

Saęlık alıřanları İin

*Arařtırma
ve
Pratik İstatistik*

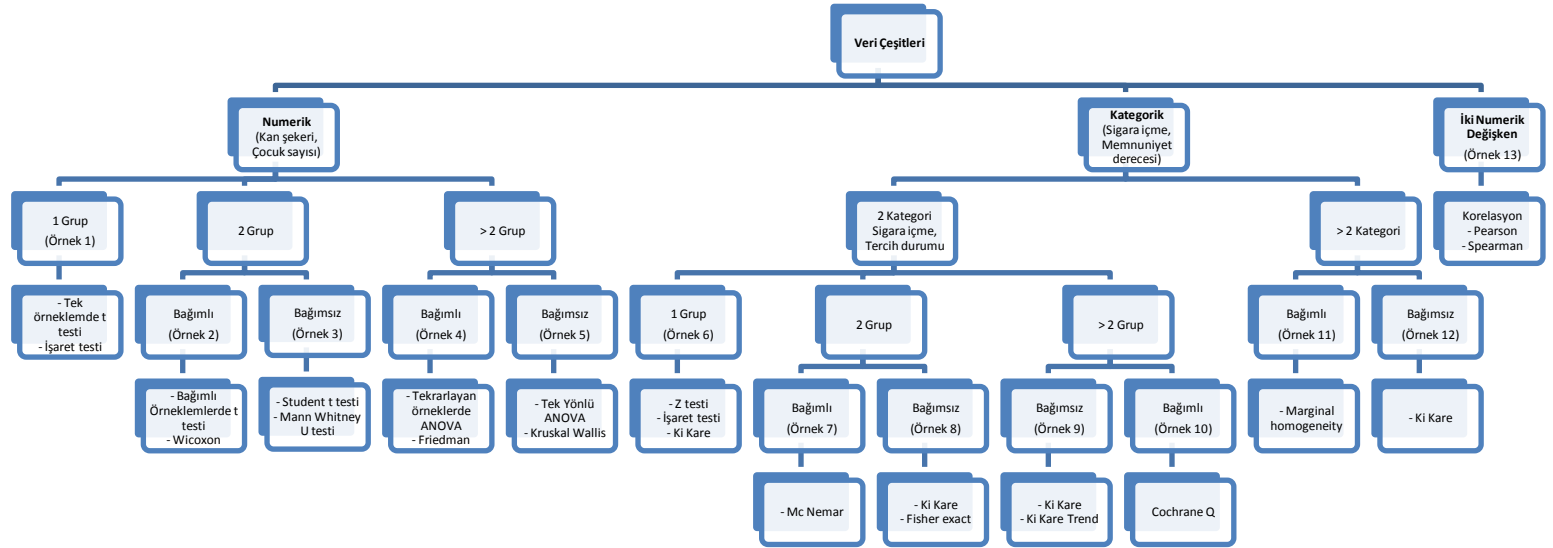
*Örnek Problemler ve
SPSS Çözümleri*

Prof. Dr. Zekeriya Aktürk Do. Dr. Hamit Acemoęlu

Erzurum, 2010

www.aile.net

Önemlilik Testleri ve Kullanım Yerleri



Örnek Problemler

<p>Örnek 1 (Sayfa 157) Nümerik Veri Tek Grup Tek örnekleme t testi</p> <p><i>Mahalleimizdeki bireylerin boylarını ölçtüğümüzü düşünelim [mahalle=tek grup; kilo=numerik veri]. Sonuçlarımızı il geneli ile karşılaştırmak istiyoruz. Daha önceki araştırmalardan ilimizdeki bireylerin boy ortalamasının 161,5 cm olduğunu biliyoruz.</i></p>	<p>Örnek 2 (Sayfa 165) Nümerik Veri İki Bağımlı Grup Bağımlı örneklemlerde t testi</p> <p><i>Diyabetik hastalarımıza bir beslenme ve egzersiz programı uyguladığımızı varsayalım. Program sonunda katılımcıların kilo verip vermediklerini araştırmak istiyoruz</i></p>	<p>Örnek 3 (Sayfa 179) Nümerik Veri İki Bağımsız Grup Bağımsız örneklemlerde t testi</p> <p><i>Diyabet hastalarımızın vücut kitle indekslerinin (VKI) erkekler ve bayanlar arasında farklılık gösterip göstermediğini araştırmak istediğimizi varsayalım</i></p>
<p>Örnek 4 (Sayfa 189) Nümerik Veri İkiden Fazla Bağımlı Grup Tekrarlayan ölçümlerde (repeated measures) ANOVA</p> <p><i>Diyabetik hastalarımıza ilaç, beslenme ve egzersizden oluşan tedavi programı uyguladığımızı varsayalım. Program süresince katılımcıların Hemogloblin A1c düzeylerini araştırmak istiyoruz. Hemogloblin A1c ölçümleri tedavi programı uygulanmadan önce, programın 3. ve 6. aylarında yapılmış olsun. Ölçmek istediğimiz değişken numeriktir. Uygulama öncesi ve sonrasında elde edilen üç ölçümler ise bağımlıdır.</i></p>	<p>Örnek 5 (Sayfa 199) Nümerik Veri İkiden Fazla Bağımsız Grup Tek yönlü ANOVA</p> <p><i>Diyabet hastalarımızın eğitim durumlarına göre boylarının farklılık gösterip göstermediğini araştırmak istediğimizi varsayalım (H0: 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.</i></p>	<p>Örnek 6 (Sayfa 209) Kategorik Veri Tek Grup Z testi</p> <p>a. <i>Diyabet hastalarında sigara içme durumunu araştırdığımızı düşünelim. Örnekleminizdeki sigara içme durumunun toplumdakinden farklı olup olmadığına bakmak istiyoruz</i></p> <p>b. <i>Örneğin diyabetli hastalarımıza iki farklı biyuanid türü ilacı verdiğimizizi ve kullanım kolaylığı, tadı, yan etkileri ve açısından hangi ilacı tercih ettiklerini sordüğümüzü düşünelim</i></p>
<p>Örnek 7 (Sayfa 215) Kategorik Veri İki Bağımlı Grup McNemar testi</p> <p><i>Diyabetik hastalarımıza kan şekeri kontrolünün önemi ve gerektiğinde insülin kullanılması konusunda bir eğitim verdiğimizizi düşünelim. Örnekleminizdeki bireylerin ilaç tercihlerinin eğitimden sonra değişip değişmediğini incelemek istiyoruz.</i></p>	<p>Örnek 8 (Sayfa 219) Kategorik Veri İki Bağımsız Grup Ki Kare testi, Fisher exact testi</p> <p><i>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</i></p>	<p>Örnek 9 (Sayfa 225) Kategorik Veri İkiden Fazla Grup Ki Kare testi, Ki kare trend testi</p> <p>a. <i>Farklı mahallelerden gelen diyabet hastalarımızın medeni durumları arasında fark olup olmadığını incelemek istiyoruz.</i></p> <p>b. <i>Diyabetik bireylerde eğitim durumu ile sigara içme durumu arasında bir trend olup olmadığını incelemek istiyoruz. Eğitim arttıkça içme yüzdesinin azaldığını varsayıyoruz.</i></p>
<p>Örnek 10 (Sayfa 233) Kategorik Veri İkiden Fazla Bağımlı Grup Cochrane Q testi</p> <p><i>Diyabetik hastalarımıza kan şekeri kontrolünün önemi ve gerektiğinde insülin kullanılması konusunda bir eğitim verdik. Örnekleminizdeki 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.</i></p>	<p>Örnek 11 (Sayfa 239) Kategorik Veri İkiden Fazla Bağımlı Kategori Marginal homogeneity testi</p>	<p>Örnek 12 (Sayfa 239) Kategorik Veri İkiden Fazla Bağımsız Kategori Örnek 14 Ki Kare testi</p> <p><i>Örnekleminizdeki bireylerin medeni durumlarının (evli, bekar, boşanmış = ikiden fazla kategori) ilaç 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.</i></p>
<p>Örnek 13 (Sayfa 241) Korelasyon analizi</p> <p><i>Diyabetli hastalarımızda diyabet süresi (x) uzadıkça HbA1c düzeylerinin (y) de arttığından şüpheleniyoruz. Diyabet süresi ile HbA1c arasında doğrusal bir ilişki olup olmadığına bakmak istiyoruz.</i></p>	<p>Örnek 14 (Sayfa 247) Lojistik regresyon analizi</p> <p><i>Diyabet hastalarımızın sigara içme durumunu belirleyen faktörleri araştırmak istiyoruz. Böylece hangi hastalarımızın sigara içmeye daha meyilli olduğunu saptayabilir, gerekli klinik yaklaşımı sunabiliriz.</i></p>	

Sađlık alıřanları İin

Arařtırma ve Pratik İstatistik

Örnek Problemler ve SPSS özümleri

Prof. Dr. Zekeriya Aktürk Do. Dr. Hamit Acemođlu

Kapak

Cumhuriyetin ilk yıllarına ait el iřlemeli kadife para kesesi. Gelin eyizinin bir parası olan para keseleri hanımlar tarafından para ve deđerli eřyalarını tařımak iin kullanılırdı.

Yöresi: Kayseri, Semra Aktürk

Arka Kapak

Türk Motifi, Hatice Dolduran

***ai*lem®**

Aktürk Z, Acemođlu H. Sađlık alıřanları İin Arařtırma ve Pratik İstatistik. Anadolu Matbaası. İstanbul 2010

ISBN: 978-975-00631-1-4

Basım Yeri: Anadolu Ofset, Çađalođlu, İstanbul. Tel: 0212 5114517, 0532 5488961, Faks: 0212 5288920, <http://www.ofsetonline.com>

Sipariř: Atatürk Üniversitesi Tıp Fakóltesi Aile Hekimliđi AD, 25240 Erzurum. Tel: 0442 2317228, Faks: 0442 2316809, www.aile.net, e-posta: zekeriya.akturk@gmail.com

Teřekkür

Kitabın hazırlanma ařamasında, dizgisinde ve tashihindeki desteklerinden dolayı Yrd. Do. Dr. Memet Iřık, Yrd. Do. Dr. Turan Set ve Hemř. Semra Aktürk'e teřekkür ederiz.

İçindekiler

Önemlilik Testleri ve Kullanım Yerleri.....	2
Örnek Problemler.....	3
İçindekiler.....	7
Önsöz.....	13
Kitap Hakkında.....	14
Giriş.....	16
1. Özet.....	17
2. İstatistiksel testler ve kullanım yerleri – akış şemaları.....	29
3. Veri çeşitleri.....	31
Türetilmiş veriler.....	33
Durdurulmuş (censored) veri.....	34
4. Veri girişi.....	35
Veri girişi yöntemleri.....	35
Veri girişinin planlanması.....	36
Kategorik veriler.....	36
Nümerik veriler.....	37
Anketlerin numaralandırılması.....	38
Boş verilerin kodlanması.....	38
Araştırma 1.....	39
5. Hata ayıklama ve uç değerler.....	43
Hata arama.....	43
Eksik veriler (missing) varsa ne yapacağız?.....	44
Uç değerler.....	45
Araştırma 2.....	48
6. Verilerin grafiklerle gösterilmesi.....	49
Frekans dağılımı.....	49
Frekans dağılımlarının gösterilmesi.....	49
Frekans dağılımının şekli.....	54
7. Merkezi dağılım (yığılma) ölçütleri.....	55
Aritmetik ortalama.....	55
Ortanca.....	56
Mod.....	56
Geometrik ortalama.....	56
Ağırlıklı ortalama.....	58
Araştırma 3.....	61
8. Yaygınlık ölçütleri.....	63
Aralık (range).....	63
Persantil aralıkları.....	63
Varyans.....	64
Standart sapma.....	64
SPSS ile hesaplama.....	65
Birey içi ve bireyler arası farklılıklar.....	67
9. Teorik Dağılımlar: Normal Dağılım.....	69
Olasılık hesapları.....	69
Olasılık kuralları.....	70

Olasılık dağılımlarının teorisi.....	70
Normal dağılım (Gauss dağılımı)	72
Standart Normal Dağılım	73
Alıştırma 4	74
10. Teorik Dağılımlar: Diğer Dağılımlar	77
Diğer sürekli teorik dağılımlar	77
t-dağılımı	77
Ki Kare (χ^2) dağılımı	78
F-dağılımı	78
Lognormal dağılım	78
Tam sayılı teorik dağılımlar	81
Binomiyal dağılım	81
Poisson dağılımı	81
11. Veri Dönüştürme	83
Veri dönüşümü nasıl yapılır?	83
Tipik dönüşümler	84
Logaritmik dönüşüm, $z = \log y$	84
Karekök dönüşümü $z = \sqrt{y}$	88
Resiprok dönüşüm, $z = 1/y$	90
Kare dönüşümü, $z = y^2$	90
Logit (logistic) dönüşüm,	90
12. Örneklem ve Örneklem Dağılımları	91
Örneklemin temsil edici olması	91
Örneklem varyasyonu	91
Ortalamanın örneklem dağılımı	92
Standart sapma mı standart hata mı?	93
Orantının örneklem dağılımı	93
Alıştırma 5	94
13. Güven Aralıkları	95
Ortalamanın güven aralığı	95
Normal dağılım bilgilerimizin kullanarak	95
t-dağılımı bilgilerimizi kullanarak	96
SPSS ile ortalamanın güven aralığının hesaplanması	96
Orantılar için güven aralığı	97
Güven aralıklarının yorumlanması	98
Serbestlik derecesi	98
Örnek	98
Farkların güven aralığı	100
Nümerik veriler için	100
SPSS ile ortalamalar arasındaki farkın güven aralığının hesaplanması	100
Orantılar için	101
Odds oranları için güven aralığı	101
Relative Deviate formülünün farkların güven aralığına uygulanması	102
14. Araştırma Tasarımı – I	103
Deneysel / gözlemsel araştırmalar	104
Araştırma çeşitleri	104
Nedensellik	106

Kesitsel / longitudinal arařtırmalar	106
Kontrol grupları	107
Hatalar (bias)	107
Alıřtırma 6	108
15. Arařtırma Tasarımı – II	111
Varyasyon	111
Kopyalama (Replikasyon)	111
Örneklem büyüklüğü	111
Özel arařtırma tasarımları	111
Bloklama	111
Paralel ve çapraz çalıřmalar	112
Faktöriyel tasarımlar	112
16. Klinik Arařtırmalar	115
Karşılařtırmalar	115
Arařtırmanın çıktıları (endpoints, outcomes)	116
Gruplara ayırma	116
Körleme	117
Etik	117
Arařtırma protokolü	117
17. Önceki Konuların Özeti	121
Alıřtırma 7	125
18. Kohort Arařtırmaları	127
Kohortun seçilmesi	127
Kohort arařtırmaların analizi	129
Kohort arařtırmaların olumlu yönleri	130
Kohort arařtırmaların olumsuz yönleri	130
Klinik kohortlar	130
19. Vaka Kontrol Çalıřmaları	135
Vakaların seçilmesi	135
Kontrol grubunun seçilmesi	136
Eřleřtirme	136
Eřleřtirilmemiş vaka kontrol arařtırmalarının analizi	137
Vaka kontrol çalıřmalarının olumlu yönleri	138
Vaka kontrol çalıřmalarının olumsuz yönleri	138
Alıřtırma 8	139
20. Hipotez Testleri	141
Hipotez testlerinde genel yaklařım	141
H_0 ve H_1 hipotezi	141
İki yönlü ve tek yönlü hipotez	142
p deęerinin yorumlanması	143
Parametrik olmayan testler	144
Hipotez testi mi güven aralıęı mı?	144
21. Hipotez Testlerinde Hatalar	147
Karar verme	147
Yanlıř karar verme	147
Arařtırmanın gücü ve iliřkili faktörler	149
Birden çok hipotezi test etme	150

Alıştırma 9	155
22. Numerik Veri Tek Grup.....	157
Tek örnekleme t-testi varsayımları	157
Örnek uygulama	157
Örneğin SPSS ile yapılması	159
Varsayımlar karşılanamıyorsa	160
İşaret testi.....	160
Alıştırma 10.....	164
23. Numerik Veri İki Bağımlı Grup	165
Bağımlı örneklemlerde t-testi.....	165
Varsayımları.....	165
Testin mantığı.....	166
Ek açıklama.....	166
Örnek uygulama	166
Örneğin SPSS ile yapılması	171
Varsayımlar karşılanamıyorsa	172
Wilcoxon işaretli sıra testi.....	172
Mantığı.....	172
Bir örnek üzerinden Wilcoxon işaretli sıra testi	173
Wilcoxon işaretli sıra testinin SPSS ile yapılması	177
24. Numerik Veri İki Bağımsız Grup	179
Bağımsız örneklemlerde t testi	179
Varsayımları	179
Mantığı	179
Not	180
Varsayımlar karşılanamıyorsa	186
25. Numerik Veri İki'den Fazla Bağımlı Grup.....	189
Varsayımları.....	189
Testin mantığı.....	190
Örnek uygulama	190
General Linear Model.....	195
Varsayımlar karşılanamıyorsa	197
26. Numerik Veri İki'den Fazla Bağımsız Grup.....	199
ANOVA.....	199
Varsayımları	199
Mantığı	200
Varsayımlar karşılanamıyorsa	204
Kruskal Wallis Testi.....	204
27. Kategorik Veri Tek Grup.....	209
Tek orantı testi (test of a single proportion)	209
Varsayımlar	209
Not	209
Mantığı	209
Eğer bir tercih durumunu araştırıyorsak (Binomiyal test).....	211
28. Kategorik Veri İki Bağımlı Grup.....	215
McNemar testi.....	215
Testin mantığı.....	215

29. Kategorik Veri İki Bağımsız Grup	219
Ki Kare (χ^2) testi	219
Varsayımlar	219
Testin mantığı	220
30. Kategorik Veri İki'den Fazla Bağımsız Grup.....	225
Çoklu gruplarda Ki Kare (χ^2) testi.....	225
Varsayımlar	225
Testin mantığı	225
Varsayımların karşılanamaması durumunda.....	226
Ki Kare Trend testi	230
31. Kategorik Veri İki'den Fazla Bağımlı Grup.....	233
Cochran's Q testi varsayımları.....	233
32. Kategorik Veri İki'den Fazla Bağımlı Kategori	237
Marginal homogeneity test varsayımları	237
33. Kategorik Veri İki'den Fazla Bağımsız Kategori	239
34. Korelasyon	241
Pearson korelasyon katsayısı	241
Özellikleri	241
r ne zaman hesaplanmamalıdır?.....	242
Pearson korelasyon katsayısı için hipotez testi	243
Spearman korelasyon katsayısı.....	245
Farkları	245
35. Lojistik Regresyon	247
Lojistik regresyon eşitliği.....	248
Modelin oluşturulması	248
Modelin uygunluğunun test edilmesi.....	249
Örnek uygulama	249
Alıştırma 11.....	255
Alıştırma 12.....	256
36. Örneklem Hesabı.....	257
Tanımlayıcı istatistikler için örneklem büyüklüğü formülleri	258
Sonuç ölçütümüz kategorik ise:	258
Sonuç ölçütümüz numerik ise:.....	259
Hipotez testleri için örneklem büyüklüğünün hesaplanması	259
Alıştırma 13.....	263
37. Kendini Değerlendirme Çalışması.....	265
38. Kendini Değerlendirme Çalışması Çözümleri.....	269
Ek 1 – İstatistik tablolar.....	279
Z tablosu (Standart normal dağılım).....	279
t tablosu.....	280
Ki Kare tablosu.....	281
F tablosu.....	282
Pearson korelasyon katsayısı.....	283
Spearman korelasyon katsayısı.....	284
Ek 2 – Alıştırmaların çözümleri.....	285
Alıştırma 1 (Sayfa 38).....	285
Alıştırma 2 (Sayfa 47).....	286

Alıştırma 3 (Sayfa 60).....	287
Alıştırma 4 (Sayfa 72).....	290
Alıştırma 5 (Sayfa 91).....	291
Alıştırma 6 (Sayfa 104).....	292
Alıştırma 7 (Sayfa 121).....	293
Alıştırma 8 (Sayfa 133).....	294
Alıştırma 9 (Sayfa 147).....	296
Alıştırma 10 (Sayfa 155).....	297
Alıştırma 11 (Sayfa 239).....	299
Alıştırma 12 (Sayfa 240).....	300
Alıştırma 13 (Sayfa 247).....	301
Kaynaklar	305
İndeks.....	307

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 sağlığın 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 eğitimcilerin 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 daha çok ihtiyaç vardır.

Geriye dönüp baktığımda, bu kitabın temellerinin akademisyenliğe ilk niyet ettiğim 1994 yıllarında Doç. Dr. Mehmet Akın ile asistanlık eğitimim sırasında yaptığımız çalışmalarda atıldığını görüyorum. 1999 yılında Trakya Üniversitesi'nde akademisyenliğe başladıktan sonra Doç. Dr. Nezih Dağdeviren hocam ve iş arkadaşarımla yaptığımız çalışmaların da çok katkısı oldu. Araştırma ve istatistik konusunda yeterli kaynak bulamamamız ve/veya var olan kaynaklara ulaşamamız 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 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 ulaşımına açık olmasının kendisini geliştirmek isteyen sağlık profesyonellerine katkıda bulunacağı, giderek iyileştirme ve yaygınlaştırma sayesinde 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.

Prof. Dr. Zekeriya Aktürk

Doç. Dr. Hamit Acemoğlu

Kitap Hakkında

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 13 adet araştırma uygulamasına yer verilmiştir. Araştırmaların ilgili konudan sonra çözülmesi ve gerekirse Ek 2’deki 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 kullanılmıştır. Analizde kullanılacak veri seti www.aile.net adresinde yayınlanmıştır.

Haftada iki saat okuma ve iki saat de pratik yapmanız halinde yaklaşık 4 ay içerisinde aşağıda belirtilen hedeflere ulaşabileceğinizi düşünüyoruz. Konular tamamlandıktan sonra “37. Kendini Değerlendirme Çalışması” Bölümündeki deneme sınavını yapmanızı ve “38. 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 araştırmaların 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
- Merkezi dağılım 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
- Örneklem 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 örneklemlerde 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
- Tek örnekleme 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
- Cochran's Q testini açıklayabilme ve kullanım yerini bilme
- Marginal Homogeneity testini açıklayabilme ve kullanım yerini bilme
- Hipotez testlerini SPSS ile yapabileme ve çıktıları yorumlayabilme
- Çeşitli senaryolarda hangi istatistik analizler yapılacağını açıklayabilme
- Korelasyon analizini açıklayabilme, SPSS'te yapabileme ve sonuçları yorumlayabilme
- Lojistik regresyon analizini açıklayabilme, SPSS'te yapabileme ve sonuçları yorumlayabilme
- Örnekleme etki eden faktörleri açıklayabilme ve çeşitli yazılımları kullanarak örneklem hesabı yapabileme

Giriř

Ülkemizde tıp alanında arařtırmaya olan ilgi, bilimsel alıřmalar ve yayınlar giderek artmaktadır. YÖK ve üniversitelerimizin akademik yükseltme ve atama kriterlerini belirlemiş olmaları ve bunların giderek zorlaşması da bilimsel arařtırmalara olan ilgiyi artırmaktadır.

Diđer taraftan, aile hekimliđi disiplininin hızla gelişmesi, birinci basamakta da şimdiye kadar hiç olmadığı kadar bilimsel alıřmalara ilgi duyulmasını sağlamıştır.

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

İstatistik ve arařtırma konusunda ok deđerli basılı eserler ve internet kaynakları vardır. Bununla birlikte, istatistik konusu özellikle klinisyenler için ürkütücü olmaya devam etmektedir. Hekimler karmařık gözükten hesaplamalar yerine pratik olarak alıřmalarına yol gösterilmesini tercih etmektedir.

Arařtırma görevlileri de asistanlık sürelerinin sonundaki tez alıřmalarında önemli zorluklarla karşılaşmakta, alıřmalarının istatistik analizleri için danışmanlığa ihtiyaç duymaktadır.

Klinisyen olarak özellikle aile hekimlerinin olmak üzere tüm sađlık alıřanlarının, arařtırma görevlilerinin ve tıp öğrencilerinin kolaylıkla anlayıp uygulayabilecekleri bir esere ihtiyaç olduğunu düşünürüz.

Temel arařtırma yöntemleri ve istatistik analizlerini içeren, örnek veri setleri ve SPSS'teki kullanımlarıyla desteklenmiş bu kitabın konusuyla ilgili ihtiyacı karşılayacağını düşünürüz.

1. Özet

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

Hedefler: Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Bir araştırmanın aşamalarını sayabilmeleri
- Yanıtlanabilir 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 araştırıyım?)
- 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 ve TÜBİTAK fonları gibi birçok kaynağa da uygun projeler sunmamız halinde ulaşabiliriz. Burada Avrupa Birliği fonlarından da bahsetmek gerekir. 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 karşılařtırma
8. Rapor yazma ve yayınlama

Arařtırma 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 karşılařtırdığımız durumlar
- alıřma ortamımız
- Hastalardan gelen fikirler
- Arkadařlarımızın önerileri ve görüşleri
- Beklenmeyen laboratuvar sonuçları
- ...

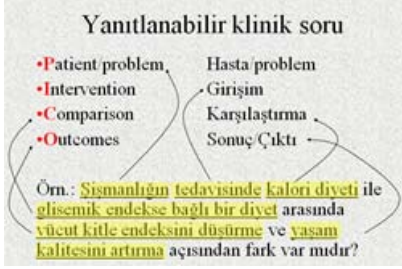
Belirlediğimiz fikirleri hemen araştırma imkanımız olmayacağından bunları not etmemiz ve gerektiğinde kullanmamız iyi olacaktır. Bu amaçla bir araştırma panosu (**Tablo 1**) oluşturabilirsiniz. Panonuz alıřma masanızın önünde olabileceğı gibi bilgisayarınızda bir dosyadan da oluşabilir. Arařtırma panosunda “arařtırma fikirleri”, “yöntem yazılma aşamasında olan konular”, “uygulama aşamasında olan konular” ve “rapor yazım aşamasında olan konular” gibi başlıklarınız olmalıdır. Arařtırma panosu sürekli dinamik olmalıdır ve konular sütunlar arasında yer deđiřtirerek yayına dođru gitmelidir.

