187,11

7-Kilo AO=71,50; Ortanca (Medyan)=70,00; Mod=67

Statistics		
Kilo		
N	Valid	194
	Missing	0
Mean		71.50
Median		70.00
Mode		67

8-71,59

Statistics		
Kilo		
N	Valid	30562
	Missing	0
Mean		71.59
Median		70.00
Mode		67

9-E=25, K=169,

Boy AO=157,54cm,

Boy cinsiyete göre ağırlıklandırılmış AO=158,73cm

Çözüm 4 (Sayfa 102)

- 1- Bu durumda 20 depresyonlu, 20 psikozlu, 60 başka tanılı hasta var. Böylece depresyonlu veya psikozlu hasta olasılığı $\%20 + \%20 = \%40$ olacaktır.
- 2- $4/10 \times 3/9 = 12/90 = \%13,3$.
- 3- $Z = (60-50)/5 = 2$. $Z=2$ için iki yönlü tablo p değeri = 0,046. Çan eğrisinin sağ tarafındaki bireyler geçmiş olacağından iki yönlü p değerinin yarısı yani $\%2,3$ 'ü sınıfta kalacak. 1000 öğrenciden 977'si geçer
- 4-
 - a. $z = (x - \mu) / \sigma$ formülüne göre $150 - 130 / 10 = 2$ yani 2SD'na denk gelmektedir, bu da $\%95$ 'lik kısma denk geldiği için bu kişi toplumun $\%95$ 'lik sınırındadır ($z=2$ için tablo p değeri 0,046. $100 - 4,6 = 95,4$. Dolayısıyla sınırın hafif dışında).
 - b. $145 - 130 / 10 = 1,5$. Tablodan baktığımızda $z=1,5$ değeri için iki yönlü p değerinin 0,134 olduğunu buluruz. $100 - 13,4 = 86,6$ olduğundan bu bireyin toplumun $\%95$ 'inin içerisinde, yani normal olduğunu söyleyebiliriz.

Çözüm 5 (Sayfa 125)

- 1-
 - a. Yaş, boy, kilo, BKİ, B/K oranı, triseps deri altı kalınlığı, kan basıncı, ve kanda bakılan tüm değişkenler nümerik sürekli;
 - b. kaybedilen akraba, anne, kardeş nümerik kesikli;
 - c. cinsiyet, koroner arter prevalansı için kullanılan EKG ve hikaye delilleri, iş sahibi olma, kategorik nominal
 - d. sosyal sınıf, eğitim kategorik ordinal.
 - e. Ayrıca K/E oranı, albumin atılımı için hız, iş sahibi olma, sosyal sınıf, eğitim, kaybedilen akraba, anne, kardeş için de yüzdelere belirtilmiş (hepsi nümerik türetilmiş veriler).
- 2- Sistolik kan basıncı ve BKİ'yi için saçılma grafiği kullanılmış. İki nümerik veri arasındaki ilişki gösterilmiş ve bu iki grup için yapılmış.

Bu iki grup arasındaki fark ta lineer, eğimi farklı iki çizgi ile gösterilmiş.

- 3- **Aritmetik ortalama:** Tablo 1’deki tüm veriler, Tablo 2’deki ve bozuk OGTT ve DM oranları, Tablo 3’teki kolesterol değerleri.

Geometrik ortalama: Tablo 2’deki insülin, proinsülin değerleri ve c-peptid, Tablo 3’teki albumin atım hızları, Tablo 4’teki tüm veriler.

Ağırlıklandırma: Tablo 1’de kan basınçları, Tablo 2 ve 3’te tüm veriler, tablo 4’de VWF ve fibrinojen cinsiyete göre; F VII, PAI Aktivitesi ve PAI antijeni cinsiyete ve analitik yığına göre ağırlıklandırılmış. Ayrıca VWF obesite, sigara içme, mevcut koroner arter hastalığına göre, glikoz toleransı ve boy yaş, cinsiyet, kuşatma maruziyeti, sosyal sınıf ve eğitime göre de ağırlıklandırılmış.

Varyans analizi: tüm sürekli veriler için yapılmış.

Varyans oranı: Sistolik kan basıncı ile BKİ ilişkisinin gösterildiği saçılma grafiğinde kullanılmış.

- 4- Değişkenleri cinsiyete göre ağırlıklandırıp sonra analiz ettikleri anlaşılıyor.
- 5- Eğimli veriler için logaritmik veri dönüşümü, varyans analizi ile korelasyon ve lineer regresyon analizleri için kullanılmıştır. Bu dönüşümün sağa eğimli veriler için normalize, linearize ve varyans stabilize edici özelliğinden yararlanılıp veriler analize uygun hale getirilmiştir.
- 6- Değişkenlerin dağılımı konusunda okuyucuyu bilgilendirmek için standart sapmayı da göstermek iyi olurdu.
- 7- Standart sapma verilerin ortalamadan ne kadar uzakta olduğunu, SEM ise verilen ya da bulunan aritmetik ortalamanın toplundaki gerçek değeri ne kadar yansıttığını gösterir. SD yerine SEM kullanırsak, okuyucu verilerin gerçek dağılımı hakkında bilgi sahibi olamaz, verilerin daha dar bir aralıkta dağıldığı kanısına varabilir.

Çözüm 6 (Sayfa 146)**Soru 1-**

Kızlar için: AO: 18.38 St sapma: 1.37 Serbestlik derecesi: 1563 n: 1564 SEM: $1.37/\sqrt{1564} = 0.035$ %95GA = $18.38 \pm 1.96 \times 0.035$ %95GA = $18.38 \pm 0.069 = 18.31-18.45$	Erkekler için: AO: 18.25 St sapma: 1.37 Ser derecesi: 1545 n: 1546 SEM: $1.30/\sqrt{1545} = 0.033$ %95GA = $18.25 \pm 1.96 \times 0.033$ %95GA = $18.25 \pm 0.065 = 18.185-18.315$

	Cinsiyet	N	Ort.	SS	%95 güven aralığı
Yaş	Kız	1564	18,38	1,37	18.31-18.45
	Erkek	1546	18,25	1,30	18.18-18.31

Soru 2-

Hastalar için: AO: 37.51 SS: 8.7 SD: 41 (t tablosundan 2.02) n: 42 SEM: $8.7/\sqrt{42} = 1.34$ %95GA = $37.51 \pm 2.02 \times 1.34$ %95GA = $37.51 \pm 2.71 = 34.80-40.22$	Kontroller için: AO: 33.73 SS: 7.16 SD: 22 (t tablosundan 2.074) n: 23 SEM: $7.16/\sqrt{23} = 1.49$ %95GA = $33.73 \pm 2.074 \times 1.49$ %95GA = $33.73 \pm 3.09 = 30.64-36.82$

Güven aralıkları kesiştiğinde her zaman arada fark olmadığı sonucunu çıkaramayız. Ancak güven aralıklarının kesişmemesi halinde arada anlamlı bir fark olduğu sonucunu çıkarabiliriz.

NOT: Bu problemi sayfa 133'te gördüğümüz Farkların güven aralığı hesabı ile de çözüünüz.

Çözüm 7 (Sayfa 166)

1.

Araştırmadaki kümeler; 5 sağlık hizmeti organizasyonu ;3 medikal grup, 2 medikal plans. Küme içindeki bireylerin birbirine benzeme ihtimali nedeniyle vaka sayısı görünenden az olacağından küme örnekleme yapılan çalışmalarda örneklem sayısını büyük tutmak zorundayız. Örneklem hesabında küme katsayısı dikkate alınmalıdır.