Araştırma Panom			
Sıradaki çalışmalar	Yöntem yazım aşamasında	Uygulama aşamasında	Rapor yazım aşamasında
<ol style="list-style-type: none">1. Halk arasında bitkisel ilaçların kullanma durumu2. Kabızlığın zihinsel fonksiyonlara etkisi3. Danışmanlığın ergenlerdeki riskli davranışlara etkisi.4. .	<ol style="list-style-type: none">1. Türkiye’de tıpta uzmanlık eğitimi2. .	<ol style="list-style-type: none">1. Yapılandırılmış egzersizin kilo vermeye etkisi2. Birinci basamak için bir ulusal kayıt yazılımı geliştirilmesi3. .	<ol style="list-style-type: none">1. Edirne’li kadınlarda osteoporoz prevalansı2. Tıp öğrencilerinin bitkileri tanıma durumu3. .

Tablo 1: Örnek bir araştırma panosu.

PICO

Araştırmaya başlarken cevaplanabilir bir klinik soru sorulmalıdır. Sorumuzda dört bileşen olmalıdır: **P**roblem, **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

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

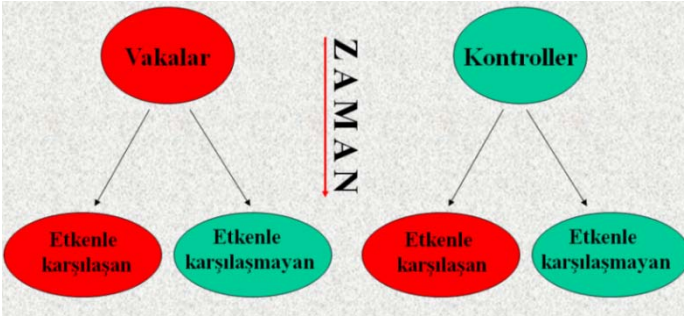
Bilimsel araştırma ve istatistik uygulamaları sırasında aşağıdaki noktalara dikkat edilmesini öneriyoruz:

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

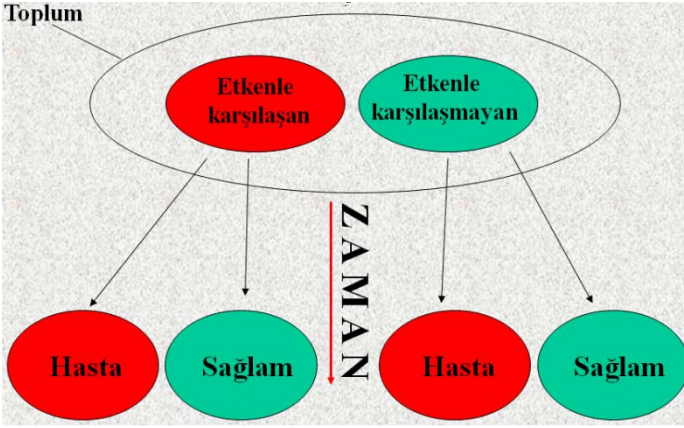
Araştırma yöntemleri

Araştırma sorumuzu ortaya koyup literatür taraması yaptıktan sonra sorumuzun cevabını hangi araştırma tasarımıyla ortaya koyabileceğimizi belirlemeliyiz. Araştırma çeşitlerini aşağıdaki gibi sınıflandırabiliriz:

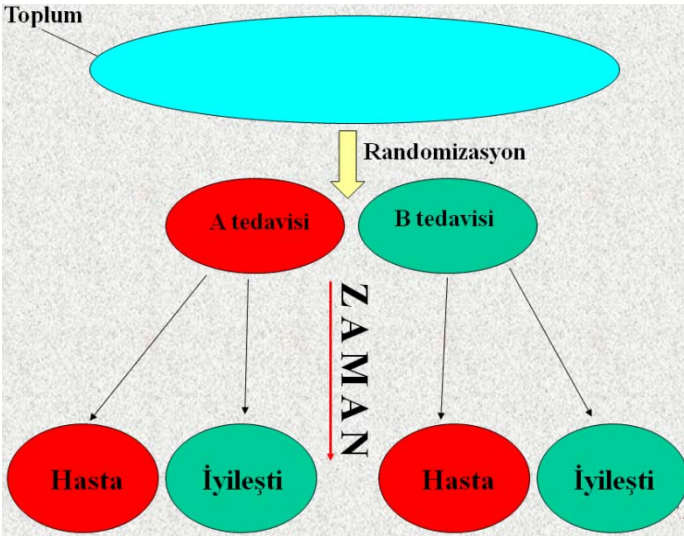
- Gözlemsel araştırmalar
 - Vaka sunumu (case report)
 - Prevalans (yaygınlık)
 - Vaka kontrol (case control) (**Şekil 2**)
 - Kalitatif araştırmalar
 - Kohort çalışmaları (**Şekil 3**)
- Deneysel araştırmalar
 - Randomize kontrollü çalışmalar (**Şekil 4**)
 - Saha deneyleri
- Metodolojik
 - Geçerlilik, güvenilirlik çalışması
 - Özgüllük, duyarlılık çalışmaları



Şekil 2: Vaka Kontrol Çalışması



Şekil 3: Kohort çalışması



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:

- Ordinal
- Nominal

Bazı önemlilik testleri:

- Parametrik
 - Student t-test
 - ANOVA
- Nonparametrik
 - Ki kare
 - Mann-Whitney U testi
 - McNemar testi
 - Wilcoxon
 - Fisher exact test

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:

Yanıtlanabilir soru oluşturmak

- Üniversite eğitimine başlayan gençlerde aile planlaması yöntemi bilenlerle bilmeyenler arasında cinsel ilişkide bulunma açısından bir fark var mıdır?
- Uzm. Dr. Turan Set. Üniversite çağındaki ergenlerin cinsel bilgi tutum ve davranışları. Uzmanlık Tezi, Edirne 2005.
(Verileri kullanmamıza izin verdiği için Dr. Turan Set'e teşekkür ederiz.)

Soru daha önce cevaplanmış mı?

- 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.
 - The most common reason for not using modern contraceptive methods among sexually active respondents was little or no knowledge of contraceptives

Hipotez oluşturma

- H_0 : Cinsel ilişkide bulunan ve bulunmayan gençler arasında aile planlaması yöntemlerini bilme açısından fark yoktur.
- H_1 : Cinsel ilişkide bulunan ve bulunmayan gençler arasında aile planlaması yöntemlerini bilme açısından fark vardır.

Yöntem

- Bu araştırma için kesitsel tanımlayıcı bir çalışma tasarımı uygundur.
- Üniversiteye kayıt için gelen gençlere bir anket doldurtulması
 - Yaş
 - Cinsiyet
 - AP yöntemi bilme durumu
 - Cinsel ilişkide bulunmuş olma durumu

SPSS veri seti

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

14:cinskim																							
anketno	a	yapı1	cinsiyet	m	li	o	e	apyonbil	k	a	e	k	e	a	s	a	cinsel	c	c	h	kacez	cinsfark	g
1	1	*	19	Kız	i		H	Bilmiyor	t	.	H	K	k	E	H		Hayır
2	2	*	22	Kız	y	D	L	E	Bilmiyor	fi	.	.	K	k
3	3	*	17	Kız	.		E	Bilmiyor	fi	H	H	K	K	H	H		Hayır
4	4	*	19	Erkek	ğ		H	Bilmiyor	t	H	H	K	k	.	E		Evet	*	P	H		1	1
5	5	*	19	Erkek	s	D		E	Bilmiyor	fi	.	.	K	K	H	H		Hayır
6	6	*	17	Erkek	t		E	Bilmiyor	fi	E	H	K	k	H	E		Evet	*	P	E		5	3
7	7	*	18	Erkek	t		E	Bilmiyor	k	.	H	T	k	H	E		Hayır
8	8	*	18	Kız	d	S	L	E	Bilmiyor	a	H	H	T	k	H	H		Hayır
9	9	*	17	Kız	d	L	E	Bilmiyor	k	H	H	K	k	H	H		Hayır
10	10	*	17	Kız	k	D		E	Bilmiyor	t	H	H	K	k	E	H		Hayır
11	11	*	17	Kız	k	D		E	Bilmiyor	fi	H	H	K	k	H	.		Hayır
12	12	*	.	Erkek	.	D	L	E	Bilmiyor	fi	H	H	K	k	H	H		Hayır
13	13	*	17	Kız	.		E	Bilmiyor	t	E	H	K	k	H	E		Evet	*	S	E		3	1
14	14	*	17	Kız	e		E	Bilmiyor	k	.	H	T	k	H	H		Hayır
15	15	*	.	Erkek	i		E	Bilmiyor	fi	H	.	T	k	H	H		Hayır
16	16	*	19	Kız	ç	S		E	Bilmiyor	t	.	.	K	K	H	H		
17	17	*	19	Erkek	e		E	Bilmiyor	t	E	H	T	T	H	.		Hayır
18	18	*	17	Kız	y		E	Bilmiyor	k	E	E	T	K	E	E		Hayır
19	19	*	.	Erkek	t	A	L	E	Bilmiyor	k	E	H	K	K	H	H		Hayır
20	20	*	20	Erkek	i	D	L	E	Bilmiyor	t	.	H	K	K	H	H		Hayır
21	21	*	18	Erkek	.		E	Bilmiyor	t	E	H	T	T	H	.		Evet

Analiz edeceğimiz verileri sınıflandıracak olursak:

- Bağımsız değişkenler
 - Yaş
 - Ordinal¹
 - Cinsiyet
 - Nominal²
 - AP yöntemi bilme durumu
 - Nominal
- Bağımlı değişkenler
 - Cinsel ilişkide bulunmuş olma
 - Nominal

¹ İng. "Order" = Sıralı² İng. "Nomination" = Atanmış

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

Tanımlayıcı istatistikler

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.

AP yöntemi * hiç cinsel ilipkiniz oldu mu? Crosstabulation

		hiç cinsel ilipkiniz oldu mu?		Total
		Evet	Hayır	
AP yöntemi	Bilmiyor	Count 3169 % within hiç cinsel ilipkiniz oldu mu? 77,9%	8762 84,1%	11931 82,4%
	Biliyor	Count 900 % within hiç cinsel ilipkiniz oldu mu? 22,1%	1653 15,9%	2553 17,6%
Total		Count 4069 % within hiç cinsel ilipkiniz oldu mu? 100,0%	10415 100,0%	14484 100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	78,644 ^b	1	,000		
Continuity Correction ^a	78,215	1	,000		
Likelihood Ratio	75,920	1	,000		
Fisher's Exact Test				,000	,000
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

- Cinsel ilişkide bulunanların %22'si AP yöntemlerini bilirken bulunmayanların %16'sı bilmemektedir. Cinsel ilişki ile AP yöntemi bilme arasında iki yönlü anlamlı bir ilişki vardır (Ki kare 78,6; $p < 0,001$)
- AP yöntemi bilme durumu her iki grupta da düşüktür.
- Gençlerin AP yöntemi bilme durumlarına göre gelecekteki davranışlarını araştıran kohort çalışmaları daha yararlı bilgi sağlayacaktır.

2. İstatistiksel testler ve kullanım yerleri – akış şemaları

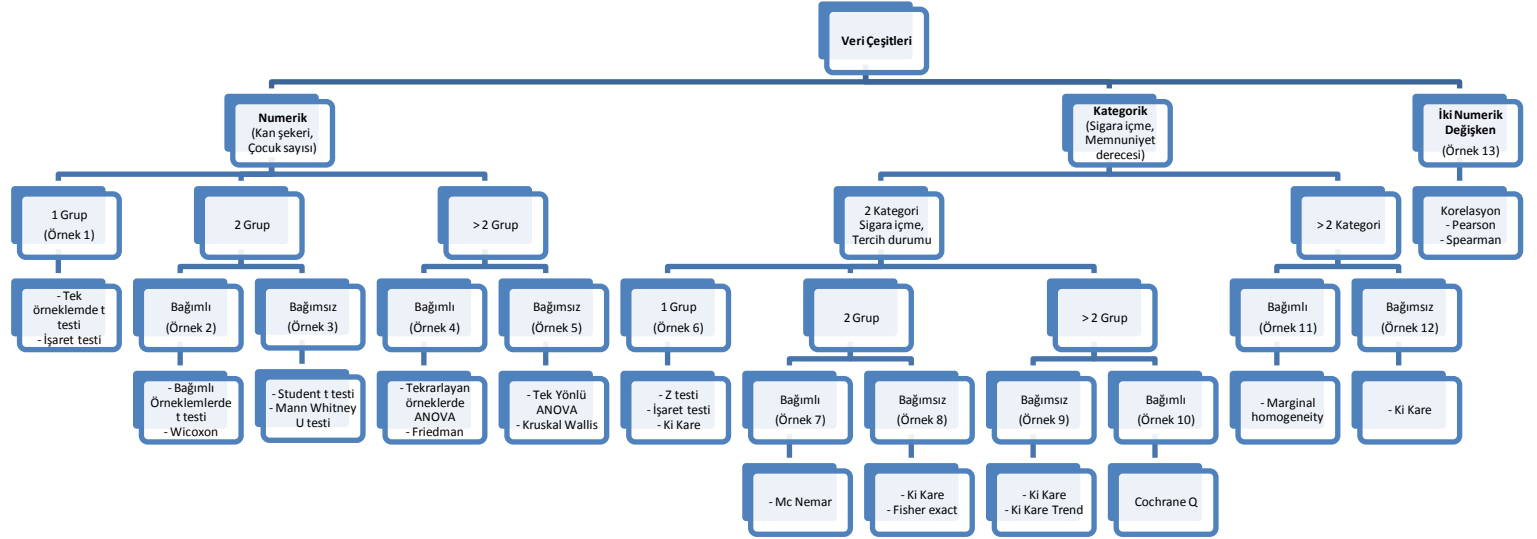
Amaç: Bu konu sonunda okuyucuların sağlık çalışmalarında kullanılan temel istatistik önemlilik testleri ve kullanım yerleri konusunda bilgi sahibi olmaları amaçlanmıştır.

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
 - Tekrarlayan ölçümlerde ANOVA
 - Friedman analizi
 - Z testi
 - Ki kare testi
 - McNemar testi
 - Ki kare trend testi
 - Fisher exact testi
 - Cochran Q testi
 - Pearson korelasyon analizi
 - Spearman korelasyon analizi
- Parametrik ve nonparametrik test ayırımını 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 belki pratik kazanılıncaya kadar araştırmacının baş ucunda asılı bulunmalıdır.

Hipotez testlerinden öte, ileri istatistiksel analizler için de bir akış şeması hazırlanmıştır.



Şekil 5: Önemlilik (hipotez) testleri ve kullanım yerleri akış şeması.

3. Veri çeşitleri

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 **numerik** 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).

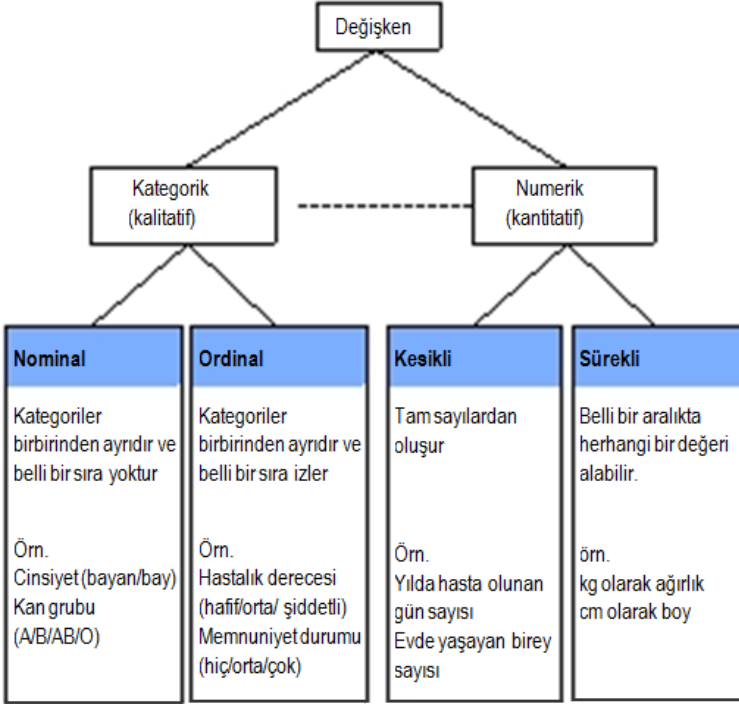
Bir kategorik değişkenin sadece iki özelliği olması halinde ona binary (**ikili**) veya dichotomous (**dikotom**) değişken denir (evet/hayır; canlı/ölü gibi).

Değişkenimiz, sınıflardan oluşmak yerine belli sayıları alan numerik 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], Hemoglobün 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 numerik değişkenler ortalama ve standart sapma gibi ölçütlerle gösterilir.

Kategorik ve numerik 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 numerik olarak ayırt edilir (tam sayılı ve sürekli numerik veriler aynı şekilde işlem görür).



Şekil 6: Veri çeşitleri.

Veri çeşitlerini ayırt etmedeki önemli bir nokta da veri toplama sırasında karşımıza çıkmaktadır. Değişkenimizin numerik olması halinde veriyi de numerik 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 değerin 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 araştırmanın sonlanma noktasındaki veriyi elde edemeyeceğiz.

4. Veri girişi

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’de 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 ayrılar 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 15’te 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 Excell formatında veri ihracı (export) yapılabilmektedir [File>Export to database].

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
Hemogloblin 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 () 3 Ortaokul () 4 Lise 5 () Üniversite
S4 Hemogloblin 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 girisinin (sekreterin) bulunduğu araştırmalarda veri girişi ve analiz aşamalarında sıkıntılarla karşılaşılabilir.

Kategorik veriler

İstatistik programlarının çoğu numerik 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 girsek 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:.....

Numerik veriler

Daha önce de belirtildiği gibi, numerik 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 () Cevap yok

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 hastasıınız?: 3,5
4. En son HbA1c değeriniz 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. Yařınız: 48
2. Cinsiyetiniz:
 - b. Erkek
3. Ka yıldır diyabet hastasıdır?: 1
4. En son HbA1c deđeriniz katı? %7,5
5. Son iki hafta ierisinde hastalıđınız nedeniyle gnlk iřlerinizi yapmakta ne kadar zorlandınız?
 - d. Hi zorlanmadım
6. Diyabet iin hangi ilaları kullanıyorsunuz? Yazınız.
 - a. Glucophage
 - b.....
 - c.....
 - d.....
 - e.....
7. Muayenenizde diyabete bađlı herhangi bir komplikasyon saptandı mı?
 - a. Hayır

özm iin bakınız s. 285

5. Hata ayıklama ve uç değerler

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 nisbeten daha kolaydır. Numerik 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. Numerik 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.:

www.aile.net/agep/istat/ders5hataayiklama.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 (34 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 alınmamıştır (örn. birey alkol kullanma durumunu belirtmek istemeyebilir), 2- bireye uygun olmadığı için veri alınmamıştır (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. Örn:

Hastalık şiddetini değerlendirmek için bir ölçüm yaptığımızı düşünelim. Eğer bazı bireylerin verilerinin eksik olduğunu farkederek, 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. Örn:

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:

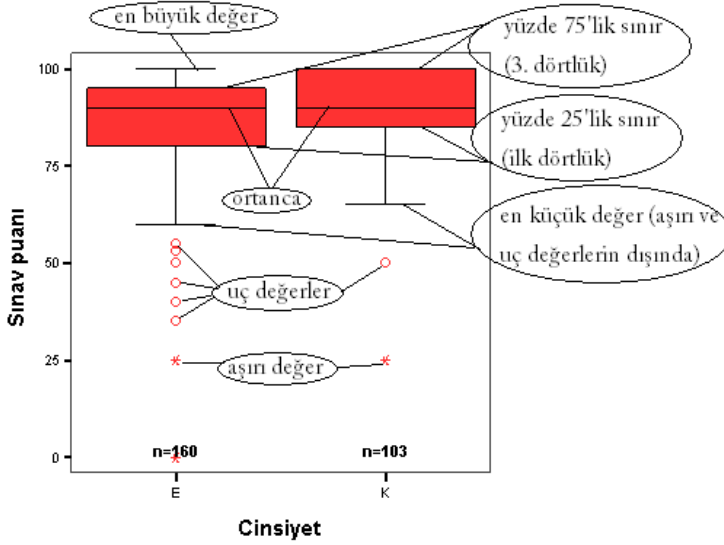
Graphs>Interactive>Boxplot [Y eksenine "Weight", X eksenine "Marital status" değişkenlerini koyalım]>OK

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 (3rd quartile) başladığı sınır. Örn.:

www.aile.net/agep/istat/ders5saplikutugrafik.sav dosyasını SPSS ile açınız.

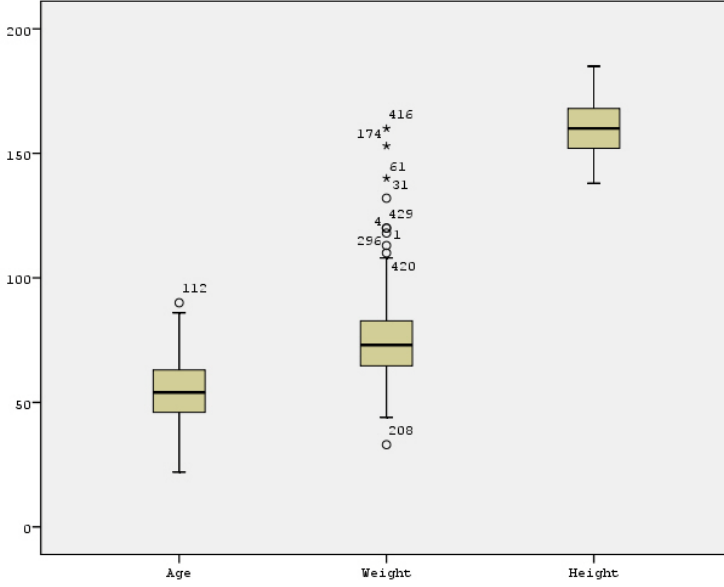
Graphs>Interactive>Boxplot [Y eksenine "Sınav puanı", X eksenine "Cinsiyet" değişkenlerini koyalım]>OK



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. www.aile.net/agep/istat/diyabet.sav 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 alıřtırmada tartıřmaya amak istediĐimiz noktalar okuduĐumuz ilk 3 konuyla ilgili. www.aile.net/agep/istat/odev2.sav adresindeki SPSS dosyasını indiriniz. Bu veriler "Suudi Arabistan'da 50 bin nüfuslu bir řehirde yařayan iřilerin hepatit B bilgisini ve ařılı olma durumlarını arařtırmak" amacıyla yapılmıř bir arařtırmadan alındı. Bu verileri

- "veri eřitleri",
- "veri giriři" ve
- "hata ayıklama ve u deĐerler"

aısından deĐerlendiriniz. DeĐerlendirirken sistematik olmaya alıřınız. İřlediĐimiz konulara uygun yapılanları ve daha iyi nasıl olabileceĐini, siz olsaydınız neleri farklı yapacaĐınızı belirtiniz.

özümü için bakınız s. 286

6. Verilerin grafiklerle gösterilmesi

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
 - Bar grafik
 - Pasta dilimi
 - Histogram
 - Dot plot
 - Boxplot (saplı kutu grafiği, asansör grafiği)
 - Scatter diagram (saçılma grafiği, nokta grafikler)
- Verilere uygun grafik seçebilmek
 - Tek değişken olduğunda
 - Çok değişken olduğunda
 - Numerik 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.

Frekans dağılımı

Ampirik (deneysel; empirical) frekans dağılımı, her bir değişkenle ilgili gözlemlerin, bunların sınıflarının, veya kategorilerinin gözlem **frekansı** (sıklığı) açısından gösterilmesidir.

Frekans dağılımlarının gösterilmesi

Frekans dağılımları kategorik veriler veya bazı tam sayılı numerik veriler için kullanılır; verilerin görsel olarak sunulmasını sağlar.

- **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.:

www.aile.net/agep/istat/diyabet.sav 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.

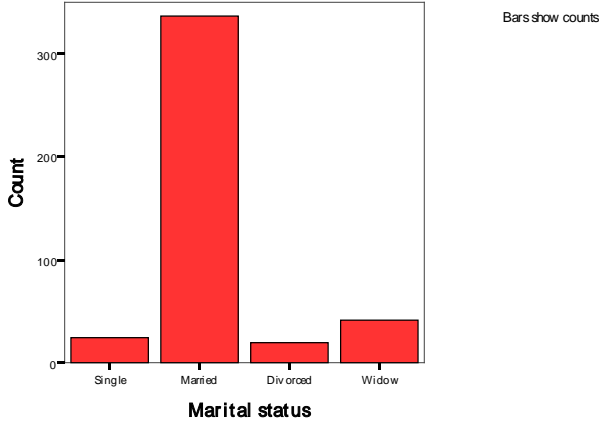
Marital status

		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:



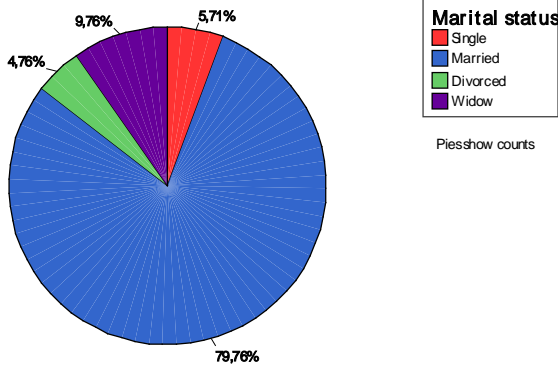
Grafiğimizde sütunların üzerinde sayı ve/veya yüzdelerin görülmesini istersek OK butonunu tıklamadan önce sekmelerden gerekli tercihleri yapabiliriz veya daha sonra grafik çıktısının üzerine çift tıklayarak da ayarları değiştirebiliriz.

- **Pasta grafiği:** daire şeklindeki bir pasta her bir dilimi değişkenin ilgili kategorisinin frekansını temsil edecek şekilde dilimlere ayrılır. Örn.:

www.aile.net/agep/istat/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:

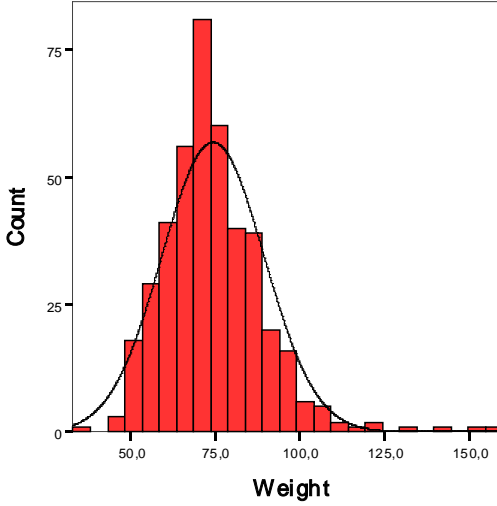


- **Histogram:** bu grafik de sütun grafiğine benzer. Ancak, sütunlar arasında boşluk yoktur. Sütunlar bir değeri değil, belli aralıktaki değerleri temsil eder. Bu grafik daha çok numerik değişkenlerde verilerin dağılım şeklini incelemek için kullanılır. Örn.:

www.aile.net/agep/istat/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:



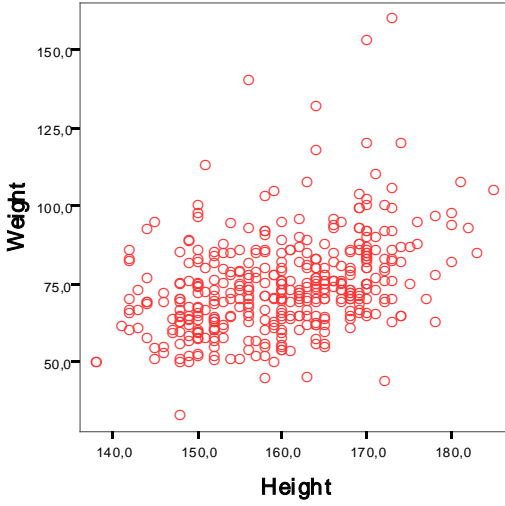
Verilerimizin 75 kg civarında yoğunlaştığını, dağılımın çan eğrisine benzediğini, çan eğrisinin sağ taraftaki kuyruğunun biraz daha uzun olduğunu görüyoruz.

- **Dot plot:** sık kullanmadığımız bir grafik şeklidir. Her bir gözlem yatay veya dikey ekseninde bir nokta olarak gösterilir. Büyük veri setlerinde bu grafikleri yorumlamak zor olabilir (noktalar üst üste binecektir).
- **Boxplot (saphı kutu grafiđi):** Asansör grafiđi, kutu ve sakal grafiđi (box and whisker plot) de denir. “Hata ayıklama ve uç deđerler” konusuna bakınız.
- **Scatter diagram** (saçılma grafiđi, nokta grafikler): iki ayrı numerik veya ordinal deđişkeni grafikte göstermenin bir yolu nokta grafikleridir. Örn.:

www.aile.net/agep/istat/diyabet.sav dosyasını SPSS ile açınız. “Height” ve “Weight” deđişkenlerinin histogram grafiđini çizelim.

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

Aşağıdaki grafiđi elde ederiz:



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

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

7. Merkezi dağılım (yığılma) ölçütleri

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

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Merkezi eğilim ölçütlerini sayabilmeli
- 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 ağırlandırma 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 dağılım 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

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad \text{veya} \quad \bar{x} = \frac{\sum x_i}{n}$$

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 $(10+13+15+16+18+19+20+20)/9 = 16,33$ 'tür.

Ortanca

Verilerimizi büyükten küçüğe doğru sıraladığımızda ortadaki değere **ortanca** denir. Eğer veri adedimiz çift sayı ise ortadaki iki değerlerin 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 dağılım 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

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.:

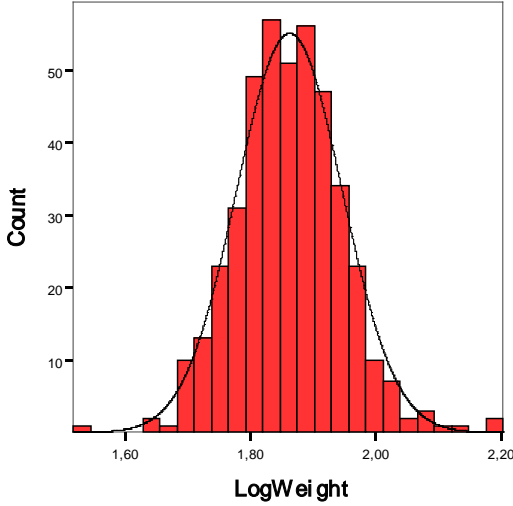
www.aile.net/agep/istat/diyabet.sav veri setinde "Weight" değişkenini incelediğimizde histogram grafiğinde çan eğrisinin kuyruğunun sağa doğru olduğunu (sağa eğimli) görmüştük. Şimdi "weight" değişkeninin logaritmasını alarak yeni bir değişken oluşturalım:

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 almamız gerekir. $10^{1.862} = 72,777$ 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.

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 değere katsayısını çarpıp bunları toplarız. Sonra değerlerin 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:

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

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

Data>Wiegth 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.

Merkezi dağılım ölçütlerinin (MDÖ) olumlu ve olumsuz yönleri aşağıdaki tabloda görülmektedir:

MDÖ türü	Olumlu	Olumsuz
Aritmetik ortalama	<ul style="list-style-type: none"> tüm değerler kullanılır cebirsal olarak tanımlanmış ve matematiksel olarak kullanılabilir bilinen örneklem dağılımı (bakınız Bölüm: veri dönüştürme) 	<ul style="list-style-type: none"> uç değerlerden etkilenir eğimli verilerden etkilenir
Medyan (Ortanca)	<ul style="list-style-type: none"> uç değerlerden etkilenmez eğimli (skewed) veriden etkilenmez 	<ul style="list-style-type: none"> bilginin büyük bir kısmı ihmal edilir cebirsal olarak tanımlanmamıştır örneklem dağılımından etkilenir
Mod	<ul style="list-style-type: none"> kategorik veriler için kolayca saptanabilir 	<ul style="list-style-type: none"> bilginin büyük bir kısmı ihmal edilir cebirsal olarak tanımlanmamıştır örneklem dağılımı bilinmez
Geometrik ortalama	<ul style="list-style-type: none"> geri dönüştürme yapılmadan önce aritmetik ortalama ile aynı avantajlara sahiptir sağa doğru eğimli veriler için uygundur 	<ul style="list-style-type: none"> sadece log dönüşüm simetrik bir dağılım oluşturuyorsa işe yarar
Ağırlıklı ortalama	<ul style="list-style-type: none"> aritmetik ortalama ile aynı avantajlara sahiptir her bir gözleme nisbi bir önem verilir cebirsal olarak tanımlanmıştır 	<ul style="list-style-type: none"> ağırlık bilinmeli veya hesap edilmelidir

Alıştırma 3

<http://www.aile.net/agep/istat/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 çiziniz 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 kaçtır? (analyze-descriptive statistics-frequencies sekmesinden). Ayrıca ortancaya verileri sıralayarak ta bakınız.
8. Boya göre ağırlıklandırılmış (düzeltilmiş) kilo değişkeninin aritmetik ortalaması kaçtır?
9. Araştırmada kaç erkek, kaç kadın vardır? Boy değişkeninin aritmetik ortalaması kaçtır? Cinsiyete göre ağırlıklandırma yapınız. Şimdi boy değişkeninin aritmetik ortalaması kaç oldu?

Çözüm için bakınız s. 287

8. Yaygınlık ölçütleri

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 numerik değişken hakkında biri merkezi dağılım diğeri de yaygınlık ölçütü olmak üzere iki özelliğini belirtmemiz halinde verilerimizin yapısını yeterince özetlemiş oluruz. Bir önceki bölümde merkezi dağılım 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:

<i>No</i>	<i>Boy</i>
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:

$$1. \text{ çeyrek} = [148 + ((154-148)*0,25)]=150$$

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

$$3. \text{ çeyrek} = [170 + ((176-170)*0,75=175)]$$

Tam %50 sınırındaki değere “ortanca” denir. 25-75. Persantiller arasına ise **interquartile range** (dörtlülük çeyrek değerler genişliği-DÇDG) denir. DÇDG, veriler sıralandığında orta kısımda kalan %50 lik bölümü gösterir. Verimisin 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 değerlerin 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 hesaplarırken 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 kare köküdür.

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

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

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

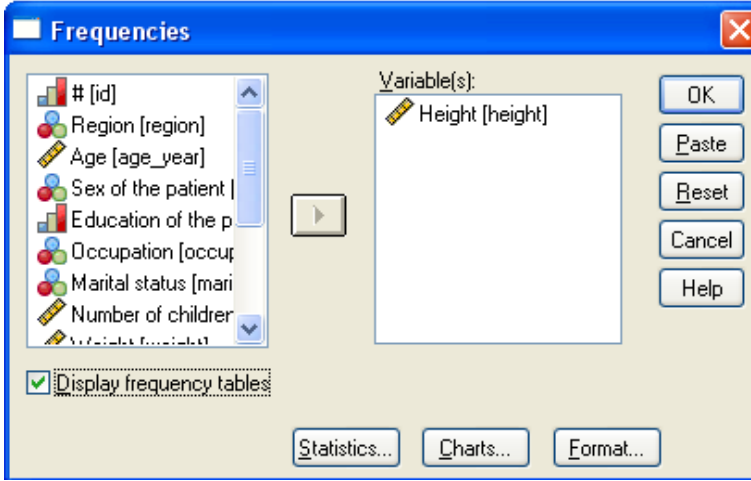
$$\text{Varyans} = 726 / 9 = 80,67$$

$$\text{Standart sapma} = 8,98$$

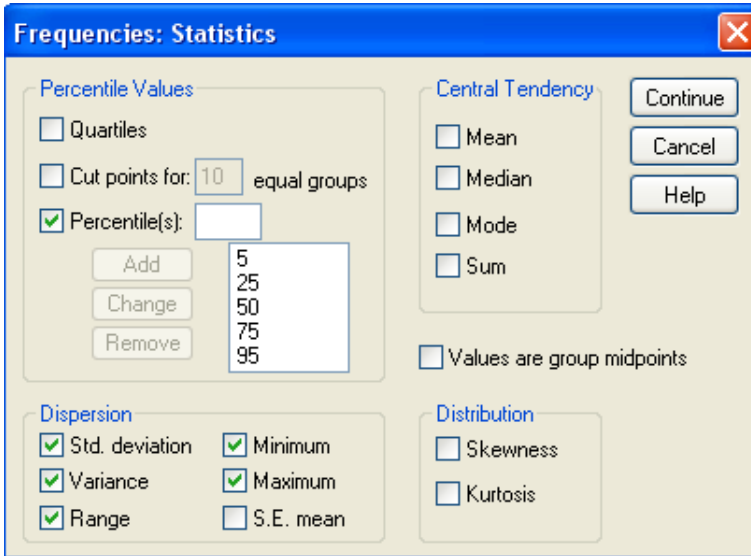
SPSS ile hesaplama

www.aile.net/agep/istat/diyabet.sav 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 $84,6 \text{ cm}^2$, standart sapması $9,2 \text{ cm}$, 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)	<ul style="list-style-type: none"> Kolayca saptanabilir 	<ul style="list-style-type: none"> Sadece iki gözlemi kullanır Uç değerlerden etkilenir Örnekleme sayısı arttıkça artma eğilimindedir
Persantillere dayalı aralık	<ul style="list-style-type: none"> Genelde uç değerlerden etkilenmez Örnekleme sayısından bağımsızdır 	<ul style="list-style-type: none"> Hesaplanması hantaldır Küçük örnekler için hesaplanamaz Sadece iki gözlem

	<ul style="list-style-type: none">• Eğimli veriler için uygundur	kullanır <ul style="list-style-type: none">• Cebirsel olarak tanımlanmamıştır
Varyans	<ul style="list-style-type: none">• Her gözlemi dikkate alır• Cebirsel olarak tanımlanmıştır	<ul style="list-style-type: none">• Ölçüm birimi ham verinin karesidir• Uç değerlerden etkilenir• Eğimli veriler için uygun değildir
Standart sapma	<ul style="list-style-type: none">• Varyansla aynı avantajlara sahip• Ölçüm birimi ham verininkiyle aynıdır• Kolayca yorumlanabilir	<ul style="list-style-type: none">• Uç değerlerden etkilenir• Eğimli veriler için uygun değildir

9. Teorik Dağılımlar: Normal Dağılım

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
 - A priori (önsel)
- Olasılık kurallarını açıklayabilmeli
 - Toplama kuralı (mutually exclusive) (veya)
 - Çarpma kuralı (ve)
- 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 gerçekleşmez. Olasılığın *bir* olması halinde ise olay kesinlikle gerçekleşir. Bir de **şartlı olasılık** (conditional probability) vardır. Şartlı 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 oluşturduğumuz 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:** A ve B olaylarının birbirini dışlayan (mutually exclusive) olmaları halinde geçerlidir. Ya A olayı gerçekleşecektir veya B. Örn.

Bir kişide uyku apne sendromu (UAS) görülme olasılığı %5, diyabet görülme olasılığı ise %9'dur. Bu kişide UAS veya diyabet görülme olasılığı %5+%9=%14 olur.

Bir kişinin bazı dişlerinin eksik olma olasılığı (BDE) %67, hiç eksik dişinin olmaması (TAM) olasılığı ise %24'tür. Bu durumda kişinin ağızında diş olma olasılığı BDE veya TAM yani $0,67+0,24 = \%91$ olacaktır.

- **Çarpma kuralı:** A ve B olaylarının birbirinden bağımsız olması durumunda geçerlidir. A olayı ve B olayının bir arada gerçekleşme olasılığıdır. Yukarıdaki örneklere göre yapacak olursak:

Bir kişide uyku apne sendromu (UAS) görülme olasılığı %5, diyabet görülme olasılığı ise %9'dur. Bu kişide UAS ve diyabet görülme olasılığı $0,05 \times 0,09 = \%0,045$ olur.

Bir kişinin bazı dişlerinin eksik olma olasılığı (BDE) %67'dir. Bu durumda diş hekiminde bekleyen iki hastada aynı anda (A ve B hastasında) BDE olma olasılığı $0,67 \times 0,67 = \%44,9$ 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.

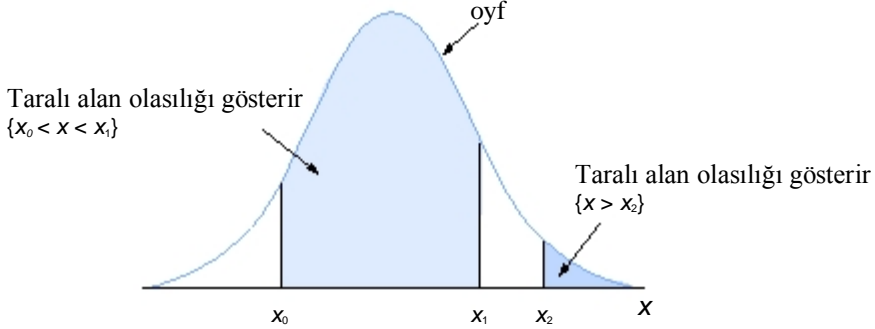
Her bir olasılık dağılımı belirli **parametrelerle** (örn. Ortalama, varyans..) tanımlanır. Bu parametreleri bilmemiz halinde ilgili teorik dağılımı tanımlayabiliriz. İşte istatistikte örneklemimizdeki bu parametreler tahmin edilerek hesap yapılır.

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

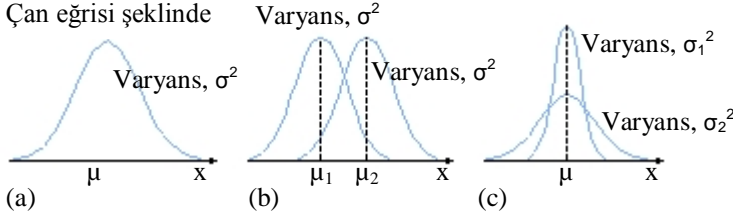
- **Tam sayı (discrete) teorik dağılımlar:** rastgele değişkenin mümkün olan her bir değeri için olasılıklar hesaplayabiliriz. Bu olasılıkların toplamı birdir. Binomiyal ve Poisson dağılımı bu tür teorik dağılımlardandır.
- **Sürekli (continuous) teorik dağılımlar:** x değişkeni sonsuz sayıda değer alabileceğinden 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)

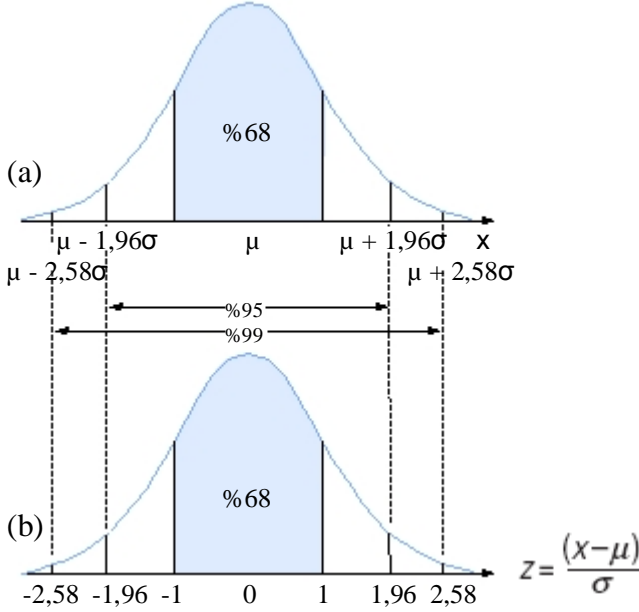


Şekil: x için olasılık yoğunluğu fonksiyonu (oyf)



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 yoğunluğu 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 eğimli olur (varyanslar sabit kaldığında);
- Varyans arttığında yassılaştır, 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 kaçını 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

Çözüm için bakınız s. 290.

10. Teorik Dağılımlar: Diğer Dağılımlar

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

- WS Gossett tarafından bulunmuştur. 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 sayısı ile ilgidir 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

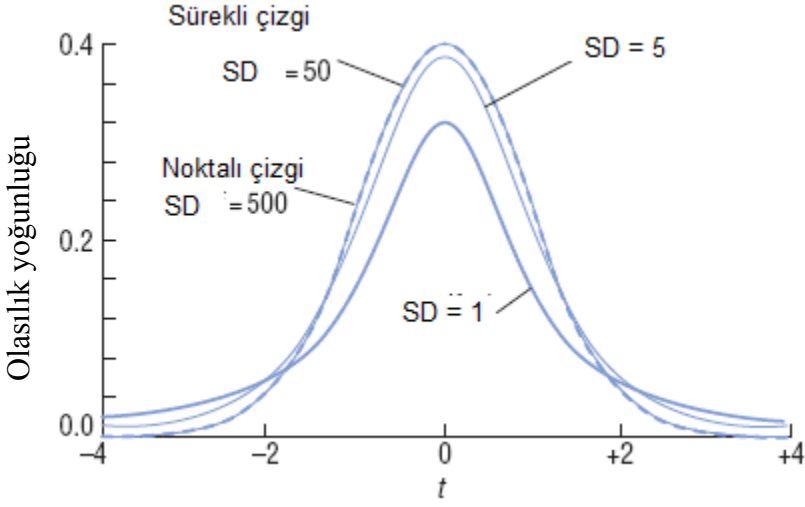
- Pozitif değerleri olan sağa eğimli bir dağılımdır.
- **Serbestlik derecesiyle** tanımlanır.
- Şekli serbestlik derecesine bağlıdır; serbestlik derecesi arttıkça daha simetrik ve normal dağılıma benzer hale gelir.
- Özellikle kategorik değişkenlerin analizinde yararlıdır.

F-dağılımı

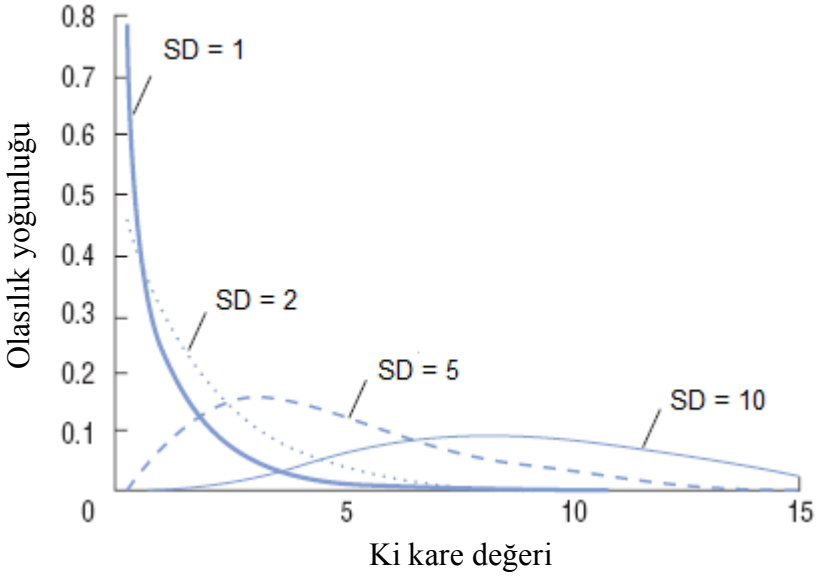
- Sağa eğimlidir.
- Bir oran tarafından belirlenir. Normal veriden hesaplanan iki tahmini varyansın birbirine oranı F -dağılımını verir.
- F -dağılımını tanımlayan iki parametre **pay ve paydasının serbestlik dereceleri**dir.
- F -dağılımı özellikle iki varyansın karşılaştırılmasında ve varyans analizini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

Lognormal dağılım

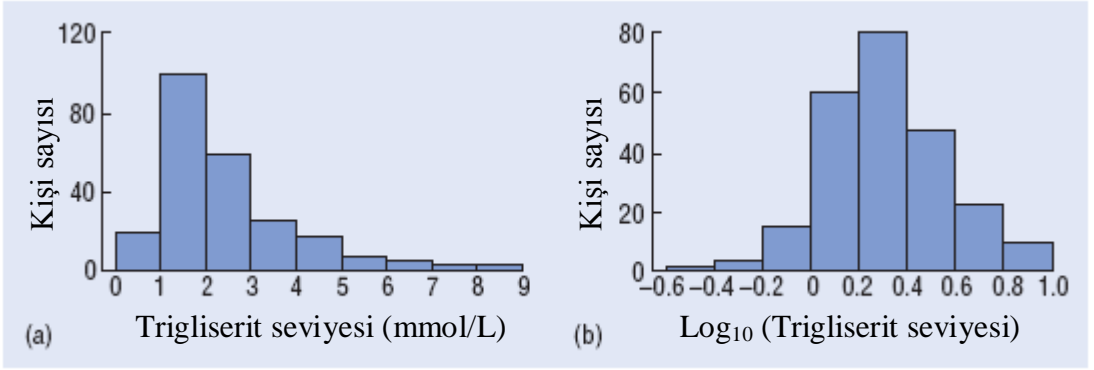
- 10 veya e tabanına göre logaritmasının normal dağılım gösterdiği rastgele bir değişkenin olasılık dağılımıdır.
- İleri derecede sağa eğimlidir.
- Eğer sağa doğru eğimli bir veriye logaritmik dönüşüm uyguladığımızda ortaya çıkan veri normal dağılıma yaklaşıyorsa verimizin lognormal dağılıma sahip olduğunu 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.



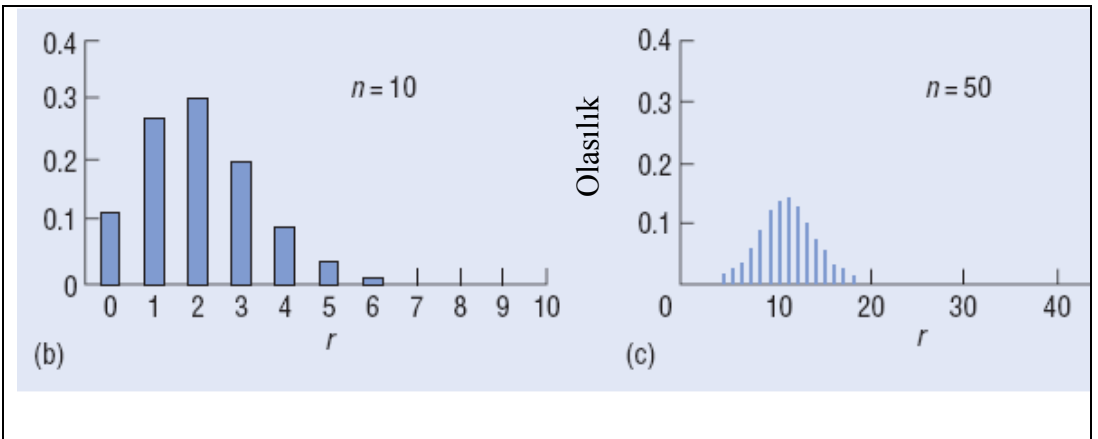
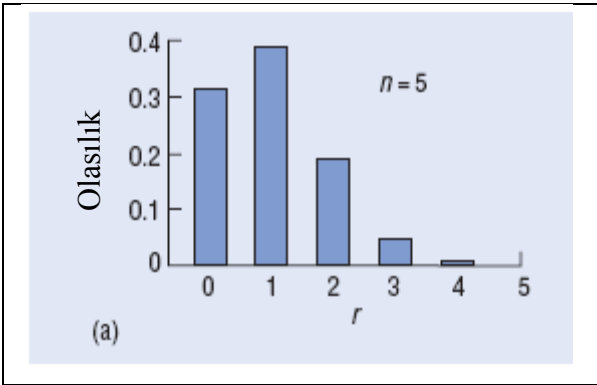
Şekil: Çeşitli serbestlik dereceleri (SD) olan t -dağılımları.



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) $n = 50$.

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 örneklemdaki 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.

11. Veri Dönüştürme

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 teste **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 uygulamalıyız (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 kare kökünü 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;

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

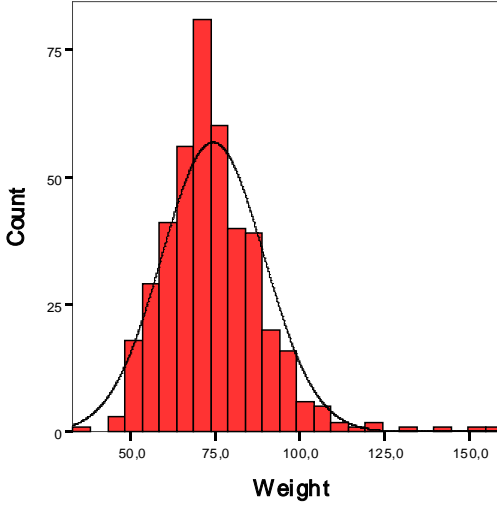
$$\text{Antilog}(2) = 10^2 = 100$$

- y sağa eğimli olduğunda “ $z = \log y$ ” genelde yaklaşık olarak **normal dağılıma** sahip olur. Bu durumda y’nin lognormal dağılıma sahip olduğundan bahsedebiliriz.
- y ile x değişkeni arasında üstsel 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. www.aile.net/agep/istat/diyabet.sav 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

<i>N</i>	<i>Valid</i>	424
	<i>Missing</i>	6
<i>Skewness</i>		1,329
<i>Std. Error of Skewness</i>		,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.