2.

Dahil etme (inclusion) kriterleri:

- a. Şubat 2002 ve Şubat 2003 yılları arasında klinisyenler tarafından 18 yaşında veya daha büyük olan depresyon başlangıç tedavisi alan veya tedavisi değiştirilen hastalar olarak tanımlanan kişiler.
- b. Katılanların telefonu olması, İngilizce konuşması ve DSM IV’e göre majör depresif bozukluk ve distimi tanılarını karşılaması gereklidir.

Hariç tutma (exclusion) kriterleri: İlk viziteden sonra 14 gün içinde ikinci bir değerlendirme görüşmesi yapılamayacaksa, hamileyse, intihar düşünceleri varsa, şizofrense, bipolar bozukluğu varsa, posttravmatik stres bozukluğu varsa, madde bağımlılığı varsa.

3.

Araştırma hasta akışı genelde CONSORT şemasına uygun olarak verilmiştir. Participants flow and follow up bölümü sonuç bölümünde verilebilirdi.

4.

Araştırmanın analizinde intention to treat yöntemi kullanılmıştır. Hastaların çalışmayı terk etmesi gibi bir olay olmadığı için çalışmanın güvenilirliği artmış ve sonuçları daha inandırıcı olmuştur.

5.

Araştırmada sonuç bölümünden önce örneklem hesabıyla ilgili bilgi verilmiştir. Baseline depresyon skorunu 1,5 olarak almışlar ve gruplar arasındaki anlamlı farkı da görüşmelerde depresyon skorundaki 0,15 (SD 0,50) lik değişim olarak alıp örneklem hesabı yapmışlar.

6.

Görüşmecilerin araştırmanın amacı hakkında bilgileri olmadığı belirtilmiş, katılan hastaların da araştırma hakkında bilgileri olmaması halinde bu araştırma çift kör olur.

7.

Ana sonuç ölçütü “baseline depresyon skoru” olarak belirtilmiş.

Çözüm 8 (Sayfa 181)

1.

Araştırma 1 için odds oranı = $(100/10)/(8100/15010) = 18,75$

2.

Araştırma 2 için rölatif risk = $600/850/400/2150 = 3,8$

3.

Birinci araştırma ile ikinci araştırmanın vaka sayılarına bakarsak;

I. Araştırma toplam katılan 23110 kişinin 110'u hastadır.

II. Araştırma toplam katılan 3000 kişinin 1000'i hastadır.

Vaka-kontrol çalışmaları nadir görülen hastalıklarda yapıldığı için 23 bin gibi sayıca çok katılımcıyı bulmak zor olur. Bir de yine ikinci çalışmada hasta olanlarda riskle karşılaşma durumu da oldukça yüksek, bu nedenle ikinci çalışma vaka kontrol için daha uygun.

Çözüm 9 (Sayfa 194)

1.

P değeri genelde %5 olarak alınır ve tip 1 hatayı yani H_0 'ın doğru olduğu halde reddedilme olasılığını ifade eder.

2.

Tip 1 hata ihtimalini azaltırsak, tip 2 hata yapma ihtimalini artırırız yani ters ilişki vardır.

3.

- 1- Örneklem büyüklüğü (arttıkça güç artar)
- 2- Effect of interest (=etki genişliği=farkın önemlilik derecesi) (arttıkça güç artar)
- 3- Standart sapma (variabilite) (arttıkça güç azalır)
- 4- Anlamlılık düzeyi (p) (arttıkça güç artar)

4.

Güç, örneklem büyüklüğü, etki genişliği, variabilite ve anlamlılık düzeyi birbiriyle ilişkili olduğundan örneklem sayısını çok artırdığımızda öteki değerleri de etkileriz. Bu durumda örneğin etki genişliğini çok küçük tutabiliriz. Dolayısıyla klinik açıdan önemli olmayan küçük farkları da istatistiksel olarak anlamlı bulabiliriz.

Çözüm 10 (Sayfa 203)

1.

Quality of life of people with Diabetes = Diyabetli insanların yaşam kalitesi

2.

H_0 =Diyabet hastası insanlar ile sağlıklı insanların yaşam kalitesi arasında fark yoktur.

H_1 = Diyabet hastası insanlar ile sağlıklı insanların yaşam kalitesi arasında fark vardır.

3.

Araştırmanın örneklemi 440 kişidir. (*Data could be obtained for 440 patients*)

4.

Bu dağılımın variabilitesi ortalamanın standart sapması veya varyansı ile hesaplanır.

5.

Bu araştırma için **effect of interest %7** alınmıştır.

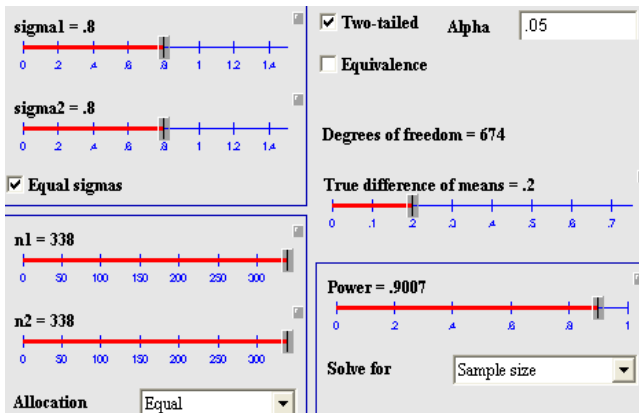
6.

A P value of less than **0.05** was regarded as statistically significant.

7.

Taking the population standard deviation as 0.8 and the effect of interest as 7%, 338 patients were needed to reach a **power of 90%**.

8.



Çözüm 11 (Sayfa 290)

	Gözlemci1				
Gözlemci2	1	2	3	4	Toplam
1	880	66	0	17	963
2	146	145	18	15	324
3	0	0	30	5	35
4	3	1	25	14	43
Toplam	1029	212	73	51	1365

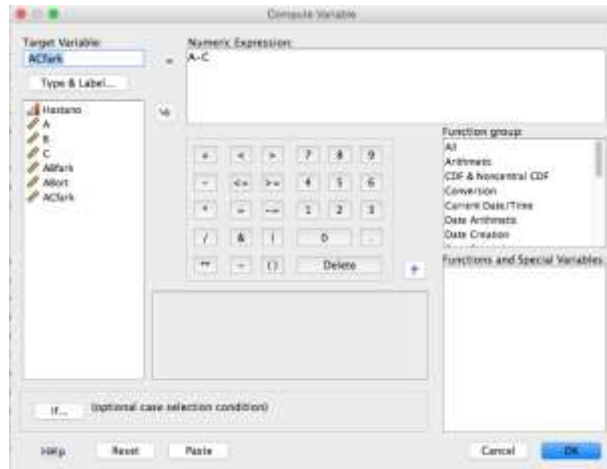
$$\text{Basit Uyum} = \frac{880+145+30+14}{1365} = 0,78$$