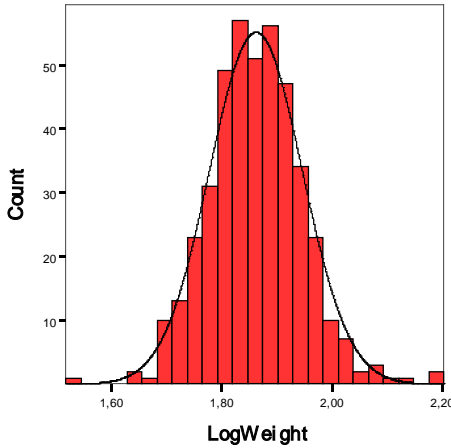
Şimdi "weight" değişkeninin logaritmasını alarak yeni bir değişken oluşturalım:

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

Stadart 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 numerik 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 (numerik 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			95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper	
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 örneklerde t testi yaptığımızda varyansların homojen olduğunu görüyoruz (Levene p=0,877):

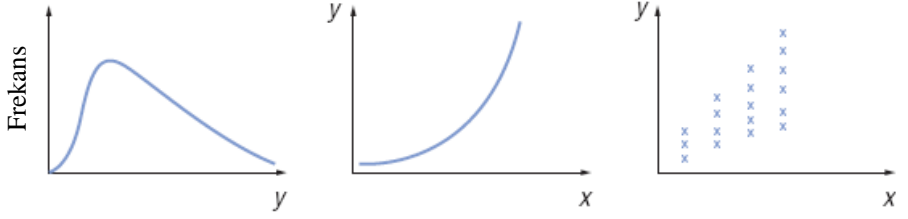
Independent Samples Test											
		Levene's Test for Equality of Variances					t-test for Equality of Means			95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper	
lognumchild	Equal variances assumed	,024	,877	1,807	319	,072	,05105	,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 dönüşünden sonra aradaki farkın anlamsız çıktığına (p=0,072) dikkat ediniz.

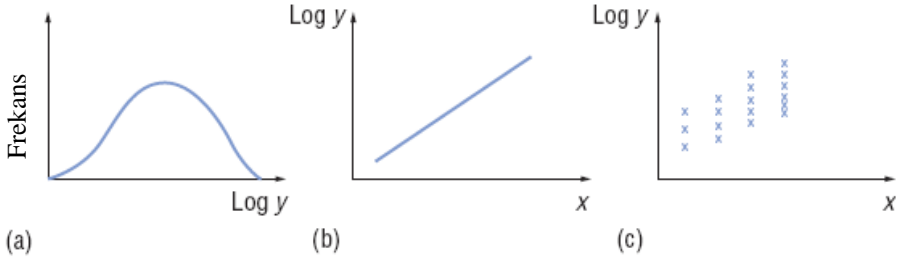
Karekök dönüşümü $z = \sqrt{y}$

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/artmetik ortalama sabit olduğunda) **varyansı stabilize edici** etkisi de vardır. Karekökü dönüşümü genelde Poisson türü değişkenler için kullanılır. Ayrıca, eksi sayıların karekökünün alınmayacağını da hatırlatmak gerekir.

Dönüşümden önce

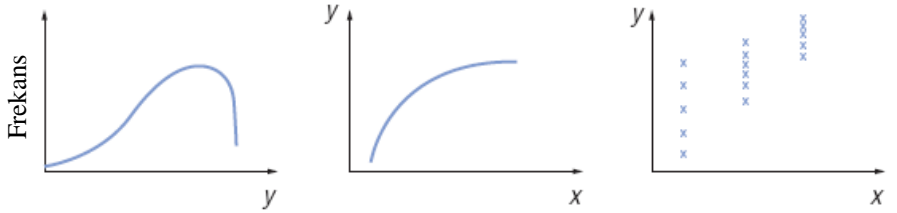


Dönüşümden sonra

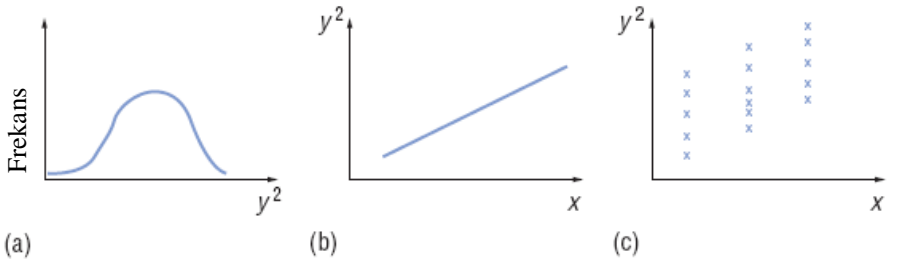


Şekil: Logaritmik dönüşümün etkileri: (a) normalize edici, (b) linearize edici, (c) varyans stabilize edici.

Dönüşümden önce



Dönüşümden sonra



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.

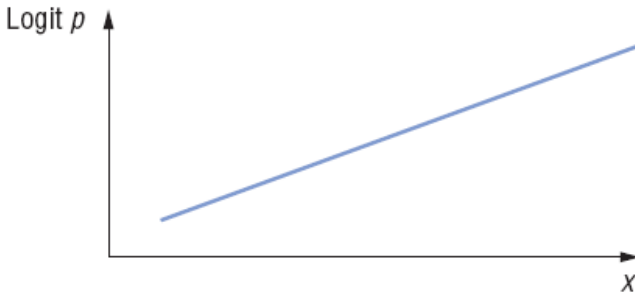
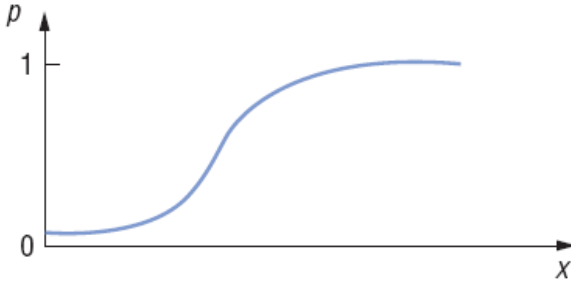
- y sola eğimli olduğunda $z=y^2$ genelde **normal** dağılır.
- x ve y arasındaki ilişkinin grafiği çizildiğinde aşağıday doğru bir eğri oluştuğunda $z=y^2$ ve x arasındaki ilişki yaklaşık olarak **lineer** olur.
- Sürekli bir değişken olan y'nin varyansı y arttıkça azalma meylindeyse kare dönüşümü $z=y^2$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

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

Logit (logistic) dönüşüm,

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.

12. Örneklem ve Örneklem Dağılımları

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 örneklemden 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 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** ve **sistemik 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 edilebilir.

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

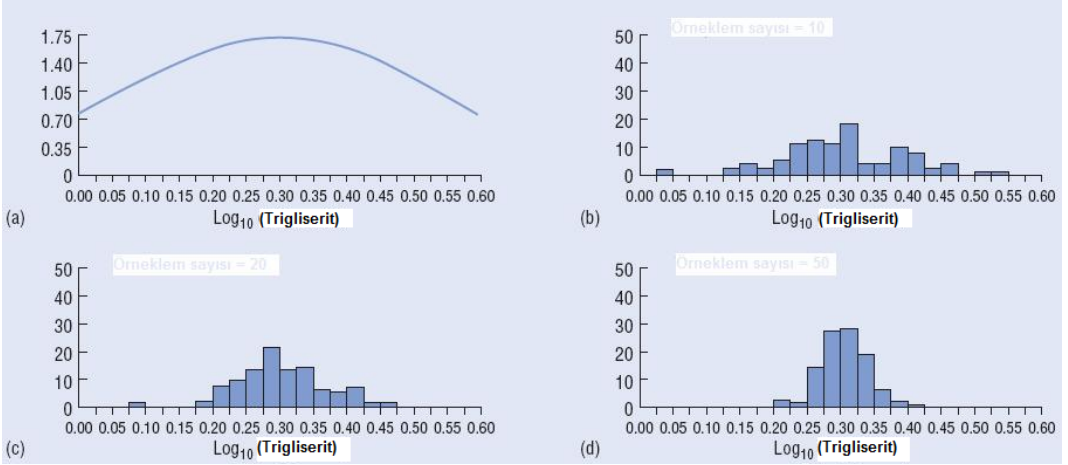
Örneklem varyasyonu

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 olabiliriz ve böylece örneklem hatamızın düzeyi hakkında fikrimiz olur.

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

Ortalamanın örneklem dağılımı

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



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 $\text{SEM} = \sigma / \sqrt{n}$ 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

$\text{SEM} = s / \sqrt{n}$ 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.

Aliřtırma 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 arařtırmada kullanılan verileri çeřitlerine göre sınıflandırınız?
2. Bu arařtırmada sistolik kan basıncı ve vücut kitle endeksi arasındaki iliřki ne tür bir grafikte gösterilmiřtir?
3. Arařtırmanın sonuçlarından merkezi daęılım ölçütleri ve yaygınlık ölçütlerine örnekler veriniz.
4. Arařtırmada belirtilen “adjusted for sex” ifadesinden ne anlıyorsunuz?
5. Arařtırmada hangi veri dönüřtürmeleri uygulanmıřtır? Neden?
6. Yetkili arařtırmacı siz olsaydınız bu arařtırmanın yönteminde herhangi bir farklılık yapar mıydınız?
7. Bařka 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 incelenmiř. Makaledeki Tablo 1 ařaęıdadır. Tablodan göreceęiniz gibi arařtırmalarda sıklıkla SD yerine SEM kullanılmaktadır. Sizce bu yanlış kullanım okuyucuda ne gibi yanlış anlamalara sebep olur? (Dilerseniz ilgili makalenin tartıřma 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>Anesthesia & 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. 291

13. Güven Aralıkları

Amaç: Bu konu sonunda okuyucunun örnekleme ve örneklem dağılımları 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 dağılım 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 bilgilerimizin 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 yararlanırız:

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). Güven aralığı İngilizce makalelerde Confidence Interval (CI) olarak geçmektedir.

Aslında toplum ortalamasının sabit bir değer olduğunu ve kendisine bir aralığın atfedilemeyeceğini biliyoruz. Ancak, bizler örneklerde 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 dağılması ve/veya toplum varyansını bilmememiz halinde eğer verilerimiz t-dağılımına uygunsa örneklemimizin standart sapmasını kullanarak hesap yapabiliriz. Bu durumda:

Ortalamanın %95 güven aralığı (ortalama – $t_{0,05} \times \text{SEM}$) ile (ortalama + $t_{0,05} \times \text{SEM}$) 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

www.aile.net/agep/istat/diyabet.sav 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	53,26	
	Lower Bound		
	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 dağılımı halinde ampirik dağılımımızın normal dağılıma benzeyeceği bilgisini de hatırlatmalıyız.

Orantılar için güven aralığı

Daha önce bahsedildiği gibi, yüzdelerle 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 örnekleme 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.

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.

- **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?**

Toplumdaki toplumdaki alınan bir örneklemimizin ortalamasının toplumun güven aralığının içerisinde olması durumunda ilgili örneklemimizin toplumdaki 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 toplumdaki 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

<http://lokman.cu.edu.tr/psikiyatri/derindex/kriz/2002/2/29.pdf>

Makalede Tablo 1'deki verilerden yaş için birinci ve ikinci gruplarda 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
<ul style="list-style-type: none"> Ortalama 1 = 51 Standart sapma 1 = 15,39 $n_1=25$ $SD = 24$; $p = 0,05$ 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] 	<ul style="list-style-type: none"> Ortalama 2 = 43,76 Standart sapma 2 = 15,07 $n_2 = 25$ kişi $SD = 24$; $p = 0,05$ 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 yaşın %95 güven aralığının nispeten geniş olduğunu ve grupların güven aralıklarının kesiştiğini görüyoruz.

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

Tablo-4: Hasta yakınlıklarının 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
<ul style="list-style-type: none"> Ortalama 1 = 5,95 Standart sapma 1 = 5,19 $n_1 = 38$ $SD = 37$; $p = 0,05$ için tablo t değeri $\sim 2,02$ $SEM_1 = 5,19 / \sqrt{38} = 0,84$ %95 GA 1 $[5,95 \pm 2,02 \times 0,84]$ [4,25 – 7,65] 	<ul style="list-style-type: none"> 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 $[0,75 \pm 2,201 \times 0,75]$ [-0,90 – 2,40]

Farkların güven aralığı

Nümerik veriler için

Nümerik veri iki grupta farkların standart hatasını hesaplamak için şu formülü kullanırız:

$$SEM(\text{fark}) = \sqrt{[(s_1^2/n_1) + (s_2^2/n_2)]}$$

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 sıfır içerdiği 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 ötekenden büyük olduğunu iddia edemeyiz).

SPSS ile ortalamalar arasındaki farkın güven aralığının hesaplanması

www.aile.net/agep/istat/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:

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

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

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

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 biri (1) içermesi durumunda risk açısından bir anlamlılık olmadığı yorumu yapılabilir.

Relative Deviate formülünün farkların güven aralığına uygulanması

Daha önceden gördüğümüz gibi relative deviate (relatif 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.

14. Araştırma Tasarımı – I

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 boşuna 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).

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 yararlanmasak 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 konulara da dikkat edilmelidir.

Gözlemsel araştırmalarda ise araştırmacı bir müdahalede bulunmayıp sadece gözlem yapar. Bu araştırma tasarımında daha az güvenilir bilgiler elde ederiz. 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

Zaman içerisindeki akışı ve tipik kullanımlarıyla birlikte çeşitli araştırma tasarımları aşağıdaki tabloda özetlenmiştir.

Çeşidi	Zamanlama	Şekli	Geçmiş	Şimdi	Gelecek	Kullanımı
Kesitsel	Zamanda bir an	Gözlemsel		Veri toplama		<ul style="list-style-type: none"> Prevalans Referans aralıkları Bir grubun sağlık durumu
Tekrarlayan kesitsel	Zamanda bir an	Gözlemsel		Veri toplama	Veri toplama Veri toplama	<ul style="list-style-type: none"> Zaman içerisindeki değişim (trend)
Kohort	Longitudinal (prospektif)	Gözlemsel		Kohortun tanımlanması ve risk faktörlerinin değerlendirilmesi	Takip → Sonuçların gözlenmesi	<ul style="list-style-type: none"> Prognoz ve hastalığın doğal seyri Etyoloji
Vaka kontrol	Longitudinal (retrospektif)	Gözlemsel	Risk faktörlerinin değerlendirilmesi	Vakalar ve kontrollerin belirlenmesi	← incele	<ul style="list-style-type: none"> Etyoloji (özellikle nadir hastalıklar için)
DeneySEL	Longitudinal (prospektif)	DeneySEL		Müdahale	Takip → Sonuçların gözlenmesi	<ul style="list-style-type: none"> Tedavi etkisini araştırmak için klinik çalışmalar Önleyici girişimlerin araştırılması (örn. aşılama) Laboratuvar deneyleri

Nedensellik

Araştırma yapmadaki amacımız genelde bir **neden – sonuç ilişkisi** bulmaktır. Neden – sonuç ilişkisini inceleyen en güvenilir araştırmalar deneysel araştırmalardır. Bununla birlikte, gözlemsel araştırmalara bakarak ta nedenler hakkında çıkarımlar yapabiliriz. Ancak, her araştırma sonunda bir nedensellik çıkarımı yapamayabiliriz. Sigara içenlerde içmeyenlere oranla akciğer kanserinin anlamlı derecede daha sık görülmesini 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 araştırma tarafından desteklenmelidir.
- Sebep ve sonuç arasındaki ilişki kuvvetli olmalıdır.
- Sebep ile 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 örnekleme farklılıklarından 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.

DeneySEL 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.
- **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 meylindedir.

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

Alıştırma 6

1. www.aile.net/agep/istat/odev6.sav isimli SPSS dosyasını kullanarak kızlarla erkeklerin yaşları arasında istatistiksel bir fark olup olmadığına bakalım. Söz konusu problemde bağımsız iki grupta numerik verileri karşılaştıracamız (ilk dersimizdeki 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 ta 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	1,37	3,47E-02
	Erkek	1546	18,25	1,30	3,30E-02

Independent Samples Test

	Levene's Test for Equality of Variances	F	Sig.	t	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Equal variances assumed	2,112	,146	2,689	3108	,007	,13	4,79E-02	3,49E-02	,22	
Equal variances not assumed			2,690	3101,685	,007	,13	4,79E-02	3,49E-02	,22	

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 sağlamlasını yapınız.

TABLO 1. Hasta ve Kontrollerin Demografik ve Klinik Özellikleri.

	Hastalar (n=42)	Kontroller (n=23)	Karşılaştırma
Yaş (yıl)	37.57±8.70	33.73±7.16	t=1.80; p>0.05
Cinsiyet (E/K)	21/21	11/12	-
BKI (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. 292

15. Araştırma Tasarımı – II

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. Replike 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:

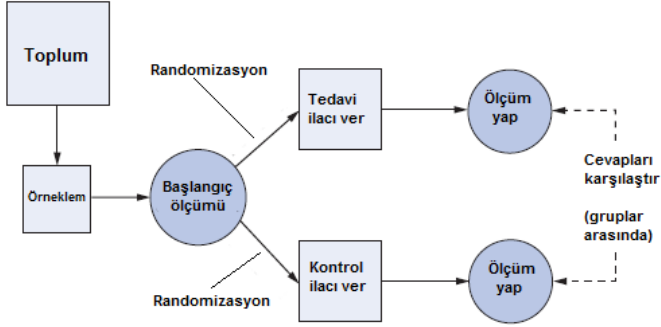
		Beta karoten	
		Hayır	Evet
Aspirin	Hayır	<i>Hiçbiri</i>	<i>Beta karoten</i>
	Evet	<i>Aspirin</i>	<i>Aspirin + beta karoten</i>

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

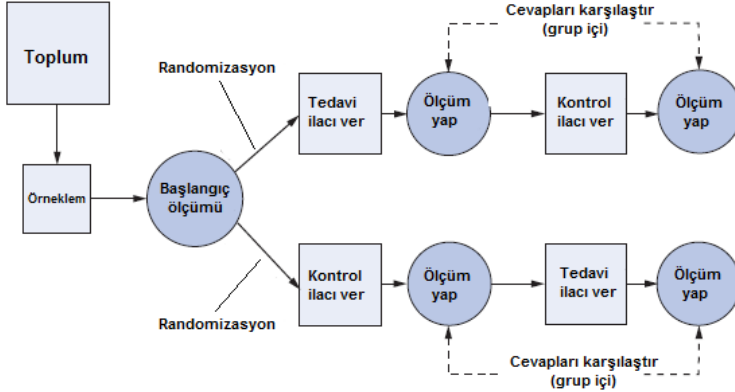
Ayrıca bu tasarımda iki faktör arasında **etkileşim** (interaction) olup olmadığını da araştırabiliriz. Etkileşim saptamamız halinde yukarıdaki örnekte aspirin ve beta karotenin birlikte verilmesinin ayrı ayrı etkilerinden daha fazla bir yarar sağlayacağı anlamına gelir. Bu yönüyle de bu araştırma tasarımı 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).

(a) Paralel



(b) Çapraz



16. Klinik Araştırmalar

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
- Faz 0, I, II, III 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 körlere çeşitlerini açıklayabilmek
 - Tek kör
 - Çift kör
- Etik onay ve bilgilendirilmiş rızanın önemini savunmak
- Araştırmaya dahil etme ve çıkarma kriterlerini açıklayabilmek
- CONSORT statement'ı 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 **klirik araştırmalar** denir. Bu araştırmalar **klirik öncesi** çalışmalar (laboratuvar, hayvan deneyleri) şeklinde olabileceği gibi, ilacın güvenilirliğini araştıran küçük klinik çalışmalar (**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; ve **Faz II** araştırmalar – 20-300 kişilik grupta doz, etkinlik ve toksisiteyi araştıran çalışmalar) veya yeni ilacın daha büyük insan gruplarında tam araştırılması (**Faz III** – 300-3000 kişilik çok merkezli gruplarda ilacın tüm yönleriyle araştırılması) şeklinde olabilir. Ruhsatlandırma işlemi bu aşamada yapılır. **Faz IV** çalışmalar ise ilacın ruhsatı alınıp kullanıma konulmasından sonra yapılan klinik araştırmalardır.

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 (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 önenebilir.

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 **sistemik gruplandırmalarda** uygulamayı yapan hekimin taraf tutma riski olabilir. Bunun yerine kapalı bir yöntemle gruplara atama (concealed allocation) tercih edilmelidir. Sistemik gruplandırmadan mümkün olduğunca kaçınmalıyız.

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 ölçütümüzün 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 araştırmacının 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 ebeveyninden) **bilgilendirilmiş olur** (informed consent) alınmalıdır.

Araştırma protokolü

Araştırmaya başlamadan önce araştırmayı tüm yönleriyle anlatan ve açıklayan bir protokol hazırlanmalıdır. Burada araştırmacının 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.

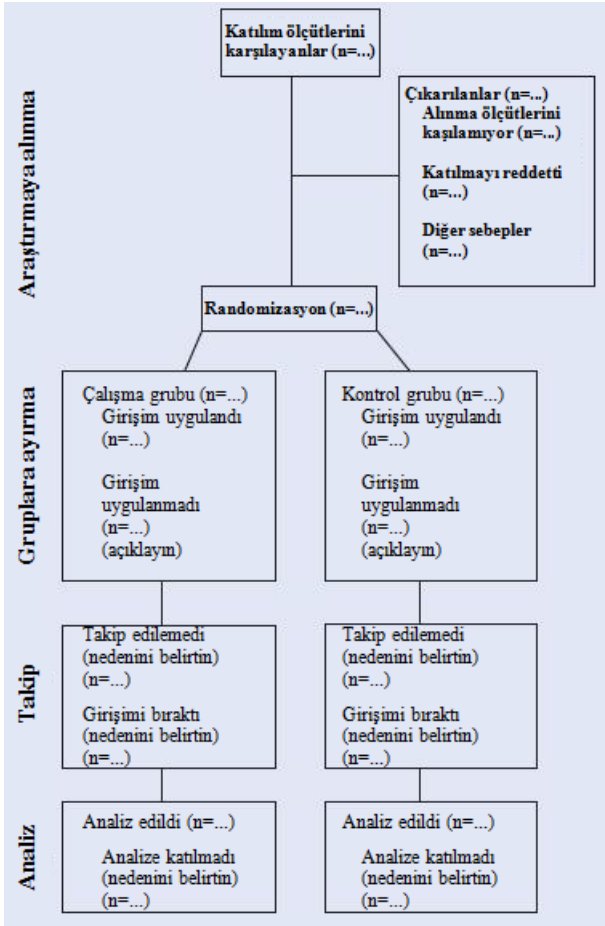
Gerçek bir tedavi etkisini ortaya çıkarabilmek için araştırmacının örneklem sayısının yeterli olması çok önemlidir. Örneklem sayısı ve nasıl hesaplandığı da protokolda belirtilmelidir.

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.

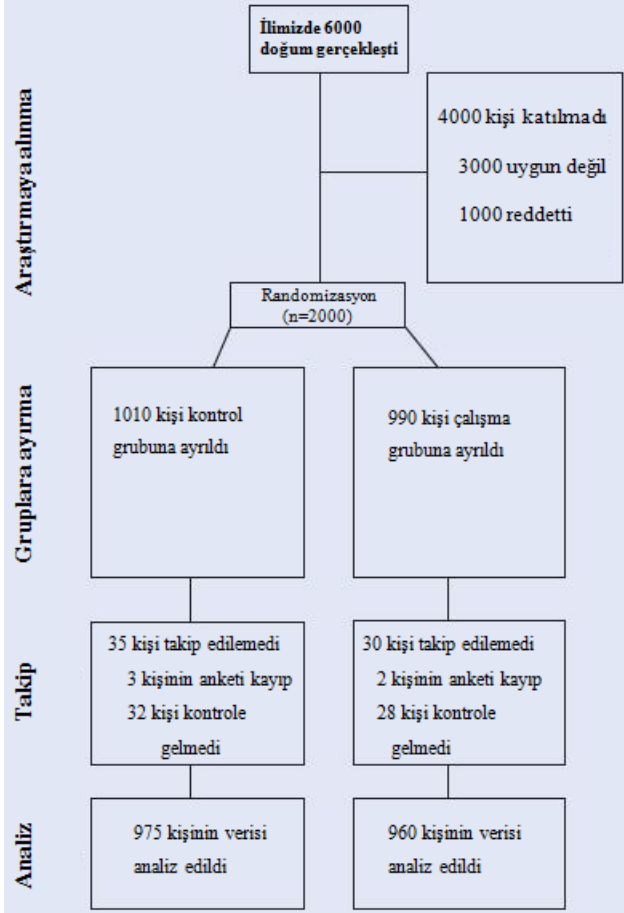
Tablo: CONSORT bildirisini (Consolidation of Standards for Reporting Trials www.consort-statement.org) kontrol listesi.

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 – ayırma gizliliği	9	Gruplar kesinleşene kadar ayırma işleminin gizli tutulması
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.



17. 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.	<ul style="list-style-type: none">• Bir araştırmada erkeklerle kadınlar arasındaki antihipertansif tedaviye uyum durumu karşılaştırılmak isteniyor.• Hangi istatistik hipotez test/testler yapılabilir?
2.	<ul style="list-style-type: none">• 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.	<ul style="list-style-type: none">• Bir araştırmada çeşitli sınıflardaki tıp öğrencilerinde girdikleri son sınavda geçer not alıp almadıkları karşılaştırılmak isteniyor.• Hangi istatistik hipotez test/testler yapılabilir?
4.	<ul style="list-style-type: none">• Bir araştırmacı mahallesindeki erişkinlerin sistolik kan basınçlarını ölçmek ve toplum ortalamasıyla karşılaştırmak istiyor.• Hangi istatistik hipotez test/testleri yapılabilir?
5.	<ul style="list-style-type: none">• Üç ayrı antihipertansif ilaç kullanan hasta grubunda kandaki pentraksin düzeylerini karşılaştırmak istiyorsunuz.• Hangi istatistik hipotez test/testleri yapılabilir?
6.	<ul style="list-style-type: none">• Veri çeşitlerini sınıflandırınız
7.	<ul style="list-style-type: none">• Aşağıdaki terimleri örnek vererek açıklayınız:<ul style="list-style-type: none">• Yüzde• Oran• Orantı• Hız• Puan
8.	<ul style="list-style-type: none">• Değişken 1: Hasta memnuniyeti:<ul style="list-style-type: none">1-hiç memnun değil,2-az memnun,3-orta halli,4-memnun,5-çok memnun• ne tür bir değişkendir?