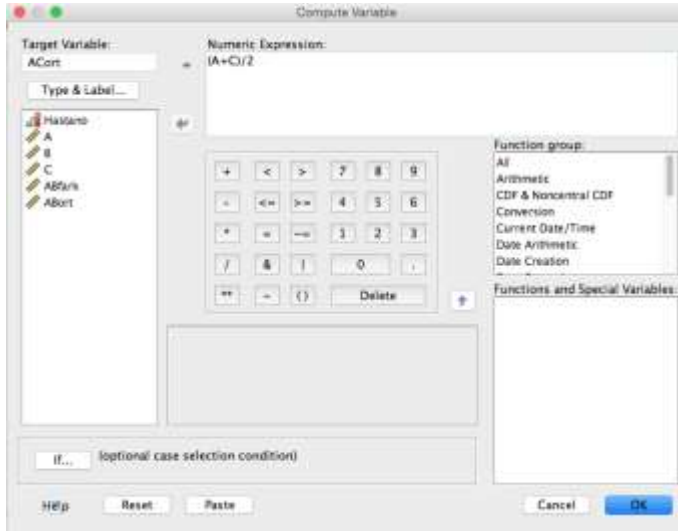
Cohen'in kappa katsayısı (κ) için önce P_e yi hesaplayalım

$$P_e = \frac{[(1029)(963)+(212)(324)+(73)(35)+(51)(43)]}{1365^2} = 0,57$$

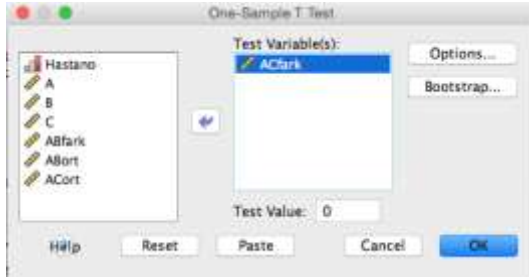
$$\kappa = \frac{0,786-0,57}{1-0,57} = 0,49$$

κ 0,49 olarak hesaplandı ve iki gözlemci arasında uyum “orta” düzeydedir.

Çözüm 12 (Sayfa 332)



Hastano	A	B	C	ACfark	ACort
1	107.61	126.17	108.01	0	108
2	91.92	87.62	76.15	16	84
3	93.77	147.21	103.37	-10	99
4	81.02	90.64	95.71	-15	88
5	46.21	77.70	84.11	-38	65
6	89.57	125.97	79.71	10	85
7	100.76	130.59	105.48	-5	103
8	81.71	91.30	52.02	30	67
9	97.29	114.99	113.32	-16	105
1	95.10	97.41	124.95	-30	110
11	93.28	89.61	103.88	-11	99
12	95.05	104.96	109.37	-14	102
13	81.73	118.78	92.00	-10	87
14	112.80	111.99	81.65	31	97
15	81.10	98.90	105.15	-24	93
16	79.88	100.66	68.10	12	74
17	80.45	118.55	80.94	0	81
18	82.85	104.86	101.30	-18	92
19	90.97	106.88	79.55	11	85

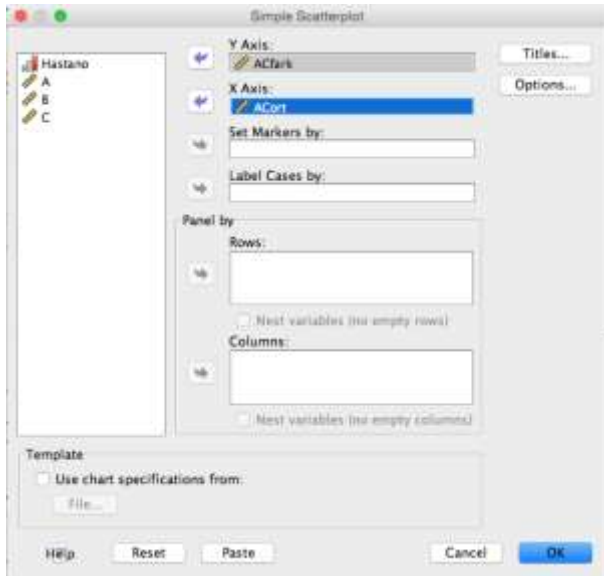


One-Sample Statistics				
	N	Mean	Std. Deviation	Std. Error Mean
ACfark	30	-3.53	21.548	3.934

One-Sample Test					
Test Value = 0					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference Lower Upper
ACfark	-.898	29	.377	-.3.531	-11.58 4.52

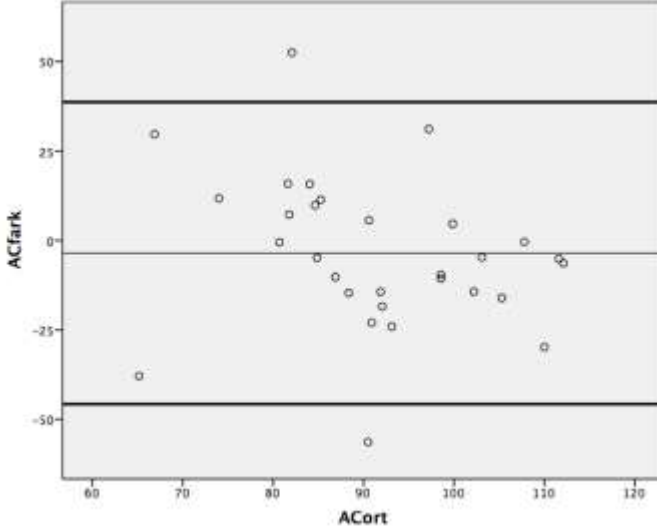
$p=0,377$ olduğundan “A ve C değişkenlerinin ortalamaları arasındaki fark istatistiksel olarak sıfırdan farklıdır”. Yorumunu yapabiliriz.

Bland-Altman grafiğini çizerek olursak:



Yatay çizgiler:

1. Çizgi: Ortalama = **-3,53**; 2. Çizgi: $-3,53 + 1,96 * 21,54 =$ **38,68**; 3. Çizgi: $-3,53 - 1,96 * 21,54 =$ **-45,74**



Yorumlayacak olursak: ACfark değişkeni sıfıra yakın olduğundan A testinin yerine C testini kullanmamız mümkün gözükmektedir. Bununla birlikte, farklı AC ortalamalarında variabilitenin yüksek olduğu dikkati çekmektedir. Dolayısıyla A testi yerine C kullanmakta yine de temkinli olunmalıdır.

Çözüm 13 (Sayfa 347)

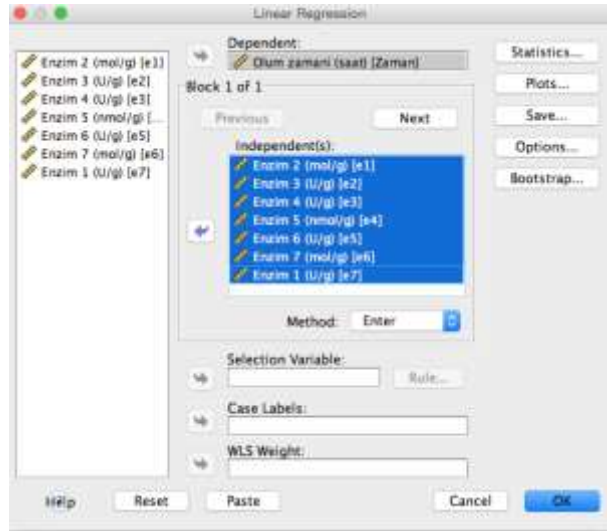
1.

Bağımlı değişken ölüm zamanıdır (saat). Bu değişken numeriktir.

2.

Bağımsız değişkenler olarak yedi adet enzim düzeyi ölçülmüştür (e1, e2, e3, e4, e5, e6, e7). Bu değişkenler de numeriktir.

3.



Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.976 ^a	.952	.940	.424
a. Predictors: (Constant), Enzim 1 (U/g), Enzim 6 (U/g), Enzim 5 (nmol/g), Enzim 2 (mol/g), Enzim 3 (U/g), Enzim 7 (mol/g), Enzim 4 (U/g)				

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	99.975	7	14.282	79.577	.000 ^b
	Residual	5.025	28	.179		
	Total	105.000	35			
a. Dependent Variable: Olum zamanı (saat)						
b. Predictors: (Constant), Enzim 1 (U/g), Enzim 6 (U/g), Enzim 5 (nmol/g), Enzim 2 (mol/g), Enzim 3 (U/g), Enzim 7 (mol/g), Enzim 4 (U/g)						

Coefficients ^a					
		Unstandardized Coefficients		Standardized Coefficients	
Model		B	Std. Error	Beta	t
1	(Constant)	2.772	.568		4.878
	Enzim 2 (mol/g)	.032	.021	.095	1.551
	Enzim 3 (U/g)	.000	.006	-.001	-.007
	Enzim 4 (U/g)	-.008	.003	-.208	-2.463
	Enzim 5 (nmol/g)	-.004	.003	-.096	-1.425
	Enzim 6 (U/g)	-.011	.010	-.094	-1.104
	Enzim 7 (mol/g)	.052	.011	.375	4.533
	Enzim 1 (U/g)	-.089	.020	-.307	-4.492

a. Dependent Variable: Olum zamanı (saat)

Analiz sonucunda Enzim 1, Enzim 4 ve Enzim 7 anlamlı çıktığından bu değişkenler kullanılarak analiz tekrar edildi:

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.969 ^a	.939	.933	.447
a. Predictors: (Constant), Enzim 1 (U/g), Enzim 4 (U/g), Enzim 7 (mol/g)				

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	98.596	3	32.865	164.214	.000 ^b
	Residual	6.404	32	.200		
	Total	105.000	35			
a. Dependent Variable: Olum zamanı (saat)						
b. Predictors: (Constant), Enzim 1 (U/g), Enzim 4 (U/g), Enzim 7 (mol/g)						

Coefficients ^a						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.654	.372		7.133	.000
	Enzim 4 (U/g)	-.012	.002	-.307	-5.362	.000
	Enzim 7 (mol/g)	.064	.010	.463	6.441	.000
	Enzim 1 (U/g)	-.095	.020	-.328	-4.863	.000
a. Dependent Variable: Olum zamanı (saat)						

4.

%93,9

5.

“Ölüm Zamanı” = 2,654-0,328 * “Enzim 1 düzeyi” - 0,307 * “Enzim 4 düzeyi” + 0,463 * “Enzim 7 düzeyi”

Çözüm 14 (Sayfa 359)

1.

Bağımlı değişken menopoza varlığıdır (var/yok).

2.

Öncelikle değişkenler biyolojik açıdan mantıklı olmalı. Ayrıca modele girecek değişkenlerin bağımlı değişkenimizle olan ilişkisini hipotez testleri ile tek tek inceleyebiliriz (Modeli oluştururken verilerin tek tek karşılaştırılmasında p'nin 0,05'e %5-10 yakın olmasını da geçerli sayabiliriz.)

3.

Duyarlılık: %88,9

Özgüllük: %94,9

Doğru sınıflandırma oranı: $(32+75)/115 \times 100 = \%93.0$

4.

-2loglikelihood ne kadar düşükse modelimiz o kadar iyidir. Tabloda modeldeki değişkenlerin likelihoodlarına baktığımızda değerlerin küçük olduğunu görüyoruz (örneğin yaşın likelihood'u: 125.832, ateşter'in: 102.185)

5.

Exp(B)=odds ratio, yani olmanın olmamaya oranının kaç kat olduğunu gösterir. Menopozu tahmin etmede en etkili olan değişken ATESTER(1) DEĞİŞKENİDİR. Çünkü katsayısı en yüksek olan değişken odur.

6.

Beta lojistik regresyon eşitliğinde ilgili değişkenlerin katsayısını gösterir. Formül oluşturulduğunda $f(x)=k+d_1b_1+d_2b_2+\dots$ ilişkisinde yerine konur.

7.

Wald tıpkı t, ki-kare, F gibi test sonucudur ve ne kadar büyükse p o kadar anlamlıdır.

8.

ATESTER Odds oranı $\text{Exp}(B) = 39.094$. Ateş basması ve terleme şikayeti olanların olmayanlara göre 39 kat daha fazla menapozda olma olasılıkları vardır. %95 güven aralığına baktığımızda 5 ile 286 arasında olabileceğini görüyoruz.

9.

Yaş için odds oranının güven aralığı 1'i içermektedir. Bu nedenle yaşın bağımlı değişkene etkisinin istatistiksel olarak anlamlı olmadığı sonucunu çıkarabiliriz.

Çözüm 15 (Sayfa 360)

1.

Burada PEFR yi tahmin etmek için oluşturulan model yaş, ağırlık ve boy değişkenlerinden oluşmaktadır.

2.

Model belirli bir sayıda örneklem üzerinde çalışılarak oluşturulmaya başlanır. Sonra çalışmadaki her kişi için her değişkenin değeri ölçülür. Daha sonra sonuç değişkeni üzerine etki edebilecek birçok açıklayıcı değişkenle model oluşturulur.

Bu modele sigara içme durumu, cinsiyet, kronik hastalık olup olmadığı, test sırasında aktif herhangi bir hastalık olup olmadığı da eklenebilirdi.

3.

Cinsiyetin dâhil edilmemesi 19-60 yaş grubunda kadın ve erkeklerde ayrı ayrı modeller oluşturulmasıyla açıklanabilir.

Aksi halde cinsiyet PEFR değerlerine etki edecek önemli bir değişken olduğundan dâhil edilmeliydi.

4.

Yapılan düzeltmelerden sonra Goodness of fit:

Kadınlarda:0.82

Erkeklerde:0.80 bulunmuştur.

Bu modelle PEFR deki değişimin kadınlarda %82 si erkeklerde %80 i açıklanabilir.

5.

PEFR ye en çok etki eden değişken yaştır. Bu modele göre yaştaki bir birimlik değişim PEFR ye 9.7 kat etki edecektir.

Çözüm 16 (Sayfa 371)

Means and Medians for Survival Time

Tedavi	Mean ^a				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Yeni ilaç	4,977	,386	4,221	5,734	3,700	,313	3,087	4,313
Mevcut ilaç	5,163	,359	4,460	5,865	4,200	,902	2,433	5,967
Overall	5,053	,261	4,542	5,563	3,900	,279	3,354	4,446

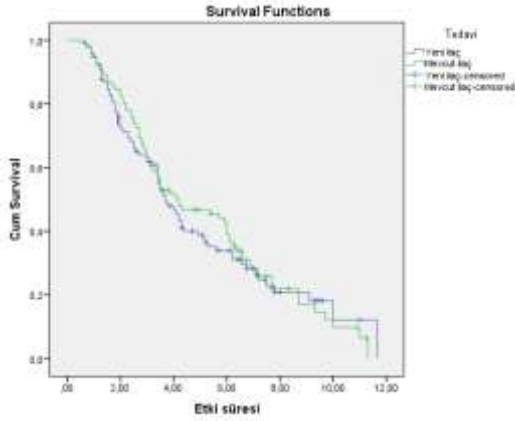
a. Estimation is limited to the largest survival time if it is censored.

Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	,040	1	,842

Test of equality of survival distributions for the different levels of
Tedavi.

“Log Rank testin kullanıldığı sağkalım analizinde, etki süreleri de göz önüne alındığında yeni ilaç ve mevcut ilacın etkinliği arasında istatistiksel olarak anlamlı bir fark bulunmadı ($\chi^2 = 0,4$, $p = 0,842$).”