9.	<ul style="list-style-type: none">Değişken 2: Öğrenim durumu: 1-ilkokul, 2-orta okul, 3-lise, 4-üniversite, 5-doktorane tür bir değişkendir?
10.	<ul style="list-style-type: none">8 ve 9. sorulardaki değişkenlerin ortalama ve standart sapmaları hesaplanabilir mi?Neden?
11.	<ul style="list-style-type: none">SPSS hangi yazılım veya standartlardan veri alabilir (import) ve hangi formatlarda veri ihraç edebilir (export) ?
12.	<ul style="list-style-type: none"><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.	<ul style="list-style-type: none">Değişkenimizde birden fazla seçeneğin olması durumunda SPSS’te hangi yöntemler izlenebilir?<ul style="list-style-type: none">Hatırlatma: Az sayıda seçenek vardır ve bunlardan çoğunun seçilmesi söz konusu olabilir Çok sayıda seçenek vardır ama bunlardan sadece birkaçının seçilmesi söz konusudur
14.	<ul style="list-style-type: none">Aşağıdaki anket sorusu hakkında ne söyleyebilirsiniz?Kaç yaşındasınız?<ol style="list-style-type: none">20’ den az,20-40,41-6060’ dan fazla
15.	<ul style="list-style-type: none">Araştırmadaki bireylere numara verilmesi ve SPSS’te de bu numaralar için bir değişken tanımlanması neden önemlidir?
16.	<ul style="list-style-type: none">Veri analizine başlamadan hata ayıklaması yapmak neden önemlidir?SPSS’te hata ayıklamasını hangi yöntemlerle yapabiliriz?
17.	<ul style="list-style-type: none">SPSS’te uç değerler ve aşırı değerleri nasıl saptarsınız?

18.	<ul style="list-style-type: none"> SPSS ile diyabet.sav veri setinde “duration” değişkeninin histogram grafiğini çiziniz ve yorumlayınız. 										
19.	<ul style="list-style-type: none"> Aşağıdaki ne tür bir grafikdir. SPSS ile diyabet.sav veri setinde bu grafiği yapınız? <div style="text-align: center;"> <p>Marital status</p> <ul style="list-style-type: none"> Single Married Divorced Widow <p>Pieshow counts</p> <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> </div>	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.	<ul style="list-style-type: none"> Merkezi dağılım ölçütleri nelerdir? 										
21.	<ul style="list-style-type: none"> Geometrik ortalama nedir? Antilog nasıl hesaplanır? 										
22.	<ul style="list-style-type: none"> Ağırlıklı ortalamayı açıklayınız? 										
23.	<ul style="list-style-type: none"> Yaygınlık ölçütleri nelerdir? 										
24.	<ul style="list-style-type: none"> Referans aralığı nedir? Nasıl hesaplanır? 										
25.	<ul style="list-style-type: none"> Standart sapmanın formülünü belirtiniz 										
26.	<ul style="list-style-type: none"> 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.	<ul style="list-style-type: none"> Normal dağılımı (Gauss dağılımı) tanımlayan parametreler hangileridir? 										
28.	<ul style="list-style-type: none"> Normal dağılımda verilerin <ul style="list-style-type: none"> – %68'i – %95'i 										

	<p>– %99'u</p> <ul style="list-style-type: none"> • ortalama \pm kaç standart sapma aralığındadır?
29.	<ul style="list-style-type: none"> • Normal Sapma (Standardized Normal Deviate) formülünü belirtiniz (z formülü)
30.	<ul style="list-style-type: none"> • Bir durum için sadece başarı ve başarısızlık şeklinde iki sonuç olabileceğini varsayalım. Bu veri hangi teorik dağılıma uyar?
31.	<ul style="list-style-type: none"> • Sağa doğru eğimli (pozitif) verileri normal dağılıma yaklaştırmak için hangi veri dönüştürme yöntemleri kullanılabilir?
32.	<ul style="list-style-type: none"> • Ortalamanın standart hatası (SEM) nasıl hesaplanır? • SEM bize ne hakkında bilgi verir?
33.	<ul style="list-style-type: none"> • Bir kategorik değişken için ortalamanın standart hatası nasıl hesaplanır?
34.	<ul style="list-style-type: none"> • Ortalamanın güven aralığı nasıl hesaplanır?
35.	<ul style="list-style-type: none"> • Kategorik veriler için ortalamanın güven aralığı nasıl hesaplanır?
36.	<ul style="list-style-type: none"> • Numerik veriler için ortalamalar arasındaki farkın güven aralığı nasıl hesaplanır?
37.	<ul style="list-style-type: none"> • Kategorik veriler için ortalamalar arasındaki farkın güven aralığı nasıl hesaplanır?
38.	<ul style="list-style-type: none"> • Ortalamalar arasındaki farkın güven aralığı nasıl yorumlanır?
39.	<ul style="list-style-type: none"> • Odds oranı nedir? • Odds oranı için güven aralığı nasıl yorumlanır?
40.	<ul style="list-style-type: none"> • Klinik olarak cevaplanabilir bir soru sormak için kullanılan PICO akronimini örnek vererek açıklayınız?
41.	<ul style="list-style-type: none"> • Araştırmamızda nedenselliği irdelemek için Hill tarafından önerilen kriterler nelerdir?
42.	<ul style="list-style-type: none"> • Araştırmalarda ne tür hatalar olabilir?
43.	<ul style="list-style-type: none"> • Confounding (karıştırıcılık) nedir? Örnek veriniz?
44.	<ul style="list-style-type: none"> • Temel araştırma tasarımlarını sayınız.
45.	<ul style="list-style-type: none"> • Varyasyonu azaltmak için ne tür araştırma tasarımları yapabiliriz?
46.	<ul style="list-style-type: none"> • Çapraz araştırma tasarımını açıklayınız.
47.	<ul style="list-style-type: none"> • Faktöriyel araştırma tasarımını açıklayınız.
48.	<ul style="list-style-type: none"> • Intention to treat analizini açıklayınız? Bu yöntemin aksi nedir?
49.	<ul style="list-style-type: none"> • CONSORT bildirisi nedir?
50.	<ul style="list-style-type: none"> • Consealed allocation (ayırma gizliliği) nedir? • Ayırma gizliliği neden önemlidir?

Alıştırma 7

Aşağıdaki bağlantıdaki makaleyi indiriniz ve sorulara cevap veriniz:

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

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 verilip 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. 293

18. Kohort Araştırmaları

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 yapılaş 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 açmayacağına araştırılmasıdır. 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, araştırmacının 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 araştırmacının değerini düşürecektir. Bu durum kohort araştırmalarına özgü olmayıp genel bir sorun oluşturmaktadır. Yine de konuya burada değinmeyi uygun gördük.

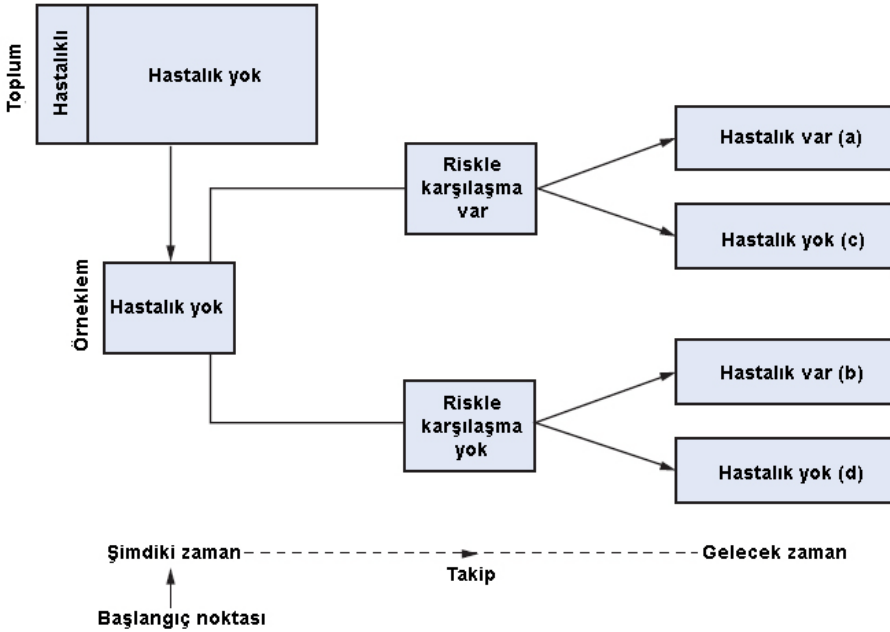
Örnek senaryo 1: 100 hastanın koroner kalp hastalığı nedeniyle medikal (50 kişi) ve cerrahi (50 kişi) tedavi almak üzere randomize edildiğini düşünelim. Sonuç ölçütümüz girişimden 1 ay sonraki efor kapasitesi olsun. Her iki gruptan da 5'er kişinin uygulama öncesinde öldüğünü düşünelim. Ölenlerin muhtemelen

durumları daha kötüydü ve cerrahi uygulansaydı ameliyat sonrasında öleceklerdi. Ameliyat öncesinde öldükleri için araştırmanın analizine dahil edilmemeleri halinde cerrahi grubunun başarısı oransal olarak daha iyi gibi gözükcektir.

Örnek senaryo 2: 100 hipertansif hastanın A ilacı (50 kişi) ve B ilacı (50 kişi) almak üzere randomize edildiğini düşünelim. Sonuç ölçtüğümüz 1 ay sonraki kan basınçları olsun. Sonuçta B ilacının istatistiksel olarak daha üstün olduğunu bulduğumuzu varsayalım. A grubunda 5, B grubunda 10 kişinin araştırmayı terk ettiğini düşünelim. Bu durumda B ilacının kullanım zorluğu, yan etkileri veya etkisiz olması gibi nedenlerle bırakıldığı spekülasyonu yapılabilir ve “Eğer araştırmadan ayrılanlar olmasaydı, B ilacı daha üstün çıkmayacaktı” denilebilir.

Yukarıdaki örneklerde verilen nedenlerle araştırmalarda vaka kayıplarının en az olması hedeflenmelidir. Bu konuda katılımcılarla iletişimi devam ettirmek, takiplere gelenleri ödüllendirmek gibi çözümler üretilebilir. Kontrole gelmeyen bireyler için sonuç ölçütünün son değerini kullanmak son çare olmakla birlikte, bu uygulama tartışmalıdır (http://www.lshtm.ac.uk/msu/missingdata/papers/psi20_oct_2004_print.pdf).

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 değerlendirilmeli ve risk durumu incelenmelidir.



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

Kohort araştırmaların analizi**Tablo:** Kohort araştırmasında gözlenen frekanslar.

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

“ $a/(a+b) - c/(c+d)$ ” 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 = $a+b/n$

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

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

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

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

$$\mathbf{RR = risk_+ / risk_- = [a/(a+c)] / [b/(b+d)]}$$

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 ölçütümüz 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 etyolojisinin 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 bazılarında oral antidiyabetik ilaçlar, bazılarında 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 karıştırmak 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 karıştırmaya yol açacaktır.

Örnek 1: risk azalması			Örnek 2: risk artışı		
	Risk pozitif (R)	Kontrol (K)	Toplam (R)	(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, veya %10	KOY = KH / KT = 0,4, veya %40		ROY = 0.5 (%50)	KOY = 0.4 (%40)
Eşitlik	Değişken	Kısaltma	Örnek 1	Örnek	
ROY – KOY	< 0: toplam risk azalması (absolute risk reduction)	TRArt (ARR)	(-)0,3, veya (-) %30		
	> 0: toplam risk artışı (absolute risk increase)	TRAz (ARI)		0,1, veya %10	
(ROY – KOY) / KOY	< 0: rölatif risk azalması (relative risk reduction)	RRAz (RR□)	(-)0,75, veya (-) %75		
	> 0: rölatif risk artışı (relative risk increase)	RRAttr (RRI)		0,25, veya %25	
1 / (ROY – KOY)	< 0: tedavi için gereken kişi (number needed to treat)	TGK (NNT)	(-)3,33		

	> 0: zarar için gereken kişi (number needed to harm)	ZGK (NNH)		10
ROY / KOY	Rölatif risk (relative risk)	RR (RR)	0,25	1,25
(RH / RS) / (KH / KS)	Odds oranı (odds ratio)	OR (OR)	0,167	1,5
$\frac{RH}{RH + KH} - \frac{RS}{RS + KS}$	Atfedilen risk (attributable risk)	AR (AR)	(-)0,34, veya (-) %34	0,095, veya %9,5
$(RR - 1) / RR$	Atfedilen risk yüzdesi (attributable risk percent)	ARY (ARP)		%20
$1 - RR$ (veya $1 - OR$)	Koruyucu kısım (preventive fraction)	KK (PF)	0,75, veya %75	

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

19. Vaka Kontrol Çalışmaları

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.

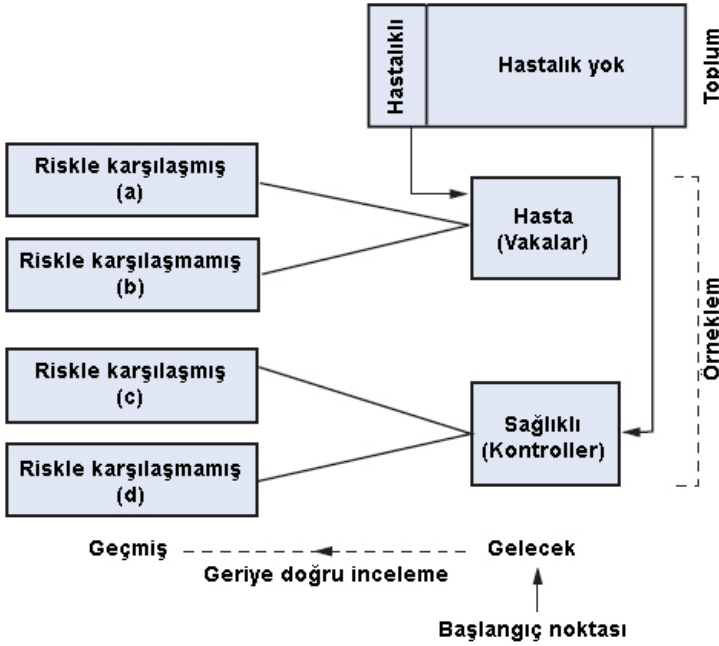
Araştırmamıza mümkün olduğunca fazla vaka almalıyız. Böylece yorumlarımızı genellememiz mümkün olur. Gerekirse hastane kayıtları incelenerek araştırma dönemindeki vakaların tümüne ulaşılmaya çalışılmalıdır.

Aralık 2008’de Edirne’de miyokard enfarktüsünden ölen bireylerde sigara içme, diyabet varlığı ve hiperlipidemi gibi risk faktörlerini araştırdığımızı düşünelim. Sadece fakülte hastanesi kayıtlarını inceleyip enfarktüstüsten 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 ortamdan 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** hesaplanır (<http://www.merriam-webster.com/dictionary/odd>: differing markedly from the usual or ordinary or accepted – normalden, alışılmıştan veya beklenenden farklı olan).

Odds = riskle karşılaşanlarda hasta olma oranı / riskle karşılaşmayanlarda hasta olma oranı

Burada “riskle karşılaşanlarda hasta olma oranı” = “riskle karşılaşanlarda hasta olma oranı” / “riskle karşılaşanlarda hasta olmama oranı” şeklinde hesaplanır.

Tablodaki frekanslar üzerinden yapacak olursak:

$$\text{Odds}_{\text{karşılaşan}} = [a/(a+c) / c/(a+c)] = a/c$$

$$\text{Odds}_{\text{karşılaşmayan}} = [b/(b+d) / d/(b+d)] = b/d$$

Bu durumda

$$\text{odds oranı} = (a/c) / (b/d) \text{ yani } ad/bc$$

olur.

Tablo: Vaka kontrol araştırmasında gözlenen frekanslar.

		Risk faktörü		Toplam
		Var	Yok	
Hastalık	Vaka	a	b	a+b
	Kontrol	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:

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

Aktürk Z, Acemoğlu H. Sağlık Çalışanları İçin Araştırma ve Pratik İstatistik

$$RR = (25/120) / (75/1880) = 5,2 \quad OR = (25/95) / (75/1805) = 6,3$$

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

		Risk faktörü		Toplam
		Var	Yok	
Hastalık	Vaka	2,5	7,5	10
	Kontrol	117,5	1872,5	1900
	Toplam	120	1880	2000

$$RR = (2,5/120) / (7,5/1880) = 5,2 \quad OR = (2,5/117,5) / (7,5/1872,5) = 5,3$$

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.

Aliřtırma 8

Arařtırma 1: Bir kohort alıřmanın sonuları.

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

Arařtırma 2: Bir vaka kontrol alıřmanın sonuları

		Risk		Toplam
		Var	Yok	
Hastalık	Var	600	400	1000
	Yok	250	1750	2000
Toplam		650	2350	3000

Sorular

1. Arařtırma 1 iin rlatif (nisbi) riski hesaplayınız.
2. Arařtırma 2 iin odds oranını hesaplayınız
3. Arařtırma trleri belirtilmemiř olsaydı Arařtırma 1'in kohort, Arařtırma 2'nin ise vaka kontrol alıřması olduėunu nasıl ayırt ederdiniz?
4. Arařtırma 2 iin odds oranının %95 gven aralıėı **[8,73 – 12,60]** olarak hesaplanmıřtır. Bunu yorumlayınız. Gven aralıėı **[0,02-12,60]** olsaydı yorumunuz nasıl olurdu?

zm iin bakınız s. 294

20. Hipotez Testleri

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

Aktürk Z, Acemoğlu H. Sağlık Çalışanları İçin Araştırma ve Pratik İstatistik 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” numerik 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:

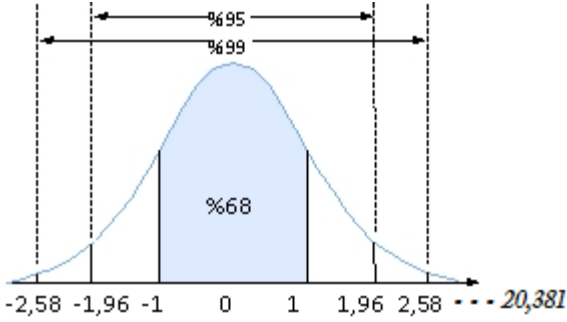
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Height	Equal variances assumed	.116	.733	20,381	404	.000	13,1159	.6435	11,8508	14,3810
	Equal variances not assumed			20,342	394,130	.000	13,1159	.6448	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 şaşırılmamalıdır.

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 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) da denir. Bu testler verilerin gerçek değerleri yerine veri setindeki büyüklük sıralarını dikkate alarak hesap yapar. Bu nedenle de olasılık dağılımlarıyla ilişkisi yoktur.

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

sađlar. Ayrıca p deđerini bilmesek bile gven aralıđına bakarak farkın anlamlı olup olmadıđı hakkında yorum yapabiliriz. rneđin numerik bir deđiřkeni iki grup arasında karřılařtırdıđımızda ortalamalar arasındaki farkın %95 gven aralıđına baktıđımızı varsayalım:

Erkeklerle kadınların sistolik kan basınları arasındaki farkın ortalaması 8,5 mmHg olarak bulunmuřtur. Aradaki fark istatistiksel olarak anlamlı çıkmıřtır ($p=0,003$). Aradaki farkın %95 gven aralıđı ise $[2,3 - 11,2]$ aralıđındadır.

%95 gven aralıđımızın sıfırını iermemesi durumunda iki ortalama arasındaki fark iin p deđerinin 0,05'ten kk olduđunu, yani anlamlı olduđunu syleyebiliriz ama tam p deđerini veremeyiz.

Gnmzde arařtırma makalelerinde p deđerisiyle birlikte genelde gven aralıđı da sunulmaktadır.

21. Hipotez Testlerinde Hatalar

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 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 numerik 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. Araştırmaların yöntem bölümünde güç hesabı ile birlikte etki genişliğinin de ne kadar alındığı belirtilmelidir.

Yanlıştırma

Kurduğumuz hipotezle ilgili karar verirken en doğru kararı vermek isteriz. Halbuki, toplumun tamamını değil de sadece bir örnekleme 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 çıkarırız. 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 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			
		H₀ reddedildi	H₀ kabul edildi
Gerçek	H₀ doğru	Tip I hata (%5)	Doğru karar (Araştırmanın gücü; Power) (%80)
	H₀ 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 te 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 4 faktör vardır:

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. **Effect of interest:** 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.

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 uygulanır. Olasılık kurallarından hatırlayacağımız gibi birden çok test yaptığımızda araştırmamızın hata ihtimali önemli ölçüde artacaktır. Örn:

Bir hipotez testi uyguladığımızda %5 yanlış yapımızın olduğunu varsayalım. Aynı araştırmada iki test uyguladığımızda biri VEYA ötekini 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:

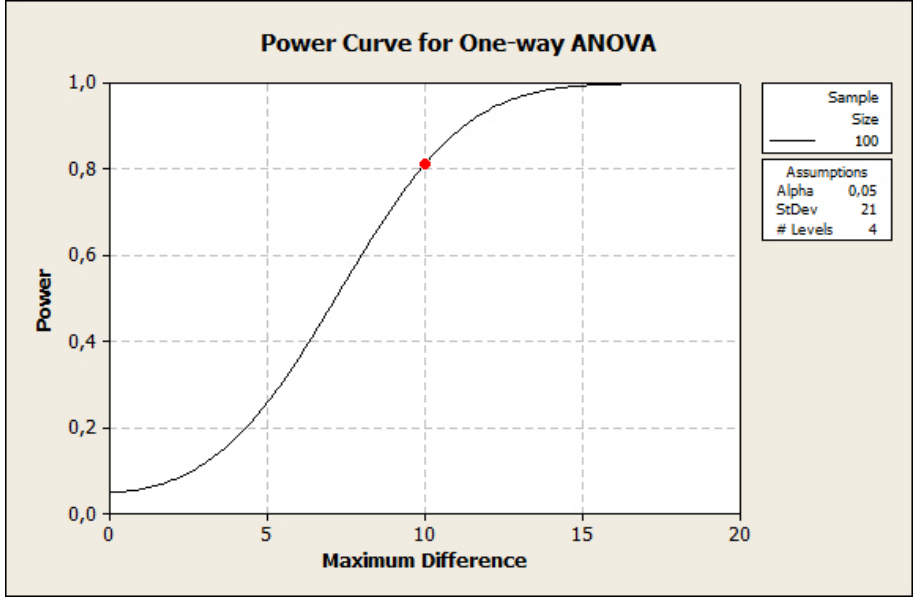
Araştırmamızda 3 adet hipotez testi uyguladığımızı düşünelim:

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.

Aşağıda MINITAB® 15 ile (<http://www.minitab.com/>) yapılan bir power hesabı örneği verilmiştir. Dikkat edilirse saptanmak istenen fark (maximum difference) azaldıkça araştırmamızın 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:

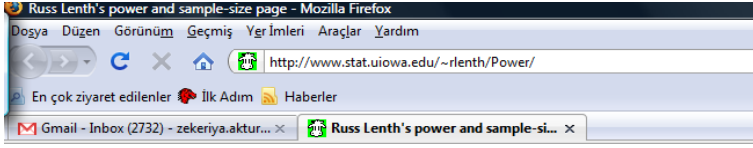
<http://www.stat.uiowa.edu/~rlenth/Power/>. Bu kez numerik veri iki grup karşılaştırması yapalım:

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.

Numerik veri – iki bağımsız grup).

Russ Lenth'in sitesinden "Two samples t-test'i" seçelim



Java applets for power

Select the analysis to be used in your study:

- CI for one proportion
- Test of one proportion
- Test comparing two proportions
- CI for one mean
- One-sample t test (or paired t)
- Two-sample t test (pooled or Satterthwaite)**
- Linear regression
- Balanced ANOVA (any model)
- Two variances (F test)
- R-square (multiple correlation)
- Generic chi-square test
- Generic Poisson test
- Online tables of common distributions
- Pilot study

This software is intended to be useful in planning sta

Each selection provides a graphical interface for stu
varying parameters, and a simple provision for graph

Each dialog window also offers a Help menu. **Pleas**

The "Balanced ANOVA" selection provides another i
model.

Note: The dialogs open in separate windows. If you'r
example, you'll have two "Help" menus there!

You may also [download](#) this software to run it on yo

Run selection'ı tıkladığımızda aşağıdaki pencere açılacaktır (bilgisayarımızda Java kurulu olmalı):

Two-sample t test (general case)

Options Help

sigma1 = 1

sigma2 = 1

Equal sigmas

n1 = 25

n2 = 25

Allocation Equal

Two-tailed Alpha .05

Equivalence

Degrees of freedom = 48

True difference of means = .5

Power = .4101

Solve for Sample size

Test of one proportion

Test comparing two proportions

CI for one mean

One-sample t test (or paired t)

Two-sample t test (pooled or Satterthwaite)

Linear regression

Balanced ANOVA (any model)

Two variances (F test)

R-square (multiple correlation)

Generic chi-square test

Generic Poisson test

Online tables of common distributions

Pilot study

Run selection

Each selection provides a graphical interface varying parameters, and a simple provision for

Each dialog window also offers a Help menu

The "Balanced ANOVA" selection provides a model.