Çözüm 17 (Sayfa 378)

Cox Regression

Block 0: Beginning Block

Omnibus Tests of Model Coefficients

-2 Log Likelihood
1293,736

Block 1: Method = Enter

Omnibus Tests of Model Coefficients^a

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi- square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
1281,970	11,705	5	,039	11,765	5	,038	11,765	5	,038

a. Beginning Block Number 1. Method = Enter

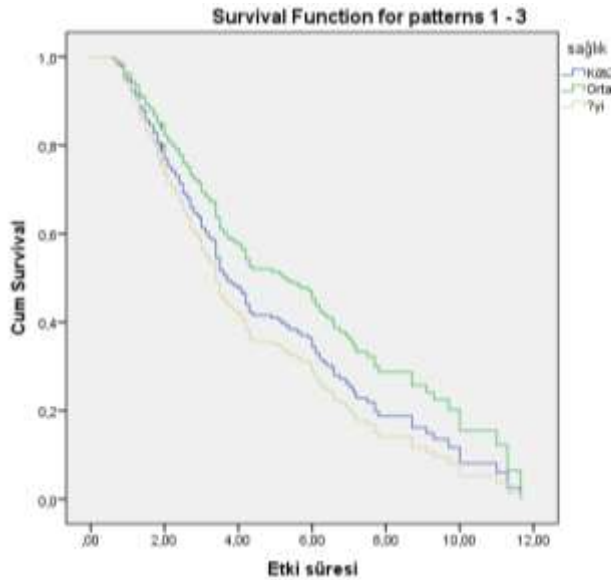
Tabloya göre modelin uyumu istatistiksel olarak anlamlı bulunmuştur (p=0,038).

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95,0% CI for Exp(B)	
							Lower	Upper
tedavi	-,072	,173	,174	1	,676	,930	,663	1,305
yaş	-,013	,008	2,445	1	,118	,987	,972	1,003
cinsiyet	-,320	,173	3,423	1	,064	,726	,517	1,019
sağlık			5,863	2	,053			
sağlık(1)	-,294	,258	1,295	1	,255	,745	,449	1,237
sağlık(2)	,160	,248	,414	1	,520	1,173	,721	1,909

Bulunan Exp(B)-HR değeri sağlık durumu için şöyle yorumlanabilir: Sağlık durumu “iyi” olan hastaların tedavi etkinliği süre açısından sağlık durumu “kötü” olan hastalardan 1,173 kat daha iyi olmakla birlikte bu durum istatistiksel olarak anlamlı değildir (HR=1,17 [0,72-1,9]; p=0,52).

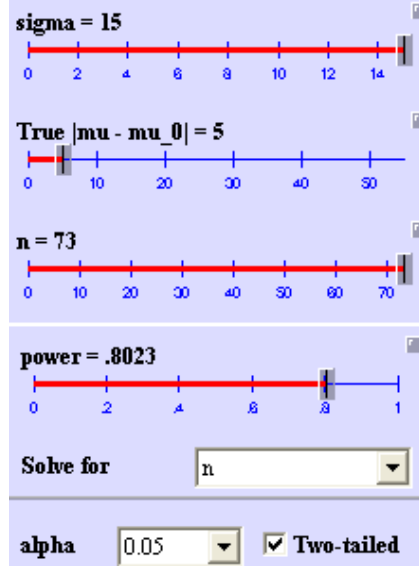
Yine tabloda tedavi için kullanılan ilaçların, yaşın ve cinsiyetin tedavi etkinliği üzerinde bir etkisinin olmadığı görülmektedir.



Çözüm 18 (Sayfa 390)

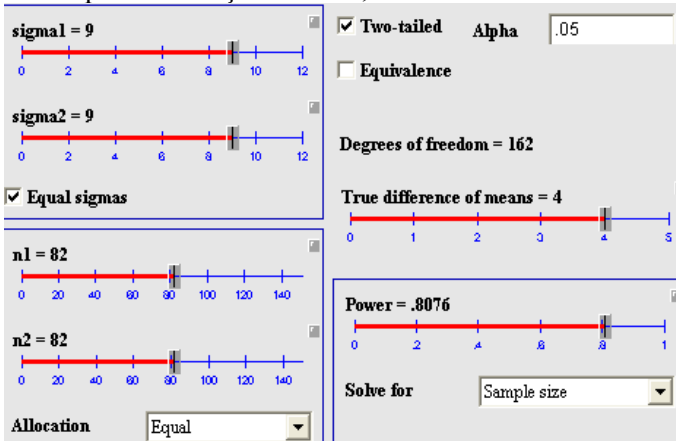
1

- H_0 Erişkin bireylerin sistolik kan basıncı ortalamaları toplumunkinden farklıdır.
- Kesitsel araştırma (Bakınız s. 138)
- Tek örneklemede t testi (Bakınız s. 196)
- 73 kişi



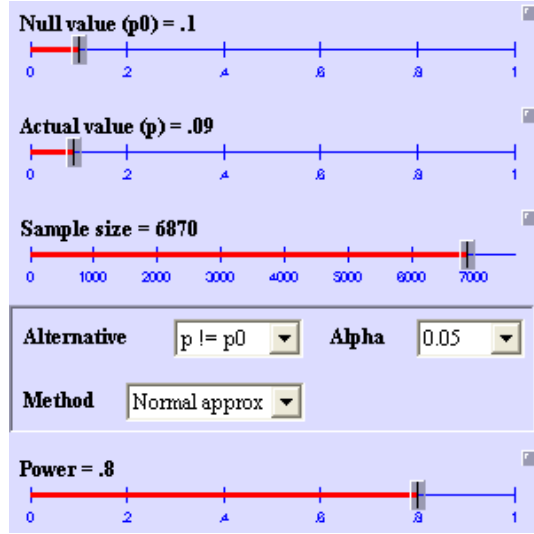
2

- H_0 : Mahalledeki erişkin erkek ve kadınların kan basınçları arasında fark yoktur.
- Kesitsel araştırma (Bakınız s. 138)
- Bağımsız örneklemelerde t testi (Bakınız s. 220)
- Toplam 164 kişi. 82 kadın, 82 erkek.



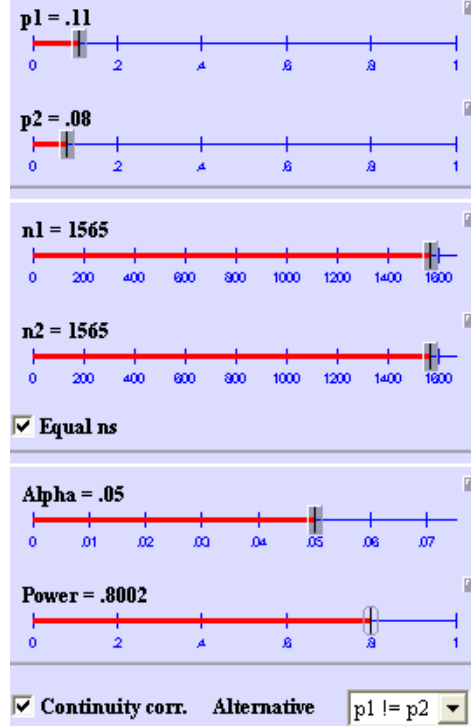
3

- a. H_0 : Mahallemizdeki şeker hastalığı yaygınlığı toplumdakinden farklı değildir.
- b. Kesitsel araştırma (Bakınız s. 138)
- c. z testi (Bakınız S. 196)
- d. 6870 kişi (Toplumdaki birey sayısı (evren) bilindiğinde daha düşük olacaktır. Bu durumda örneklem hesabı formülü değişecektir. Bakınız s. 384)

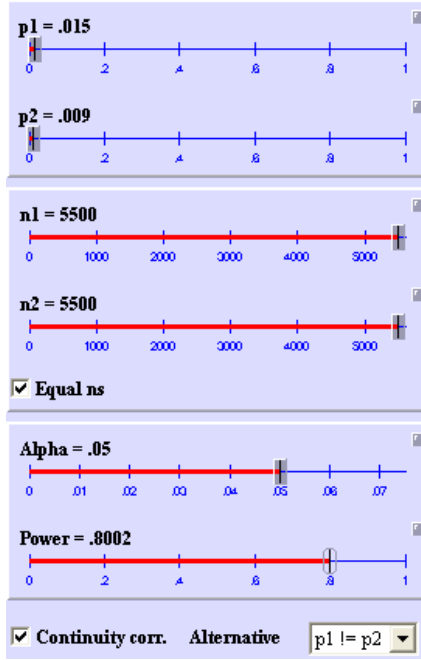


4

- a. H_0 : Mahalledeki erkekler ve kadınlar arasında şeker hastalığı yaygınlığı eşittir.
- b. Kesitsel araştırma (Bakınız s. 138)
- c. Ki kare testi (Bakınız s. 280)
- d. 3130 kişi. 1565 kadın, 1565 erkek.

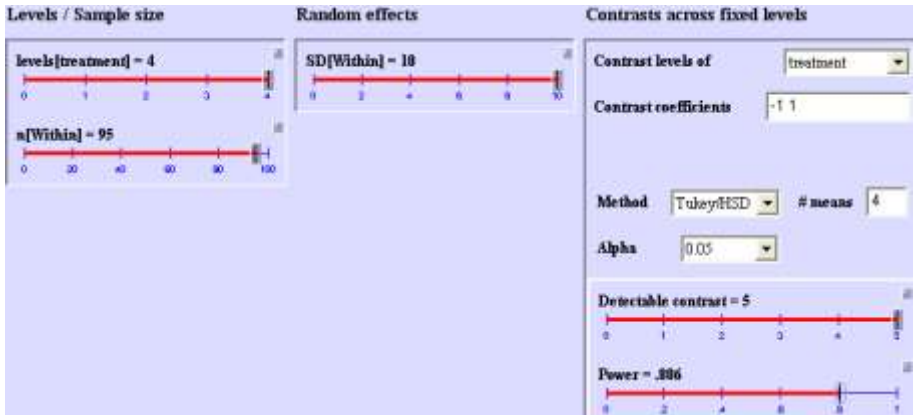


- H_0 : İnme riski açısından A ve B ilaçları arasında fark yoktur.
- Vaka kontrol çalışması (Bakınız s. 175)
- Ki kare testi (Bakınız s. 280)
- 11000 kişi. 5500 vaka ve 5500 kontrol.



6

- H_0 : Tedavi grupları arasında kan şekerini düşürme açısından fark yoktur.
- Randomize kontrollü çift kör çalışma (Bakınız s. 153)
- Tek yönlü varyans analizi (ANOVA) (Bakınız s. 248)
- 380 kişi. Her grupta 95 kişi.



Kaynaklar

1. SPSS Statistics Customer Resources. <http://www-01.ibm.com/software/analytics/spss/products/statistics/customer-resources.html> Elektronik Erişim 30.07.2015
2. SPSS Türkiye. <http://www.spss.com.tr/> Elektronik Erişim 30.07.2015
3. Lenth R. Java applets for power and sample size. <http://www.stat.uiowa.edu/~rlenth/Power/> Elektronik Erişim 30.07.2015
4. Özdamar K. Paket Programlar İle İstatistiksel Veri Analizi. Kaan Kitabevi, Eskişehir 2010.
5. Yusuf S, Cairns J, Camm J, Fallen EL. Evidence-Based Cardiology (Evidence-Based Medicine). BMJ Books, London 2009.
6. Kocaçalışkan İ, Bingöl NA. Biyoistatistik. Nobel Yayın Dağıtım, İstanbul 2008.
7. Özdemir O. Medikal İstatistik. İstanbul Medikal Yayıncılık, İstanbul 2006
8. Petrie A, Sabin C. Medical Statistics at a Glance. 3rd Edition. Wiley-Blackwell Publishing, London 2009.
9. Hayran M. Özdemir O. Bilgisayar İstatistik ve Tıp. Hekimler Yayın birliği, Ankara 1996.
10. Bekiroğlu N. Açıklamalı Biyoistatistik Terimleri Sözlüğü. Nobel Tıp Kitabevi, İstanbul 1998.
11. Özdamar K. SPSS İle Biyoistatistik. Nisan Kitabevi, Eskişehir 2015
12. Sümbüloğlu K, Sümbüloğlu V. Biyoistatistik Uygulama Kitabı. Hatiboğlu Yayınevi, Ankara 1991.

İndeks

A

ADaMSoft 415
Ağırlıklı ortalama 88
Akış şemaları 43, 44
Alternatif hipotez 184
Altın oran Bakınız İstatistikte altın oran
altın standart 408
Amaç 23
Ampirik frekans dağılımı 101
anlamlılık düzeyi (p) 186
ANOVA 27, 36, 108, 190, 194, 252, 253,
254, 255, 256, 257, 259, 260
Aralık (range) 93
Araştırma fikri 30
Araştırma Fonları 427
Araştırma ölçekleri 401
araştırma panosu 30, 31
Araştırma panosu 31
Araştırma protokolü 158
Araştırma Raporu Bölümleri 424
Araştırma Tasarımı 139, 151
Araştırma yöntemleri 33
Araştırmanın çıktıları 157
ARDEB 428
Area Under the Curve Bakınız Eğrinin
altında kalan alan
Aritmetik ortalama 84
asansör grafiği Bakınız Saplı kutu
ASCII 56
aşırı değer 67
AtademiX 21
AUC Bakınız Eğrinin altında kalan alan
ayırma hataları 157

B

Bağımlı değişken 375
Bağımlı örneklemelerde t testi 46
Bağımlı örneklemelerde t-testi 209, 210
Bağımsız değişken 375
Bağımsız örneklemelerde t testi 46, 190,
195, 224
Bakanlık Araştırma Fonları 433
BAP fonları 427
Bar grafik 70
Bar grafikleri 70
Basit uyum katsayısı 288
Başlık Düzeni 422
Bayes Teoremi 410
Beck depresyon ölçeği 401
Between-Subjects Effects 246, 248, 270
BİDEB Bakınız TÜBİTAK Bilim İnsanı
Destekleme Programları
Binomiyal dağılım 110
Birden çok hipotezi test etme 193
birey içi farklılık 97
bireyler arası farklılık 97
Bland-Altman analizi
SPSS ile yapılması 328
Bland-Altman Analizi 327
Bland-Altman analizinin kullanım yeri 327
bloklama 157
Bloklaama 152
BMDP 416
Bonferroni 194, 257
Bonferroni düzeltmesi 241, 249
Boş verilerin kodlanması 59
Boxplot Bakınız saplı kutu