Note: The dialogs open in separate windows. For example, you'll have two "Help" menus there

You may also [download](#) this software to run

Standart sapmayı (σ_1 ve σ_2) önceki örnekte olduğu gibi 21 mmHg alalım. Varyansları eşit kabul edelim (Equal sigmas). Örneklem sayısını her iki grup için 100 alalım ($n_1=100$, $n_2=100$). İki uçlu alfa 0.05 kalsın. Ortalamalar arasında saptamak istediğimiz fark (True difference of means) 10 mmHg olsun. Power %91,79 olarak çıkacaktır:

Two-sample t test (general case)

Options Help

signal
Value 21 OK

sigma2 = 21
0 5 10 15 20

Equal sigmas

n1
Value 100 OK

n2
Value 100 OK

Allocation Equal

Two-tailed Alpha .05

Equivalence

Degrees of freedom = 198

True difference of means
Value 10 OK

Power = .9179
0 2 4 6 8 10 1

Solve for Sample size

Test of one proportion
Test comparing two proportions
CI for one mean
One-sample t test (or paired t)
Two-sample t test (pooled or Satterthwaite)
Linear regression
Balanced ANOVA (any model)
Two variances (F test)
R-square (multiple correlation)
Generic chi-square test
Generic Poisson test
Online tables of common distributions
Pilot study

Run selection

Each selection provides a graphical interface varying parameters, and a simple provision

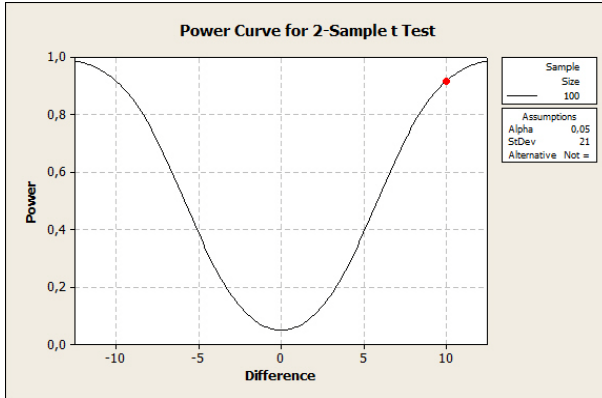
Each dialog window also offers a Help menu

The "Balanced ANOVA" selection provides model.

Note: The dialogs open in separate windows; example, you'll have two "Help" menus there

You may also [download](#) this software to run

Aynı alıřtırmayı MiniTab® ile yapsak ařađıdaki sonucu elde ederiz:



Alıştırma 9

Hipotez testlerinde yaptığımız hataları yargıda yapılan hatalara benzetelim. Bir şüpheli hakimin karşısına çıkarıldığında iki ihtimal vardır: ya beraat eder, ya da ceza alır. Her iki durumda da hukuk sisteminin yanılma olasılığı vardır.

<i>Avukatın hipotezi (H_0):</i> “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 Hata payı %5	Doğru karar Hukuğun gücü (%80)
	H_0 yanlış (red): Suçlu	Doğru karar (%95)	Tip 2 hata Hata payı %20

Dikkat ederseniz yargının gücü suçluya suçlu demesinde değildir; esas güç, suçsuz 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. 296

22. Nümerik Veri Tek Grup

Amaç: Bu konu sonunda okuyucunun tek gruptan elde edilen nümerik verilerin toplum ortalamasıyla karşılaştırılmasında kullanılan tek örnekleme 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 örnekleme t -testi varsayımlarını açıklayabilmek
- SPSS ile tek örnekleme t -testi yapabilmek
- İşaret testinin kullanım yerini tartışabilmek
- İşaret testinin nasıl yapılacağını açıklayabilmek

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

*Bu karşılaştırmayı **tek örnekleme t -testi** (one-sample t -test) ile yapabiliriz. Tek örnekleme t -testini yapabilmemiz için verilerimizin bazı varsayımları karşılaması gerekir. Bu varsayımları karşılamamız halinde alternatifimiz **nonparametrik bir test olan işaret testi** olabilir.*

Tek örnekleme t -testi varsayımları

Tek örnekleme t -testinin varsayımları;

- Değişkenimizin toplumda normal dağılması ve
- Yeterli bir örnekleme sahip olmamızdır.

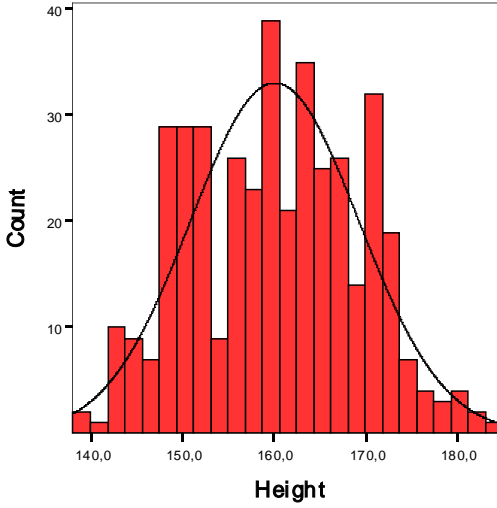
Toplumun varyansını (σ) bilmememiz durumunda verilerimizin dağılımı teorik dağılımlardan t dağılımına uyacaktır. Eğer toplumun varyansını bilirsek veya örnekleme sayımız çok yüksekse bu durumda normal dağılım bilgilerine dayanan z testini kullanabiliriz. Ancak, her iki durumda da sonuçlar aynı çıkacaktır.

Örnek uygulama

Diyabet.sav verisetini kullanarak bir uygulama yapalım:

Öncelikle tek örnekleme 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 (H₀) ve alternatif hipotezin (H₁) tanımlanması:

H₀: İlimizdeki bireylerin boy ortalaması ile mahallemizdeki bireylerin boy ortalaması arasında fark yoktur.

H₁: 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 örneklemde *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
s = örneklemimizin standart sapması	9,2
n = örnek sayısı	406

Buradan $t = -3,285$ olarak hesaplanır.

4. Test istatistiğinden elde edilen değer bilinen bir olasılık dağılımı ile karşılaştırılması:

Sonucumuzu t tablosu ile (Appendix A2) karşılaştıracak olursak $t = 3,285$ ve serbestlik derecesi = 405 için p değerinin 0,01'den küçük olduğunu görürüz.

df	Two-tailed P-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
$SD = 405$		$t = 3,285$		
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 örnekleme 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 örnekleme 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 örnekleme veri topladığımızı düşünelim:

1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması

H_0 : İlimizdeki bireylerin boy ortalaması ile örnekleminizdeki bireylerin boy ortalaması arasında fark yoktur.

H_1 : İlimizdeki bireylerin boy ortalaması ile örnekleminizdeki 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:

$$z = \left[\left| 7 - (27/2) \right| - (1/2) \right] / \left[\sqrt{(27/2)} \right] = \mathbf{1,632}$$

4. Test istatistiğinden elde edilen değerlerin bilinen bir olasılık dağılımı ile karşılaştırılması

r değerini Appendix A6'dan, z değerini ise Appendix A1'den bakarak p 'yi bulabiliriz.

z	2-tailed P -value
0.0	1.000
0.1	0.920

1.4	0.162
1.5	0.134
1.6	0.110

$z = 1,6$ için iki yönlü p değerinin **0,110** olduğunu görüyoruz.

5. *P deęerinin ve sonuların yorumlanması*

P deęeri 0,05'ten byk olduęundan sıfır hipotezini kabul etmemiz gerekir.

Dolayısıyla rnekleminizle toplum ortalaması arasında boy aısından istatistiksel olarak anlamlı bir fark yoktur.

Alıştırma 10

Bu alıřtırmada daha önceki konuda işlediğimiz araştırmanın gücü (power) ve buna etki eden faktörleri tekrar gözden geçirmek istiyoruz.

Aşğıdaki 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=1bxrilhxrip9s.victoria>

Aşğıdaki 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 örnekleme 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. 297

23. Nümerik Veri İki Bağımlı Grup

Amaç: Bu konu sonunda okuyucunun iki bağımlı gruptan elde edilen nümerik verilerin ortalamalarının karşılaştırılmasında kullanılan bağımlı örneklerde 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 örneklerde t -testi varsayımlarını açıklayabilmek
- SPSS ile bağımlı örneklerde 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 birbirinden 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 ötekini 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 örneklerde t -testi veya Wilcoxon işaretli sıra testi olacağını görürüz.

Bağımlı Örneklerde t -testi

Varsayımları

- Değişken, toplumda normal dağılımalı
- 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 örnekleme t-testine indirgenebilir. Değişkenimiz, farkların ortalaması, toplum ortalaması ise sıfırdır.

Ek açıklama

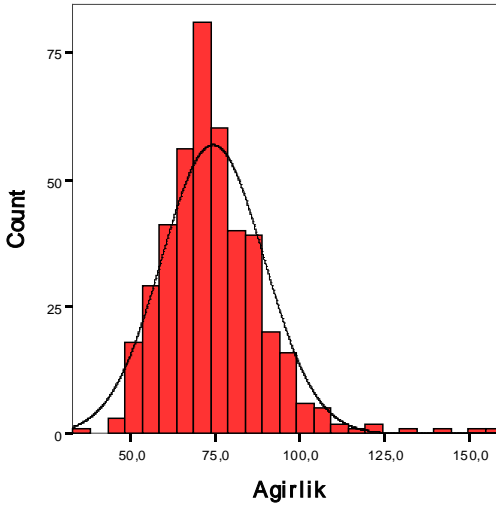
Veriler bağımlı (eşleştirilmiş) olduğundan her iki ölçümde de örneklem sayısı eşit olmalıdır. n adet bireydeki iki ölçüm arasında n adet fark olacaktır. Bunların aritmetik ortalaması \bar{x} ve tahmin edilen standart sapması s_d 'dir.

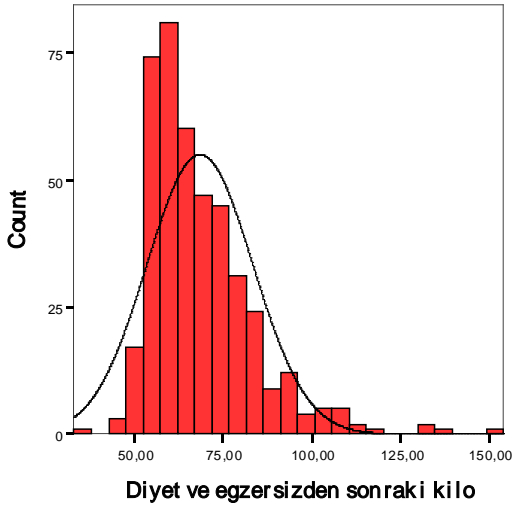
Örnek uygulama

Diyabet.sav verisetini kullanarak bir uygulama yapalım:

Öncelikle verilerimizin bağımlı örneklemlerde t-testi varsayımlarını (şartlarını) karşılayıp karşılamadığına bakmamız gerekecek:

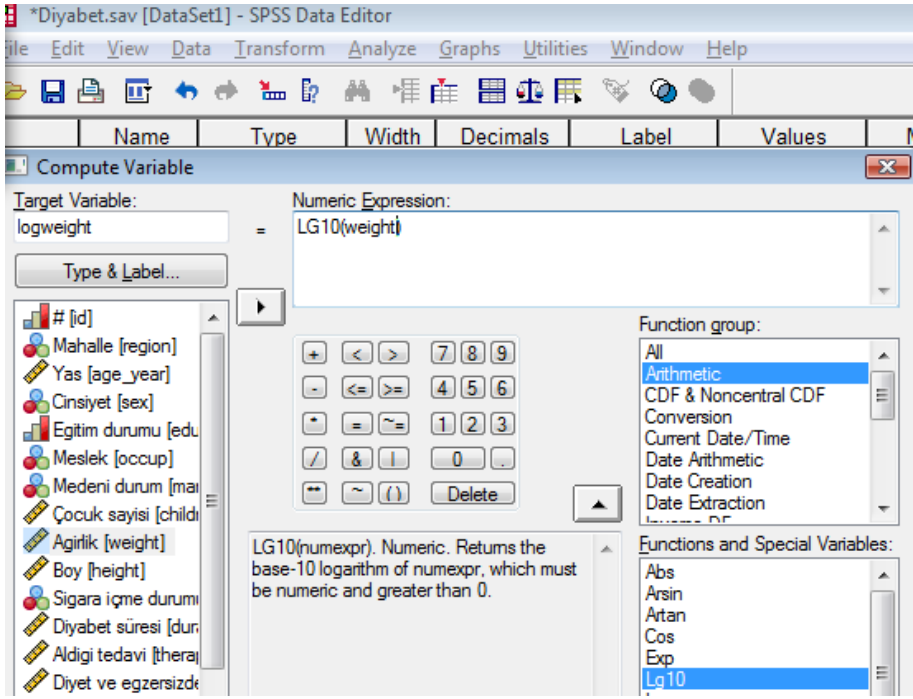
- Ağırlık (weight ve weight2) değişkenlerimizin histogram grafiklerime baktığımızda normal dağılımdan hafif bir sapma olduğunu görüyoruz:





Bu durumda sağa eğimli verilerimizi normal dağılıma yaklaştırmak için logaritmik dönüşüm yapmayı deneyebiliriz. Weight ve weight2 değişkenlerimize logaritmik dönüşüm uyguladıktan sonra

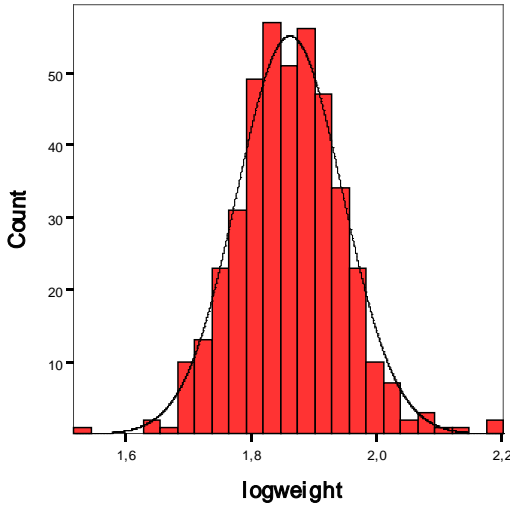
Transform>Compute variable>[“Target Variable” alanına “logweight”, “Numeric expression” alanına “LG10(weight)” girelim>ok.

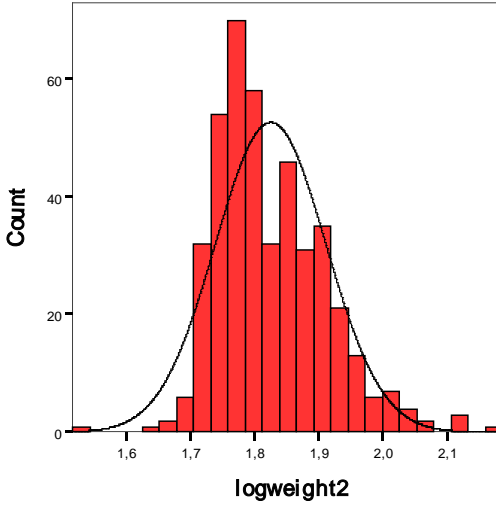


Aynı işlemi weight2 için tekrarlayalım:

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 örnekte t testine benzerdir:

$$t = \frac{(\bar{d} - 0)}{SE(\bar{d})} = \frac{\bar{d}}{s_d/\sqrt{n}}$$

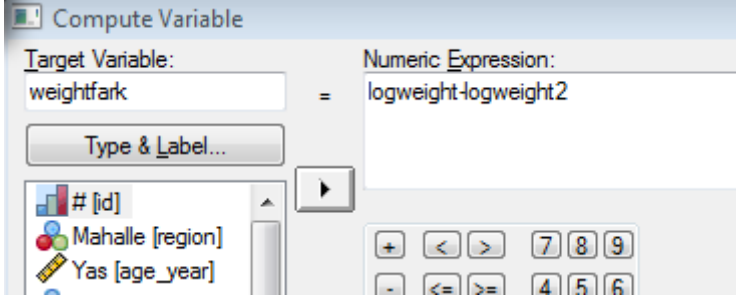
\bar{d} = öncesi ve sonrasında yapılan ölçümler arasındaki farkın aritmetik ortalaması	0,0372
0 = toplum ortalaması (aradaki farkın sıfır olduğunu varsayıyoruz)	0
s_d = öncesi ve sonrasında yapılan ölçümler arasındaki farkın standart	0,0614

sapması	
$n =$ 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 değerlerin bilinen bir olasılık dağılımı ile karşılaştırılması:

Sonuçlarımız (n-1) serbestlik derecesinde t dağılımı gösterir. t tablosu ile (Appendix A2) karşılaştıracak olursak $t = 12,475$ ve serbestlik derecesi = 423 için p değerinin 0,01’den küçük olduğunu görürüz.

df	Two-tailed P-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
SD = 423				
t = 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 P-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 yapılacak olursak:

Aktürk Z, Acemoğlu H. Sağlık Çalışanları İçin Araştırma ve Pratik İstatistik
*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) örneklekimizden 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 örneklemdaki *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 örneklemden 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
25	72,0	73
26	73,0	62,05
27	65	68
28	76,0	64,60
29	94,0	79,90
30	91,5	77,78
31	61,5	52,28
32	54	65
33	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
25	72,0	73	-1
26	73,0	62,05	10,95
27	65	68	-3
28	76,0	64,60	11,4
29	94,0	79,90	14,1
30	91,5	77,78	13,72
31	61,5	52,28	9,22
32	54	65	-11
33	88,0	74,80	13,2

Şimdi hesapladığımız farkları sıraya koyalım (sıfır olan farkları atacağız):

Öncesi-Sonrası	Sıra no
-0,5	1
-1	3
-1	3
-1	3
-3	4,5
9,22	4,5
9,75	6,5
9,75	6,5
10	8
10,51	9
-10,8	10

-11	11
10,95	12,5
10,95	12,5
11	14
11,17	15
11,32	16
11,4	17
12	18,5
12	18,5
12,29	20
12,75	21
12,9	22
13,2	23
13,72	24
14,1	25
14,25	26
14,35	27
15	28 = n'

Şimdi pozitif (T₊) ve negatif (T₋) farkların sıra değerlerini toplayalım:

T ₊ =	5+6,5+6,5+8+9+12,5+12,5+14+15+16+17+18,5+18,5+20+21+22+23+24+25+26+27+28	=3
T ₋ =	1+3+3+3+4+10+11	=

n' ≤ 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' = 28 değeri 10'dan büyük olduğuna göre z formülünü uygulayalım:

$$z = \left[\left| 35 - \frac{(28 \times 29)}{4} \right| - (1/2) \right] / \left[\frac{(28 \times 29) \times (56 + 1)}{24} \right] = \mathbf{0,086}$$

4. Test istatistiğinden elde edilen değer bilinen bir olasılık dağılımı ile karşılaştırılması

$n' \leq 25$ olması durumunda T testi değeri Appendix A8'den karşılaştırılır

$n' > 25$ olması durumunda z değeri Appendix A1'den bakılır.

z	2-tailed P-value
0.0	1.000
0.1	0.920
1.4	0.162
1.5	0.134
1.6	0.110

z tablosundan $z = 0,08$ için iki yönlü p değerinin **1,000** civarında olduğunu görüyoruz.

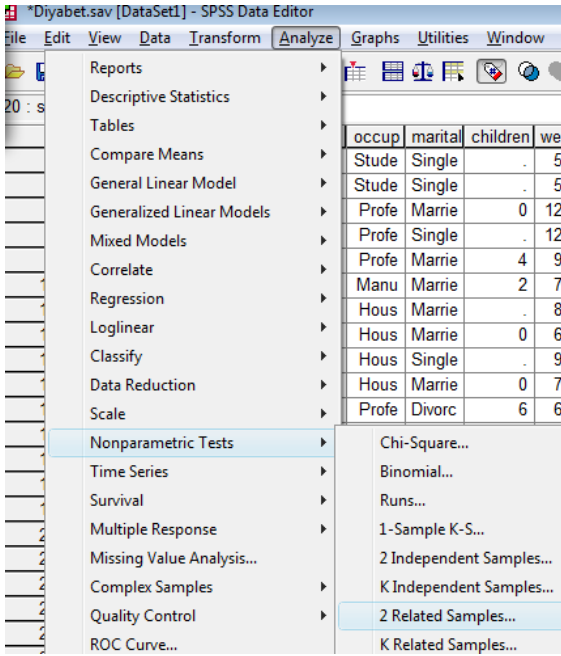
5. P değerinin ve sonuçların yorumlanması

P değeri 0,05'ten büyük olduğundan sıfır hipotezini kabul etmemiz gerekir.

Dolayısıyla tedavi öncesi ve sonrasındaki kilo farkları arasında istatistiksel olarak anlamlı bir fark yoktur.

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

24. Nümerik Veri İki Bağımsız Grup

Amaç: Bu konu sonunda okuyucunun iki bağımsız gruptan elde edilen nümerik verilerin ortalamalarının karşılaştırılmasında kullanılan bağımsız örneklerde 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 (H_0 : 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 uygulayamamız halinde nonparametrik alternatifi olan Wilcoxon sıra toplamı testi (Wilcoxon rank sum test) veya Mann-Whitney U testini yapabiliriz.

Bağımsız örneklerde t testi

Varsayımları

- Değişken, toplumda normal dağılımalı
- 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

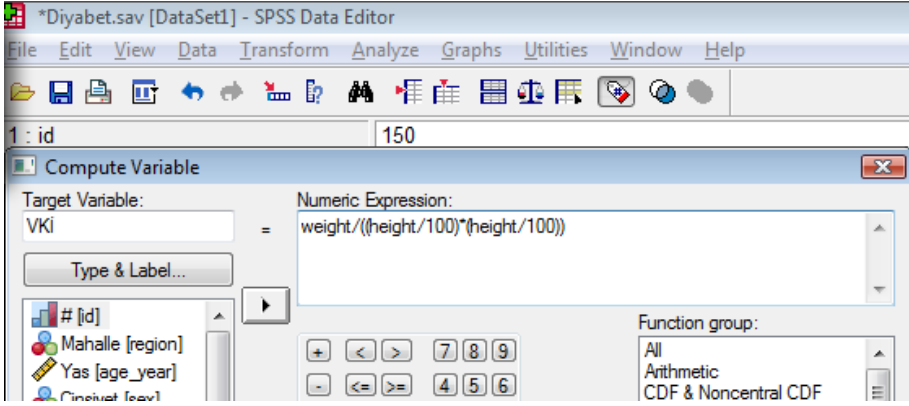
t testi daha önce gördüğümüz tek örneklemede t testi ve bağımlı örneklemlerde t testi gibi ortalamalar arasındaki farka dayanmaktadır. Bu testin verileri teorik dağılımlarından t dağılımına benzemektedir.

Not

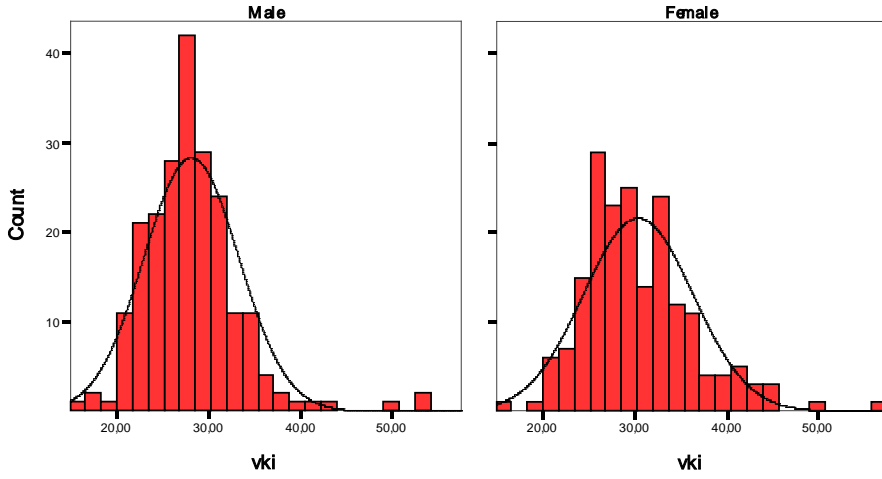
İki grubumuzun örneklemleri n_1 ve n_2 , ortalamaları \bar{x}_1 ve \bar{x}_2 , standart sapmaları ise s_1 ve s_2 'dir.

Diyabet.sav veri setimizde boy ve ağırlık değişkenleri olmakla birlikte VKİ hesaplanmamıştır. Aşağıdaki komutları kullanarak yeni bir VKİ değişkeni hesaplayabiliriz (VKİ = [ağırlık (kg)] / [boy (m)]² olduğunu hatırlayınız).