Breslow 366

C

central limit teorem 188

Cevaplanabilir bir klinik soru sormak

Bakınız PICO

Cochran Q testi 48

Cochran's Q 313, 314

Cochrane's Q

SPSS ile yapılması 313

varsayımları 313

coefficient of variation See varyans katsayısı

Cohen'in kappa katsayısı 288, 289, 479

confounding Bakınız Karıştırıcı faktörler

CONSORT bildirisi 159

Construct validity Bakınız yapı geçerliliği

Content validity Bakınız içerik geçerliliği

Cox regresiyon analizi 380

Cox regresyon analizi 363, 375

Çok değişkenli 376

Criterion validity Bakınız tahmin ettirici geçerlilik

CSPRO 415

CumFreq 415

Ç

Çarpma kuralı 100

çıkarılma kriterleri 158

çizgi grafik 70

Çok seçenekli soruların SPSS'e girilmesi 58,
Bakınız SPSS'e veri girişi

Çoklu Çözümlemeler 49

Çubuk grafik Bakınız Bar grafik

D

dahil edilme kriterleri 158

DAP 415

değişken 52

Deneyisel araştırmalar 140

desen etkisi 385

Devlet Planlama Teşkilatı 433

dikotom 53, 272

Dipnot 422

Dot plot 70

dörtlük çeyrek değerler genişliği 67, 94

DPT Bakınız Devlet Planlama Teşkilatı

Durdurulmuş (censored) veri 55

Durdurulmuş veri 364

duyarlılık 408

Duyarlılık 27, 395, 396

Duyarlılık - Özgüllük 49

E

Eğrinin altında kalan alan 397

Eksik veriler 66

endpoints Bakınız Araştırmanın çıktıları

eşdeğerlik (non inferiority) 185

Etik 158

etki genişliği 190, 384

Etkileşim 153

evren 52

Excel dosyası 56

F

Face validity Bakınız görünüm geçerliliği

Faktöriyel tasarımlar 152

Faz 156

Fazla Test İstenmesinin Sonuçları 409

Fisher exact testi 48

Frekans dağılımının şekli 73

Friedman 27, 43, 233, 241, 242, 249, 250

G

Gauss dağılımı *Bakınız Normal dağılım*

Geçerlilik

eşzamanlı 404

Goodness of fit 405

görünüm 405

içerik geçerliliği 404

Likelihood ratio 405

tahmin ettirici 404

yapı 404

Geçerlilik Çeşitleri 404

Geçerlilik ve Güvenilirlik 401

Geçerlilik ve Güvenilirlik Analizi 50

Genel Sağkalım 364

General Linear Model 153, 235, 238, 244, 246

Genstat 416

Geometrik ortalama 85

gerçek negatiflik *Bakınız özgüllük*

gerçek pozitiflik *Bakınız duyarlılık*

Glomeruler filtrasyon hızı 327

GNU Octave 415

GNU PSPP 415

Gözlemsel araştırmalar 140

güç ve ilişkili faktörler 192

Güven aralığı

orantılar için 131

Güven Aralıkları 129

Güvenilirlik

Cronbach alfa 404

gözlemciler arası 403

paralel formlar 403

test/tekrar test 403

yarıya bölme 403

H

Ham veriler olmadan SPSS ile Ki Kare analizi 310

Harmonik ortalama 89

Hata arama 64

Hata ayıklama 64, 76

Hata marjı 387

Hatalar (bias) 146

Hayalet Yazarlık 439

Hazard ratio *Bakınız Zarar oranı*

Hedefler *Bakınız Amaç*

Hediye Yazarlık *Bakınız Onursal Yazarlık* hız 55

Hipotez 37

Hipotez kurma 30, 32

Hipotez Testleri 184

Hipotez Testlerinde Hatalar 190

Histogram 70, 73

Homogeneous Subsets 269, 270

I

ICC *Bakınız Sınıf içi korelasyon katsayısı*

Intraclass Correlation Coefficient *Bakınız Sınıf içi korelasyon katsayısı*

i

İHAK *Bakınız İnternet Üzerinden Herkese Açık Kurs*

İki Nümerik Değişken 49

İki yönlü ANOVA *Bakınız İki yönlü varyans analizi*

İki yönlü varyans analizi 234, 264

İki yönlü ve tek yönlü hipotez 185

intention-to-treat 159

interaction *Bakınız Etkileşim*

İnternet Üzerinden Herkese Açık Kurs 21

İstatistiksel testler ve kullanım yerleri 43
İstatistik Programlarının Karşılaştırılması 417
İstatistik tablolar 460
İstatistiksel Gösterimler 423
İstatistikte altın oran 90
İşaret testi 199, 204, 215

J

JMP 415

K

Kalkınma Ajansları Destek Programları 434
Kaplan Meier Sağkalım Analizi 363
Kappa 403
Kappa analizi 48
Kappa uyum analizi 287
Kare dönüşümü 120
Karekök dönüşümü 119
Karekök örnekleme 123
Karıştırıcı faktörler 147
Kartopu örnekleme 123
kategorik 53
Kategorik Veri İki Bağımlı Grup 47, 287
Kategorik Veri İki Bağımsız Grup 48, 295
Kategorik Veri İkiden Fazla Bağımlı Grup 48
Kategorik Veri İkiden Fazla Bağımlı Kategori 48, 313, 317
Kategorik Veri İkiden Fazla Bağımsız Grup 301
Kategorik Veri İkiden Fazla Bağımsız Kategori 48, 321, 323
Kategorik Veri İkiden Fazla Grup 48
Kategorik Veri İkiden Fazla Kategori, 1 Grup 49
Kategorik Veri Tek Grup 47

Kategorik veriler 58, 166
Kenar Boşlukları 422
Kendall's tau 403
Kendini Değerlendirme 25, 443
Kesme Değeri 397
Ki Kare 262, 279, 280, 283, 285, 295, 296, 297, 298, 301, 302, 303, 307, 308, 315, 321, 323
Ki Kare testi 48, 49
Ki Kare Trend testi 48
Klinik Araştırmalar 155
Klinik kohortlar 173
Kohort Araştırmaları 169
Kohort çalışması 35
Kolmogorov Smirnov 115, 117, 226
Kontrol grupları 146
Korelasyon 335
Korelasyon analizi 49, 335
Körleme 158
Kruskal Wallis Testi 259

L

Levene 118, 229, 255, 260, 456
LIMDEP 416
LISREL 416
likelihood ratio Bakınız Olabilirlik oranı
Linear trend for factor levels 369
linearite testi
 Tek yönlü ANOVA 257
Lineer Regresyon 341
Lineer Regresyon Analizi 49
Literatür tarama 32
Logaritmik dönüşüm 114
Logit (logistic) dönüşüm 120
Logrank 366
Lojistik Regresyon 353

Lojistik regresyon analizi 49, 58

Longitudinal araştırmalarda 146

LR+ Bakınız Olabilirlik oranı

M

Mann-Whitney U testi 36, 223, 230, 231

Mantel-Cox 372

Marginal homogeneity test 317

SPSS ile yapılması 318

varsayımları 317

Marginal Homogeneity testi 48

Mauchly's Test 237, 238, 246

McNemar Bowker testi 50

McNemar testi 47, 278

McNemar-Bowker testi 282

MedCalc 327, 416

MedCalc™ 274, 386

Merkezi eğilim ölçütleri 84

Metodolojik Çalışmalar 50

MINITAB 194, 274, 386

Minitab 415

MiniTab® 196

Mod 85

N

NCSS 416

Nedensellik 145

Nokta prevalansı 146

Nominal 53

Nonparametrik testler 36

Normal Dağılım 99

NPV 396

nümerik 53

Nümerik Veri İki Bağımlı Grup 46, 209

Nümerik Veri İki Bağımsız Grup 46, 223

Nümerik Veri İki Kategorik Bağımsız

Değişken 47

Nümerik Veri İki'den Fazla Bağımlı Grup 46, 233

Nümerik Veri İki'den Fazla Bağımsız Grup 47

Nümerik Veri Tek Grup 46, 199

Nümerik Veri Uyum Analizi 50

Nümerik Veri Uyum Analizi 381

Nümerik veriler 58, 135, 166

O

odds oranı 136, 180, 181, 354, 361

Odds oranları için güven aralığı 136

Olabilirlik oranı 397, 410

olasılık kuralları 409

Olasılık kuralları

Çarpma kuralı 100

Toplama kuralı 100

olasılık yoğunluk fonksiyonu 186, 187

Olay 364

Ondalıklar 80

on-treatment 159

Onursal Yazarlık 439

OpenMx 415

OpenStat 415

oran 55

orantı 55

Orantının örneklem dağılımı 125

Ordinal 53

Ortalamanın örneklem dağılımı 124

ortalamanın standart hatası 125, 129, 166

Ortanca 85

Ortanca için güven aralığı 132

Otit tanısı 411

outcomes Bakınız Araştırmanın çıktıları

Ö

Öğrenim hedefleri 26
Önemlilik testleri 36
Örnekleme Hesabı 384
Örnekleme varyasyonu 124
Örnekleme 123
 kota örnekleme 123
 küme örnekleme 123
 sistematiik 123
Özel Tablolar 82
özgüllük 408
Özgüllük 34, 395

P

p Bakınız Anlamlılık düzeyi
Paket Programları 415
Paralel ve çapraz çalışmalar 152
parametre 53, 103, 108, 111, 113, 125
Parametrik olmayan testler 187
Parametrik testler 36
 varsayımları 37
PASS 386, 390
Pasta dilimi 70
Pasta grafiği 72
pbdR 415
Pearson korelasyon katsayısı 335, 341, 342
Persantil aralıkları 93
Phi 403
PICO 32, 139, 140, 157, 166
PICOS 32
Pie graph Bakınız Pasta grafiği
Pivot Tables 81
Poisson dağılımı 111
Posttest odds 412
power 151, 191, 194, 205, 478

Pozitif tahmin ettirici değeri 408
PPV 396
Pretest odds 412
Profile Plots 237, 241, 246, 248, 269
Psikometrik 50, 402, 405
PSPP Bakınız GNU PSPP
punto 80, 421, 422

R

R İstatistik Programı 415
Randomize Kontrollü Çalışma 35
Rapor Yazma 421
Receiver Operating Characteristics Bakınız
 ROC Analizi
referans aralığı 94
relatif risk 354
Relative Deviate 129, 137
Replikasyon 151
Resiprok dönüşüm 120
retrospektif kohort 169
ROC analizi 49, 395, 396
ROC Analizi 395
Rosenberg 401
Russ Lenth 194, 195, 207, 386

S

saçılma grafiği 70, 76, 472
Saçılma grafiği 327
Sage Stats 415
Sağa eğimli 73
Sağkalım eğrileri 366
sağlıklı birey etkisi 169
sansürlü veri Bakınız Durdurulmuş veri
saplı kutu grafiği 70, 76
Saplı kutu grafikleri 66, 444
SAS 415

SAS University Edition 415

Satır Aralıkları 422

Sayfa Numaraları 423

Scatter diagram Bakınız saçılma grafiği

SEM Bakınız Ortalamanın standart hatası

sensitivity Bakınız Duyarlılık

Serbestlik derecesi 133

SF-36 401

Shogun 415

Sıfır hipotezi 184

Sınav

geçerlilik güvenirlik 252, 402, 406

Sınıf içi korelasyon katsayısı 381

SigmaStat 416

Skewness 115, 116, 117, 131, 200, 223

Sola eğimli 74

Son Not 422

Spearman korelasyon katsayısı 339

specifity Bakınız Özgüllük

Split half Bakınız yarıya bölme

SPSS syntax 417

SPSS Syntax 311

SPSS'e veri girişi 56

Standart Normal Dağılım 104

Standart Normal Sapma 104

Standart sapma 95

Stata 416

STATISTICA 416

StatsDirect 416

Student t test Bakınız Bağımsız
örneklerde t testi

Sütun grafikleri 70

Syntax Bakınız SPSS Syntax

TableLook 81

Tablo Ayarları 81

Tablolar 78, 423

Tanı Testleri 407

Tanımlayıcı istatistikler 39, 386, 416

Tarone-Ware 366

Tek orantı testi 272

Tek örnekleme t-testi 199

Tek yönlü ANOVA 47

tekrarlayan kesitsel araştırmalar 146

Tekrarlayan ölçümlerde ANOVA 46

Teorik dağılımlar

F dağılımı 108

Ki kare dağılımı 108

Lognormal dağılım 108

t-dağılımı 107

Test/retest Bakınız test/tekrar test

Test/tekrar test 403

Testin “teknik özellikleri” 408

Tests of Within-Subjects Contrasts 237,
239, 246, 248

Tests of Within-Subjects Effects 237,
239, 246, 247

Teşekkür 441

Tip I hata 191

Tip II hata 191

Toplama kuralı 100

toplum 52, 94, 104, 123, 124, 125, 130,
140, 163, 199, 201, 202, 203, 205, 210,
213, 214, 215, 272, 274

TÜBİTAK Araştırma Destek Programları
428

TÜBİTAK Bilim İnsanı Destekleme
Programları 432

Türetilmiş veriler 54

T

Tabakalı randomizasyon 157

U

uç değerler 64, 67, 68, 76, 165, 444

Uç değerler 66

Uluslararası Fon Kuruluşları 434

Unistats 416

Ü

Ücretli istatistik programları 415

Ücretsiz istatistik programları 415

V

Vaka Kontrol 34, 178

Vancouver stili 426

Varyans 94

varyans katsayısı 95

Varyasyon 151

Veri çeşitleri 36, 52

Veri Dönüştürme 113

Veri giriş yöntemleri 56

Veri girişi 57, *Bakınız SPSS'e veri girişi*

Verilerde bulunması gereken özellikler 36

Verilerin grafiklerle gösterilmesi 70

W

Wilcoxon işaretli sıra testi 209, 215, 216

WOMAC 401

Y

Yaşam tabloları 363

Yaygınlık ölçütleri 93

Yazar Paslaşması 439

Yazarlık Ölçütleri 437

Yazarlık Sırası 440

Yazarlık Sorumlulukları 440

yazı tipi 421

Yöntem 38

yüzde 54

Z

Z testi 47

Zamana bağlı değişim 375

Zarar oranı 375, 376

"İstatistikten korkmanıza gerek yok!

Eğitimimiz boyunca istatistiğin rakamlar ve formüllerle ilişkilendirilmesinin de etkisiyle istatistik, sağlık çalışanları için hep korkutucu olmuştur.

Oysa günümüzde tıbbi istatistik neredeyse tamamen bilgisayar yazılımlarıyla yapılmaktadır.

Sağlık çalışanı olarak bizim sadece araştırma ve biyoistatistik yöntemlerine hakim olmamız yetecektir.

Bu kitapta başlangıçta verilen örnekler ve akış şemaları işinizi çok kolaylaştıracak. Çok az bilgisayar bilgisine de sahip olsanız temel tıbbi araştırmalarınızı yardım almadan kolayca yapabileceksiniz.

Öğrencisinden akademisyenine kadar tüm sağlık çalışanlarının anlayıp uygulayabileceği bir anlatım tarzı seçilen bu kitabın size de faydalı olmasını diliyoruz.