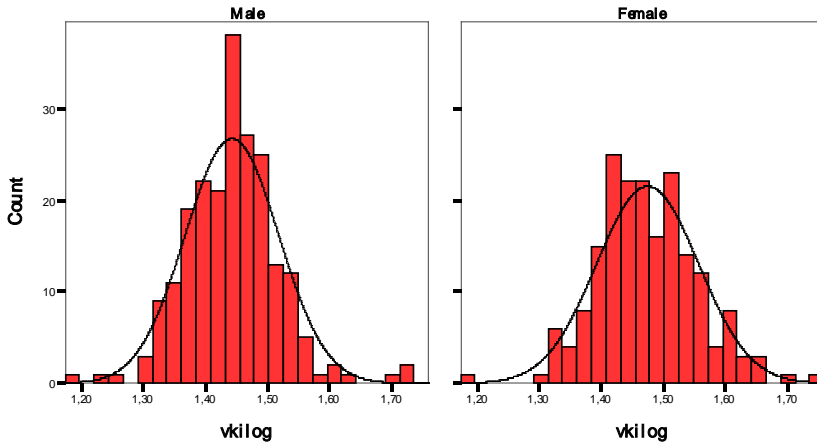
Transform>Compute variable>[“Target Variable” alanına “VKİ”, “Numeric expression” alanına “weight/((height/100)(height/100))” girelim>ok.*



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:



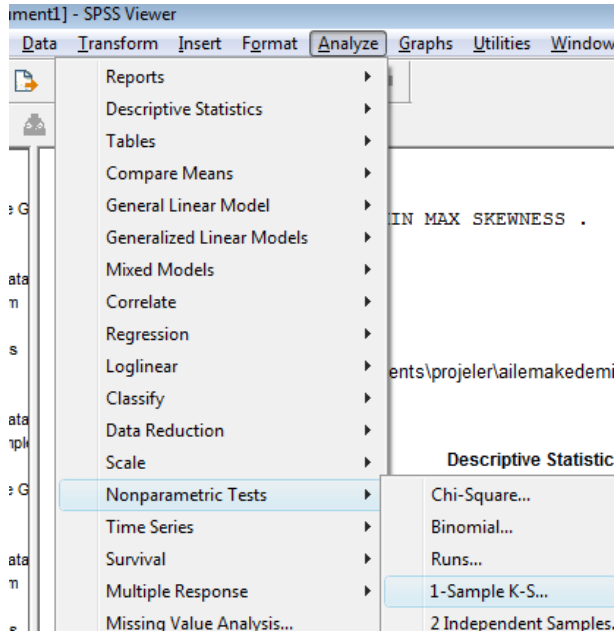
Yeni değişkenimizin ($\log VKI$) skewness değerine de baktığımızda normal dağılıma yaklaştığını görüyoruz:

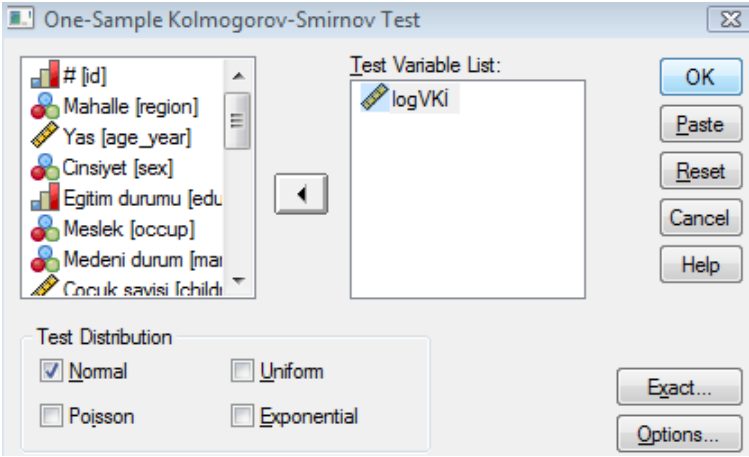
Statistics

vkilog			
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>[“logVKİ” 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
		Std. Deviation	,0755
	Most Extreme Differences	Absolute	,073
		Positive	,073
		Negative	-,040
	Kolmogorov-Smirnov Z		1,070
	Asymp. Sig. (2-tailed)		,202
Female	N		189
	Normal Parameters ^{a,b}	Mean	1,474
		Std. Deviation	,0824
	Most Extreme Differences	Absolute	,046
		Positive	,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, "logVKI" değişkenimizi kullanarak Student t testini uygulayalım:

1. Sıfır hipotezi (H₀) ve alternatif hipotezin (H₁) tanımlanması:

H₀: Diyabet hastası erkekler ve bayanlar arasında vücut kitle indeksleri açısından fark yoktur

H₁: 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:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{SE(\bar{x}_1 - \bar{x}_2)} = \frac{(\bar{x}_1 - \bar{x}_2)}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

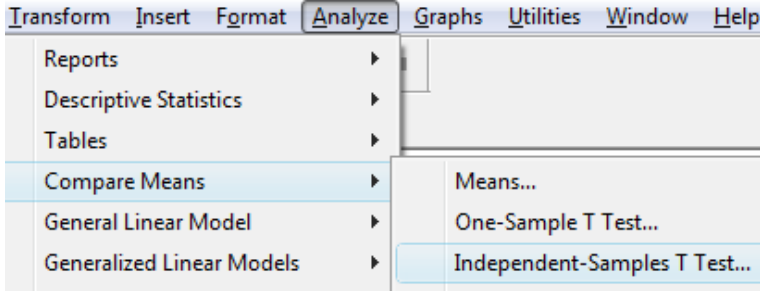
Burada *s*₁ ve *s*₂’den “pooled standard deviation” *s* hesaplanması gerekmektedir:

$$s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

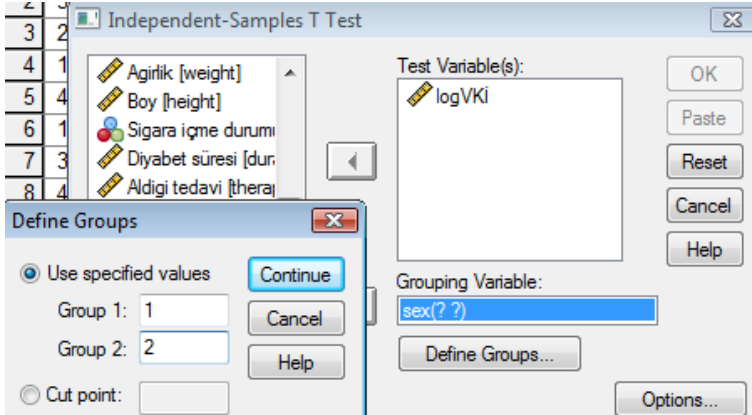
\bar{x}_1 (Erkek)	1,44
\bar{x}_2	1,47
<i>s</i> ₁	0,076
<i>s</i> ₂	0,082
<i>n</i> ₁	215
<i>n</i> ₂	189

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
logVKY	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	
									Lower	Upper
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 değerlerin 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

Aktürk Z, Acemoğlu H. Sağlık Çalışanları İçin Araştırma ve Pratik İstatistik 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 $VKİ_{\text{erkek}} - VKİ_{\text{bayan}}$ farkının hep eksi olduğu, sıfırı içermediği görülmektedir.

Elde ettiğimiz değer logaritmik bir değer olduğunu unutmamalıyız. Sonucumuzu klinik olarak anlaşılır hale dönüştürmek için antilog dönüşüm uygulamalıyız:

$$\text{Antilog}(0,048) = 10^{0,0486} = 1,12 \text{ kg/m}^2. \text{ Antilog}(0,017) = 10^{0,017} = 1,04 \text{ kg/m}^2.$$

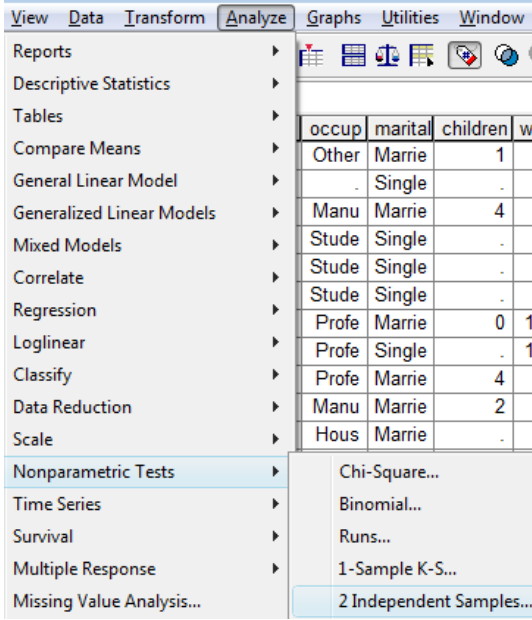
%95 GA: [-1,04 – -1,12]

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.

25. Numerik Veri İkidenden Fazla Bağımlı Grup

Amaç: Bu konu sonunda okuyucunun ikiden fazla bağımlı gruptan elde edilen numerik 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:

- Numerik 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

Diyabetik hastalarımıza ilaç, beslenme ve egzersizden oluşan tedavi programı uyguladığımızı varsayalım. Program süresince katılımcıların Hemogloblin A1c düzeylerini araştırmak istiyoruz (H_0 : Bireylerin uygulanan tedavi programı öncesi ve sonrasındaki Hemogloblin A1c düzeyleri arasında fark yoktur). Hemogloblin A1c ölçümleri tedavi programı uygulamadan önce, programın 3. ve 6. aylarında yapılmış olsun. Ölçmek istediğimiz değişken numeriktir. 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 numerik ölçüm söz konusudur. Kitabın başında da verilen istatistik testlerin uygulama yerleri akış şemalarına baktığımızda seçeceğimiz testin 2’den fazla bağımlı örneklemlerde tekrarlayan ölçümlerde varyans analizi (repeated measures ANOVA) veya Friedman varyans analizi olacağını görürüz.

Varsayımları

Tekrarlayan ölçümlerde varyans analizi, bağımsız gruplarda olduğu gibi, bağımsız değişken sayısının bir olduğu koşullarda tek yönlü varyans analizi (örn. cinsiyetin kan trigliserid düzeyine etkisi), birden çok olduğu koşullarda iki yönlü varyans analizi (örn. Cinsiyet ve mesleğin kan trigliserid 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

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:

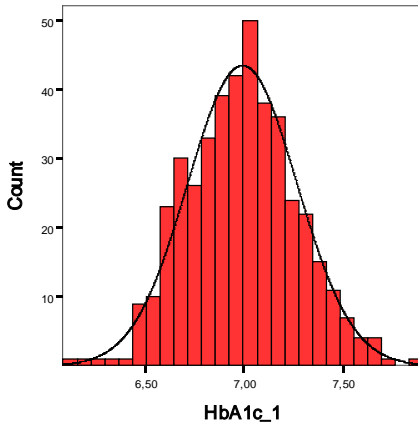
1. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

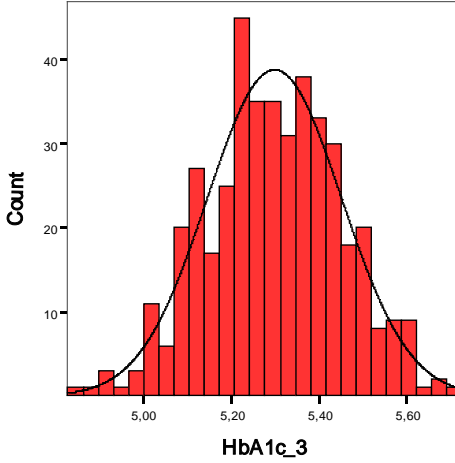
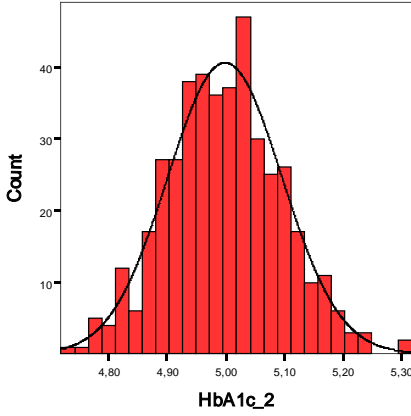
H_0 : Bireylerin uygulanan tedavi programı öncesi ve sonrasındaki 3. ve 6. aylardaki vücut ağırlıkları ortalamaları arasında fark yoktur

H_1 : Bireylerin uygulanan tedavi programı öncesi ve sonrasındaki Hemogloblin 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:





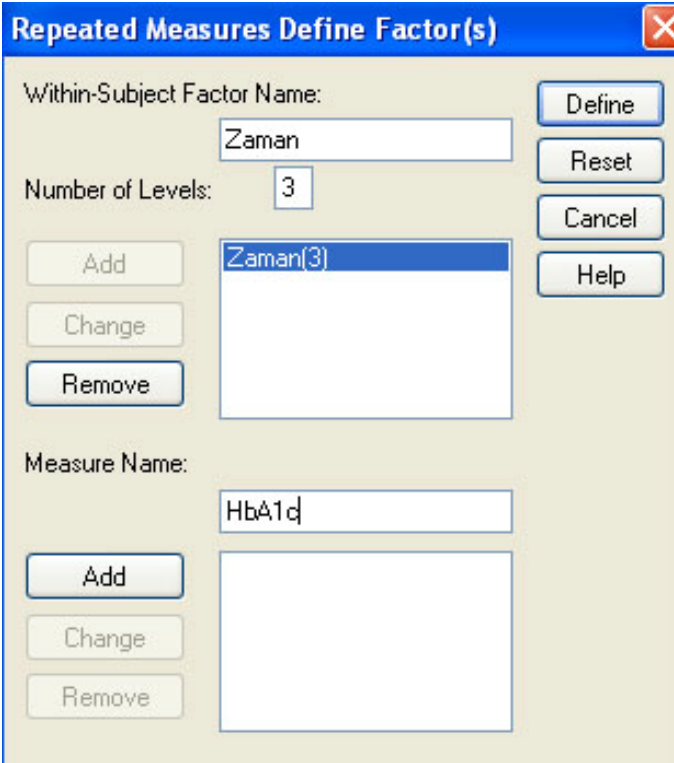
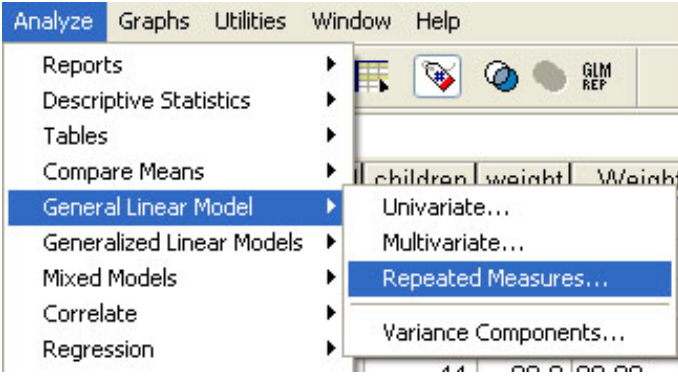
Varyansların homojenliği varsayımı ANOVA testi sırasında incelenecektir.

3. İlgili sıfır hipotezi için *test istatistiğinin* hesaplanması:

ANOVA testi grup için ve gruplar arası varyansların birbirine oranlanmasından (F oranı) oluşur. Bu oran F dağılımına uymaktadır.

SPSS ile yapacak olursak:

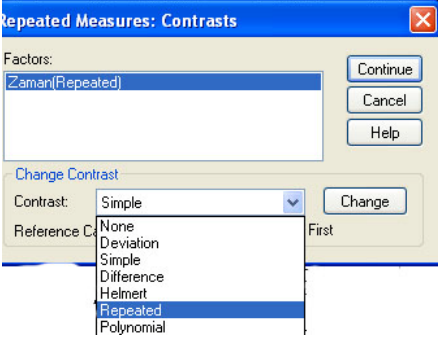
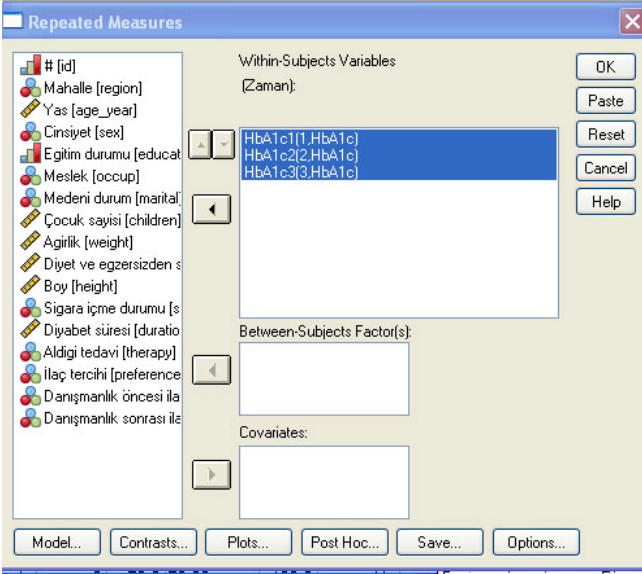
Analyze > General Linear Model > Repeated Measures > [Within-Subject Factor Name kutusuna, birden çok kez ölçülmüş olan bağımlı değişkeni; zaman (ay) ve Number of Levels kutusuna da ölçüm sayısı (3) yazılmalıdır. Add butonunu klikleyerek listeye alalım. Measure name kutusuna değişkenin adını örneğimizde "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.

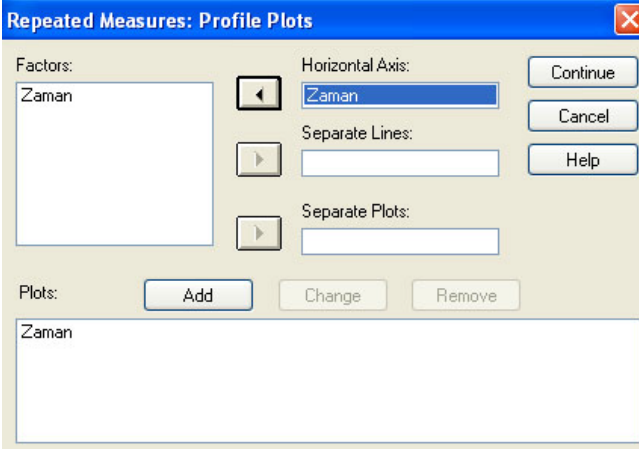
Contrast menüsü simple seçeneğini tercih edebileceğimiz gibi bu örnekte repeated seçeneği (birinci ile ikinci ikinci ile üçüncü ölçümler karşılaştırılır) seçilebilir.

>Contrast >Repeated>Change>Continue>



Grafik çizimi için;

>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 ($p=0.000$).

Mauchly's Test of Sphericity(b) (Mauchly'nin küresellik testi) incelendiğinde $p < 0.05$ olduğu için küresellik varsayımı sağlanmıştır. 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 ($p=0.000$).

Profile Plots HbA1c ölçüm ortalamaları çizgi grafik olarak verilmiştir.

Repeated Measures: Options

Estimated Marginal Means

Factor(s) and Factor Interactions: (OVERALL)
Zaman

Display Means for:

Compare main effects

Confidence interval adjustment: LSD (none)

Display

Descriptive statistics
 Estimates of effect size
 Observed power
 Parameter estimates
 SSCP matrices
 Residual SSCP matrix

Transformation matrix
 Homogeneity tests
 Spread vs. level plots
 Residual plots
 Lack of fit test
 General estimable function

Significance level: .05 Confidence intervals are 95%

Continue Cancel Help

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 Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon(a)						
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound	Greenhouse-Geisser	Huynh-Feldt	Lower-bound	Greenhouse-Geisser
Zaman	,704	150,311	2	,000	,772	,774		,500			

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

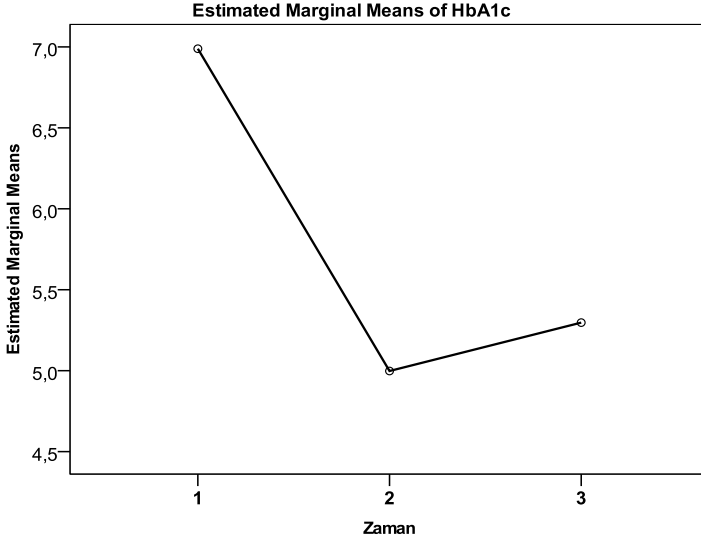
Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Zaman	Sphericity Assumed	991,215	2	495,608	13627,043	,000
	Greenhouse-Geisser	991,215	1,543	642,384	13627,043	,000
	Huynh-Feldt	991,215	1,548	640,515	13627,043	,000
	Lower-bound	991,215	1,000	991,215	13627,043	,000
Error (Zaman)	Sphericity Assumed	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



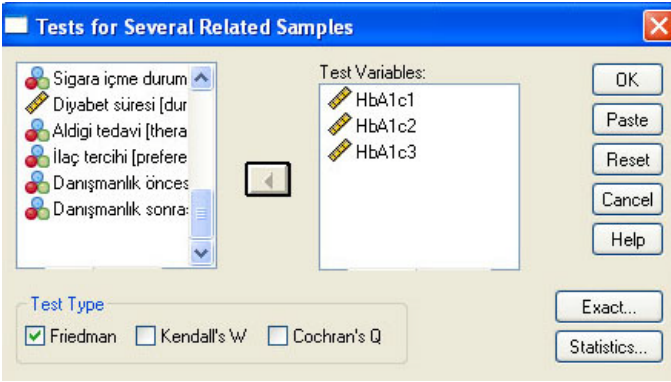
Varsayımlar karşılanamıyorsa

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 $p=0.000$ olduğundan üç farklı zamanda ölçümlerden en az birisinin diğerlerinden farklı olduğunu görürüz. Hangi grup/lar'ın 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 grupta 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

26. Numerik Veri İkiden Fazla Bağımsız Grup

Amaç: Bu konu sonunda okuyucunun ikiden fazla bağımsız gruptan elde edilen numerik 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:

- Numerik 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) numerik 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 uygulayamamız 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 dağılımalıdır.
- 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:

4. Sıfır hipotezi (H_0) ve alternatif hipotezin (H_1) tanımlanması:

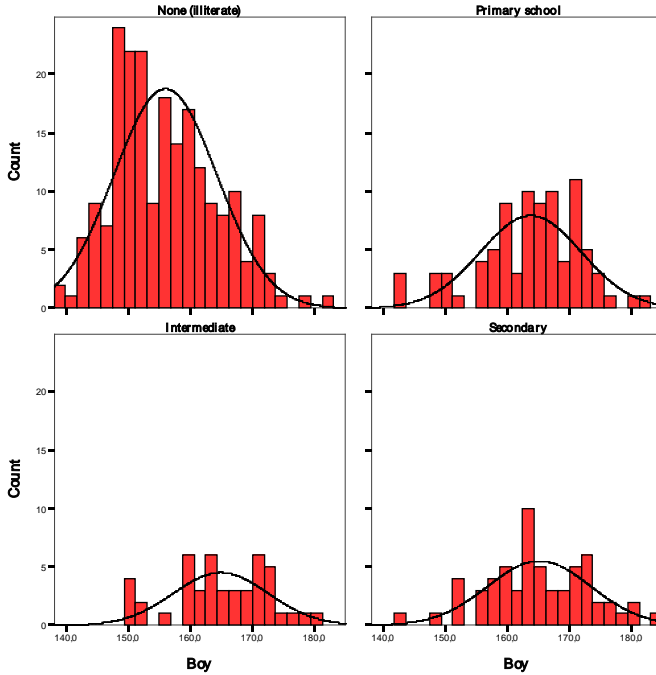
H_0 : Eğitim durumları farklı olan diyabet hastalarının boyları arasında fark yoktur.

H_1 : Eğitim durumları farklı olan diyabet hastalarının boyları arasında fark vardır.

5. Verilerin toplanması:

Veriler diyabet.sav verisetinde “height” ve “educat” değişkenlerinin altında kaydedildi.

Örneklem sayısı yeterlidir. Değişkenimizin gruplar açısından histogram grafiğine bakıldığında normal dağıldığı görülmektedir:



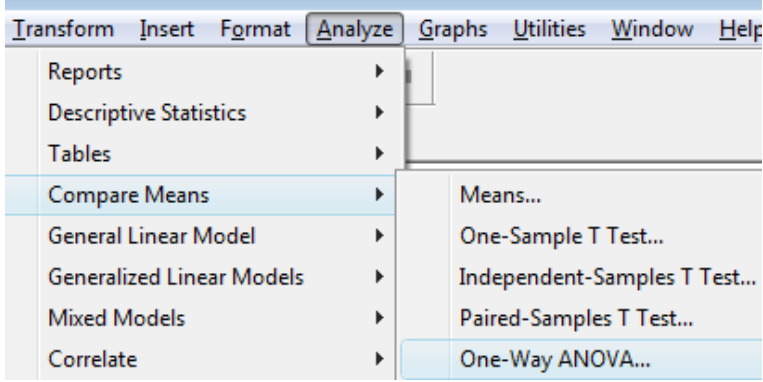
Varyansların homojenliği varsayımı ANOVA testi sırasında incelenecektir.

6. İlgili sıfır hipotezi için test istatistiğinin hesaplanması:

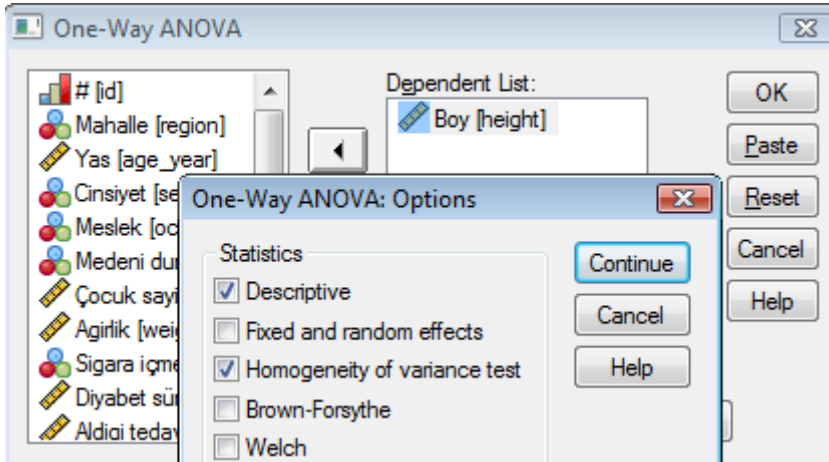
ANOVA testi grup için ve gruplar arası varyansların birbirine oranlanmasından (F oranı) oluşur. Bu oran F dağılımına uymaktadır.

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

	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.

7. Test istatistiğinden elde edilen değerlerin bilinen bir olasılık dağılımı ile karşılaştırılması:

F oranı için appendix A5'e baktığımızda p değerinin 0,05'ten küçük olduğunu görürüz. Zaten SPSS'te p 'yi sıfıra yakın bir değer olarak vermiştir.

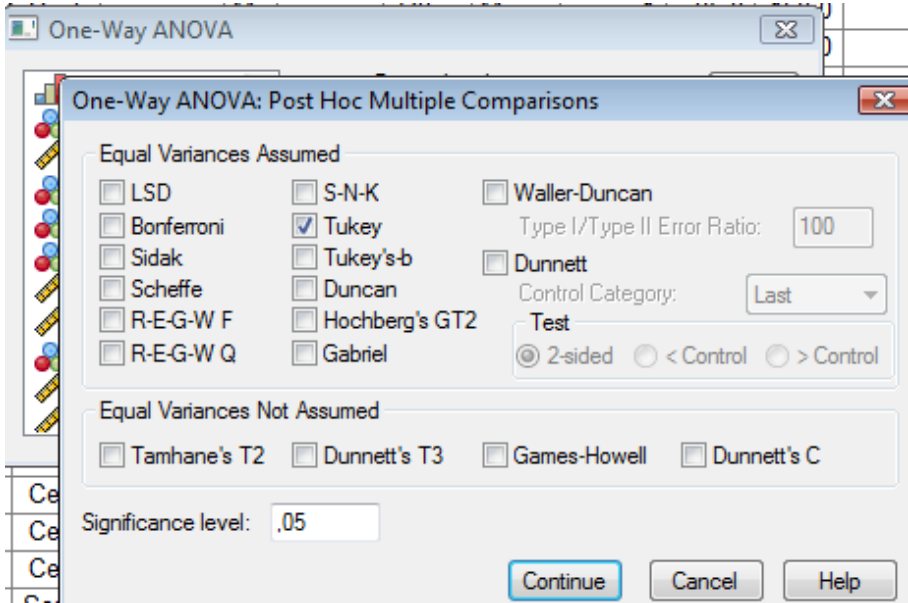
8. P değerinin ve sonuçların yorumlanması

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”

Aktürk Z, Acemoğlu H. Sağlık Çalışanları İçin Araştırma ve Pratik İstatistik
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

Aktürk Z, Acemoğlu H. Sağlık Çalışanları İçin Araştırma ve Pratik İstatistik 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.

Varsayımlar karşılanmıyorsa

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:

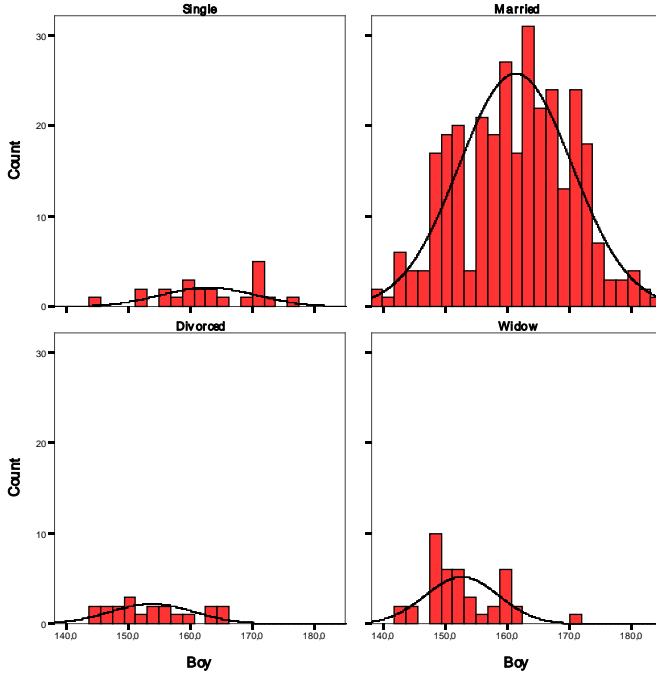
H_0 : Medeni durumları farklı olan diyabet hastalarının boyları arasında fark yoktur.

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:

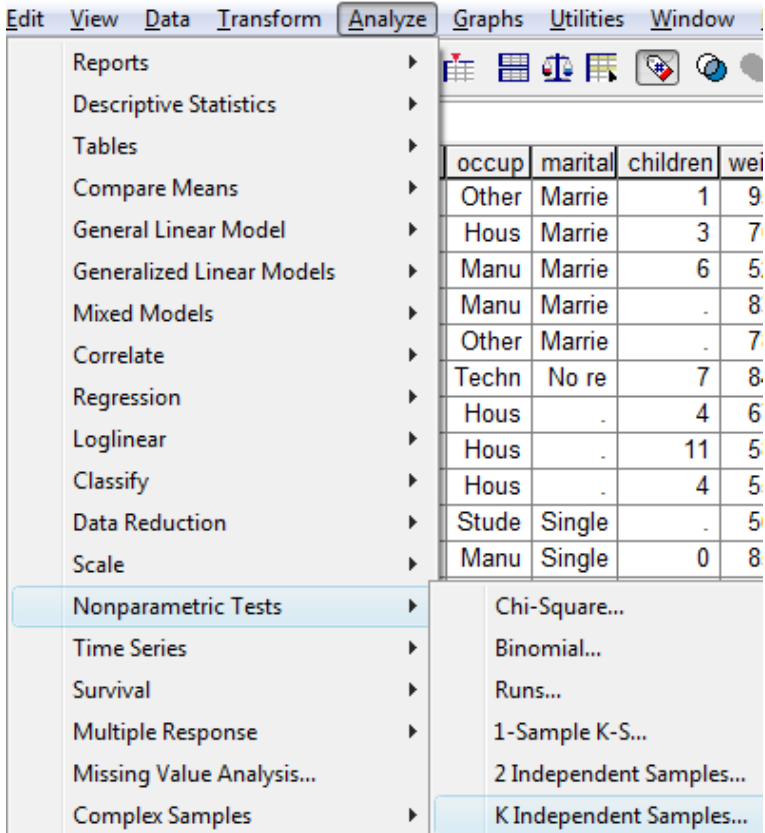
Test of Homogeneity of Variances

Boy

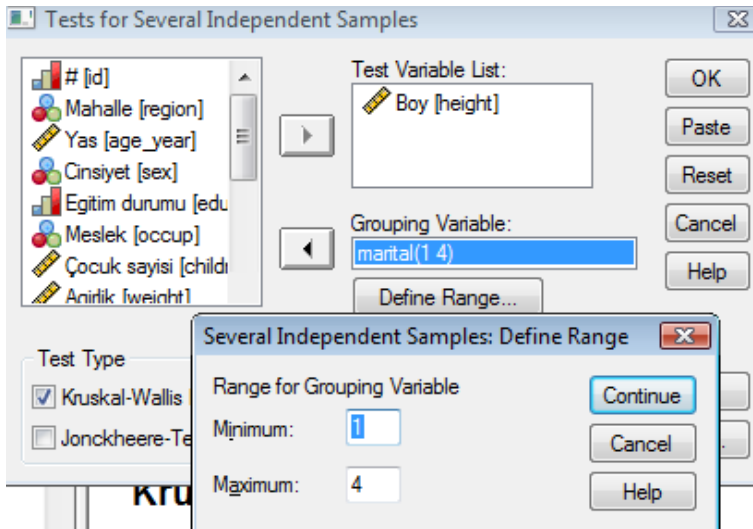
Levene Statistic	df1	df2	Sig.
4,852	3	392	,003

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.



Aktürk Z, Acemoğlu H. Sağlık Çalışanları İçin Araştırma ve Pratik İstatistik
Aşağıdaki çıktıları elde ederiz:

Ranks		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 değer bilinen bir olasılık dağılımı ile karşılaştırılması:

Kruskal-Wallis testi *Ki Kare* dağılımı göstermektedir. 47,3 değerini Tablo A3'ten bakacak olursak:

sd	İki yönlü p			
	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 →47,34
4	7.779	9.488	13.277	18.466

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.

27. Kategorik Veri Tek Grup

Amaç: Bu konu sonunda okuyucunun tek bir gruptan elde edilen kategorik verilerin belli bir toplum oranıtısı ile karşılaştırılmasında kullanılan tek oranıtıyı toplumla karşılaştırma ve Sign testinin oranıtılar için kullanımını konusunda bilgi sahibi olması ve bu analizleri yapabilmesi amaçlanmıştır.

Hedefler: Bu konu sonunda okuyucuların aşağıdaki hedeflere ulaşması beklenmektedir:

- Kategorik veri, tek grubu tanımlayabilmeli
- Tek oranıtı testinin varsayımını ve mantığını açıklayabilmeli
- Oranıtılar için işaret testinin kullanım yerini açıklayabilmeli
- Oranıtılar 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 oranıtı 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 oranıtısı $p = r/n$ olacaktır. İncelenen durumun toplumdaki π_1 oranıtısı 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 oranıtının ortalaması p , standart hatası ise

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

olacaktır.

www.aile.net/agep/istat/diyabet.sav 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 karşılaştırmanın 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:

z	İki yönlü p
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 verdiğimizizi 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 (H₀) ve alternatif hipotezin (H₁) tanımlanması:

H₀: Diyabetik hastalarda A ve B ilacı tercihi arasında fark yoktur.

H₁: 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 numerik 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' ≤ 10 olması durumunda r dikkate alınır.
- n' > 10 olması durumunda z değeri hesaplanır.

$$z' = \frac{\left| p - \frac{1}{2} \right| - \frac{1}{2n'}}{\sqrt{\frac{p(1-p)}{n'}}$$

¹ <http://www.medcalc.be/>

² <http://www.minitab.com/en-TR/default.aspx>

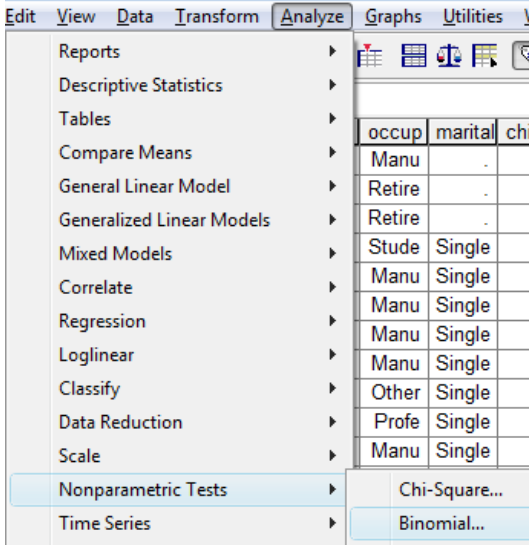
Aktürk Z, Acemoğlu H. Sağlık Çalışanları İçin Araştırma ve Pratik İstatistik
(burada π_1 yerine %50 yani $\frac{1}{2}$ alındığına dikkat ediniz).

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)	
Ylaç 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 değerin bilinen bir olasılık dağılımı ile karşılaştırılması:

- $n \leq 10$ olması durumunda Appendix A6'dan r değerine bakılır.
- $n > 10$ olması durumunda Appendix A1'den z değerine bakılır.

Appendix A1'den z değerini baktığımızda:

z	<i>İki yönlü p</i>
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.

28. Kategorik Veri İki Bağımlı Grup

Amaç: Bu konu sonunda okuyucunun iki gruptan elde edilen kategorik verilerin karşılaştırılmasında kullanılan Ki Kare (χ^2) ve 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
- 2x2 çapraz tabloda beklenen frekansları hesaplayabilmek
- Ki kare testinin varsayımlarını açıklayabilmek
- Kategorik veri iki bağımlı gruba örnek verebilmek
- SPSS'te Ki Kare testi yapabilmek
- SPSS'te McNemar testini yapabilmek
- Fisher exact testinin kullanım alanını açıklayabilmek

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) 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 sonrasında 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:

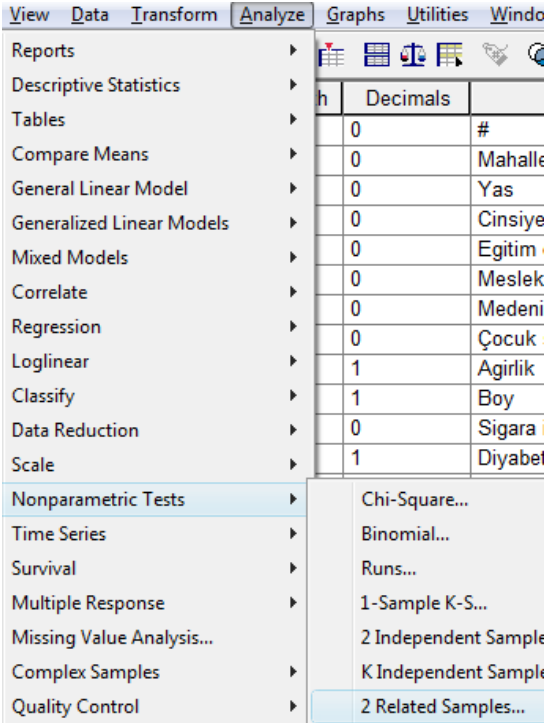
$$\chi^2 = \frac{(|b - c| - 1)^2}{b + c}$$

Bu test istatistiği, serbestlik derecesi 1 olacak şekilde Ki Kare dağılımı göstermektedir.

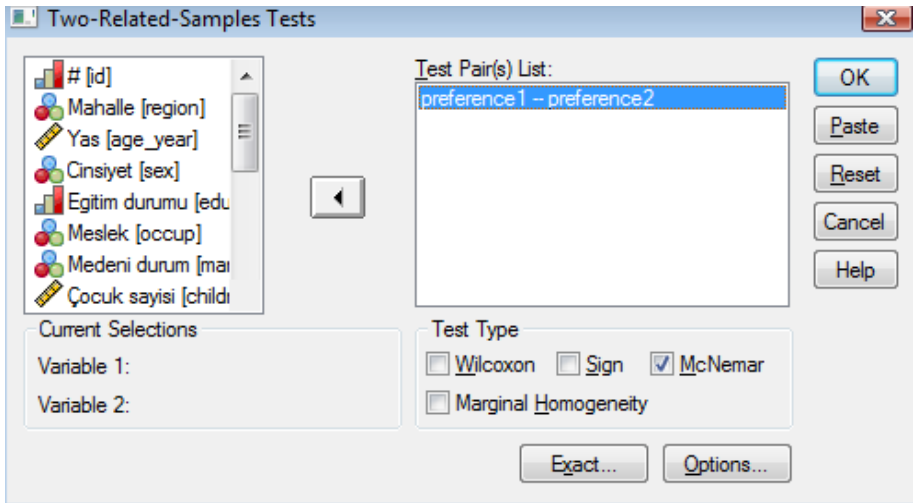
$$\begin{aligned}\chi^2 &= (|6 - 23| - 1)^2 / (6 + 23) \\ &= 256 / 29 \\ &= \mathbf{8,83}\end{aligned}$$

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.	,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 değerlerin bilinen bir olasılık dağılımı ile karşılaştırılması:

Appendix A3'ten χ^2 değerini baktığımızda:

SD	<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.

29. Kategorik Veri İki Bağımsız Grup

Amaç: Bu konu sonunda okuyucunun iki gruptan elde edilen kategorik verilerin karşılaştırılmasında kullanılan Ki Kare (χ^2) ve 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
- 2x2 çapraz tabloda beklenen frekansları hesaplayabilmek
- Ki kare testinin varsayımlarını açıklayabilmek
- Kategorik veri iki bağımlı gruba örnek verebilmek
- SPSS'te Ki Kare testi yapabilmek
- SPSS'te McNemar testini 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 seçmemeli).
- Biraz muhafazakâr da olsa, genel yaklaşım, beklenen frekansların tablonun her bir gözü için 5'ten büyük olması şeklindedir.

Testin mantığı

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	$p_1=a/n_1$	$p_2=b/n_2$	$p=(a+b)/n$

Tablo: gözlenen frekanslar

Ki Kare testi gözlenen frekanslarla beklenen frekanslar arasındaki farkı bakar. 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	$p_1=0,30$	$p_2=0,48$	$p=0,38$

Tablo: gözlenen frekanslar

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

Tablo: beklenen frekanslar

* Beklenen frekansların hesaplanması:

430 kişiden 236'sı erkek ise;

165 insülin kullandıktan 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

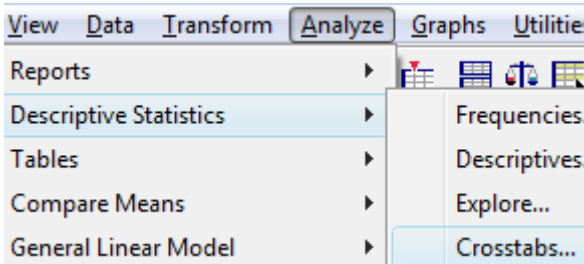
B = Dört gözlü tablonun her gözündeki beklenen frekanslar

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

Aktürk Z, Acemoğlu H. Sağlık Çalışanları İçin Araştırma ve Pratik İstatistik
*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 değer **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:

Appendix A3'ten χ^2 değerini baktığımızda:

SD	<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 sifıra 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.

30. Kategorik Veri İki'den Fazla Bağımsız Grup

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:

- *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.)*
- *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 (sıtır-1)x(sütun-1) olacak şekilde Ki Kare dağılımı göstermektedir.

Ki Kare tablosuna baktığımızda;

Mahalle * Medeni durum Crosstabulation

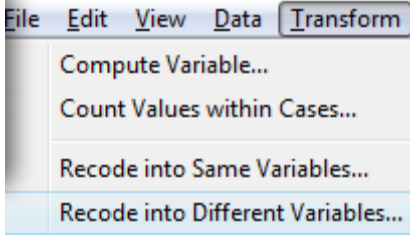
			Medeni durum				Total
			Bekar	Evi	Bosanmış	Dul	
Mahalle	Güney	Count	6	151	10	20	187
		Expected Count	13,4	142,0	11,1	20,5	187,0
		% within Mahalle	3,2%	80,7%	5,3%	10,7%	100,0%
	Dogu	Count	7	23	5	5	40
		Expected Count	2,9	30,4	2,4	4,4	40,0
		% within Mahalle	17,5%	57,5%	12,5%	12,5%	100,0%
	Merkez	Count	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.

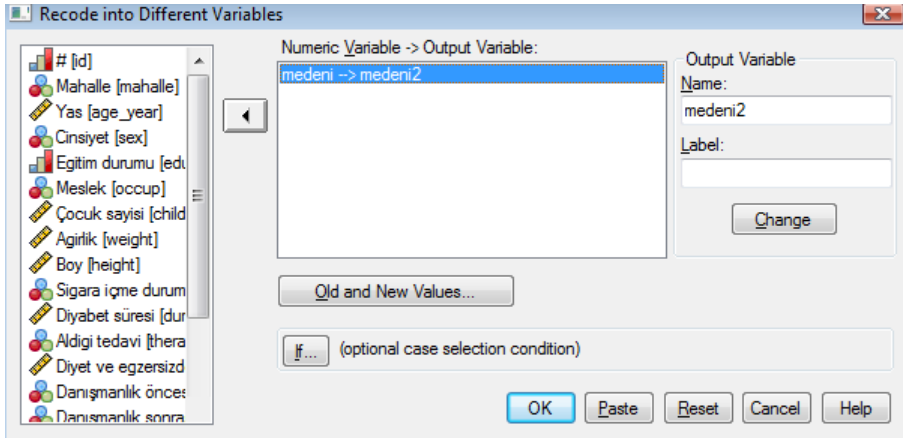
Dolayısıyla Ki Kare için gerekli varsayım karşılanmamaktadır.

"Boşanmış" ve "Dul" gruplarını tek bir grupta birleştirmemiz çözüm olabilir:

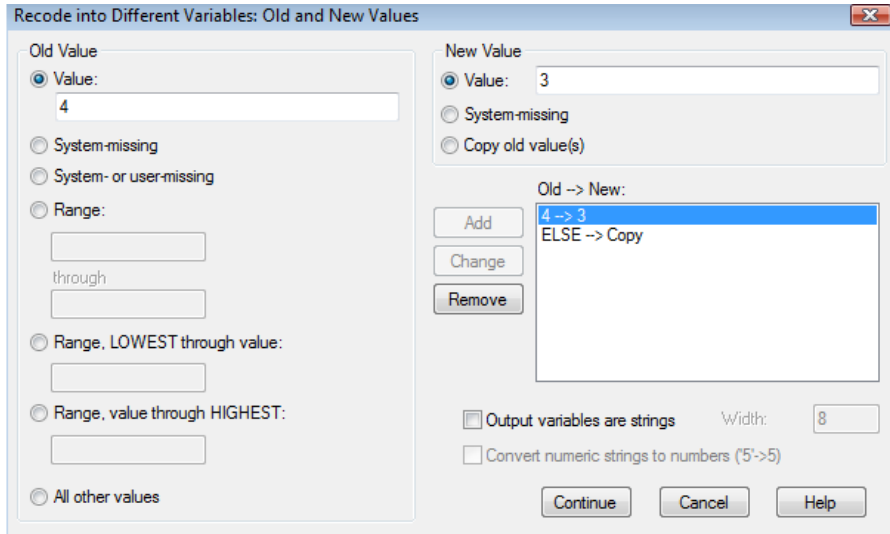
Transform>Recode into Different Variables>



[“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]



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	Egitim durumu	{1, None (illiter	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	Agirlik	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	Diyabet süresi	None	None	6	Right	Scale
13	therapy	Numeric	8	0	Aldigi 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	Danyımanlık ö	{1, Ynsülin}...	None	12	Right	Nominal
16	preference	Numeric	8	0	Danyımanlık s	{1, Ynsülin}...	None	12	Right	Nominal
17	yasgrup	Numeric	8	0	Yaş grupları	{1, 39 yaş ve	None	10	Right	Ordinal
18	medeni2	Numeric	8	0	Medeni durum	{1, Bekar}...	9	10	Right	Nominal

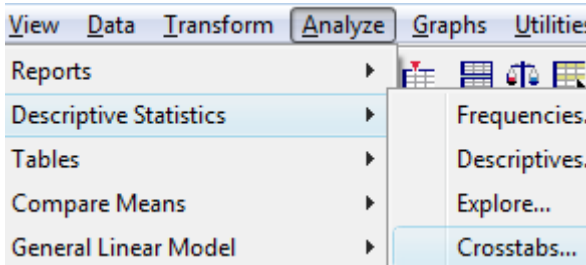
Yeni “medeni2” değişkenimizde sadece 1 hücrede (%11,1) beklenen frekans 5’ in altında olduğundan işlem yapabiliriz.

$$\chi^2 = (6-13,4)^2/13,4 + (151-142)^2/142 + (30-31,6)^2/31,6 + (7-2,9)^2/2,9 + (23-30,4)^2/30,4 + (10-6,8)^2/6,8 + (17-13,8)^2/13,8 + (145-146,6)^2/146,6 + (31-32,6)^2/32,6$$

$$=14,68$$

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.

Aktürk Z, Acemoğlu H. Sağlık Çalışanları İçin Araştırma ve Pratik İstatistik
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:

Appendix A3'ten χ^2 değerini baktığımızda:

SD	<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)x(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 bekarların 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 oranısının 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:

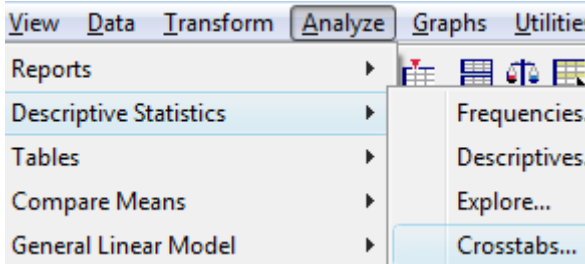
Veriler diyabet.sav verisetinde “educat” ve “smoking” değişkenleri altında kaydedildi.

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 görülmeyle 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:

Appendix A3'ten 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 değerinin ve sonuçların yorumlanması

P değeri 0,05'ten küçük olduğu için H_0 hipotezini reddederiz. Dolayısıyla diyabetik bireylerde yaşla birlikte sigara içme oranları artmaktadır.

31. Kategorik Veri İkidenden Fazla Bağımlı Grup

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

İkidenden fazla bağımlı gruplardan elde edilen veri dikotom ise örneğin evet-hayır, var-yok gibi. Bu durumda 2*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

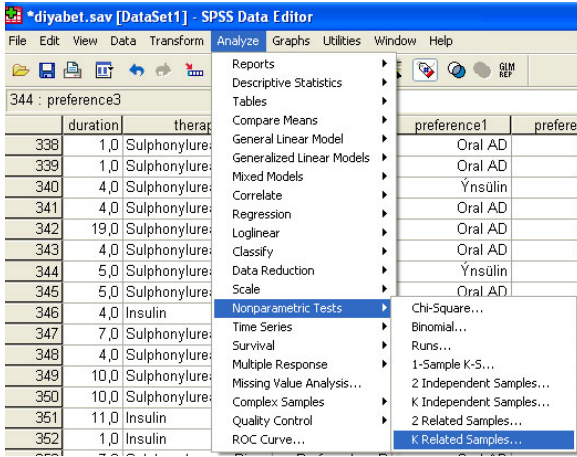
H_1 : Diyabet hastalarının ilaç tercihlerinde eğitim öncesi, sonrası birinci ve altıncı aylardan en az birinde fark vardır.

2. Verilerin toplanması:

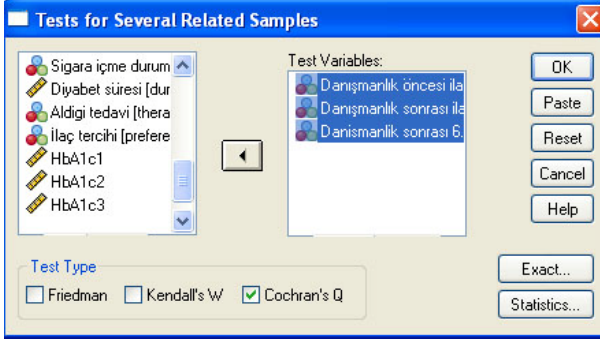
Veriler diyabet.sav verisetinde “preference1”, “preference2” ve “preference3” değişkenleri altında kaydedildi.

SPSS ile yapacak olursak:

Analyze>Nonparametric Tests> Related Samples>



[“preference1”, “preference2” ve “preference3” 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.

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.

32. Kategorik Veri İki'den Fazla Bağımlı Kategori

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 memnunuz, memnunuz ve çok memnunuz” 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*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.

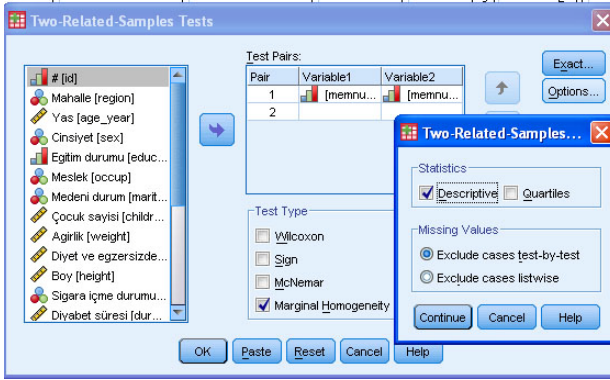
2. Verilerin toplanması:

Veriler diyabet.sav veri setinde “memnuniyet1” ve “memnuniyet2” 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.

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

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.

33. Kategorik Veri İkiden Fazla Bađımsız Kategori

Örnekleminizdeki bireylerin medeni durumlarının (evli, bekar, boşanmış = ikiden fazla kategori) ilaç antidiyabetik ilaç tercihleri (oral antidiyabetik, insülin = iki grup) ile ilişkili olup olmadığını incelemek istiyoruz.

Bađımsız deđişkenimiz ikiden fazla kategoriden oluştuğundan bu örnekte de örnek 8 ve 9'da olduğu gibi Ki kare testi uygulanmalıdır.

Bakınız sayfa 225

34. Korelasyon

Amaç: Bu konu sonunda okuyucunun iki numerik 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 numerik 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

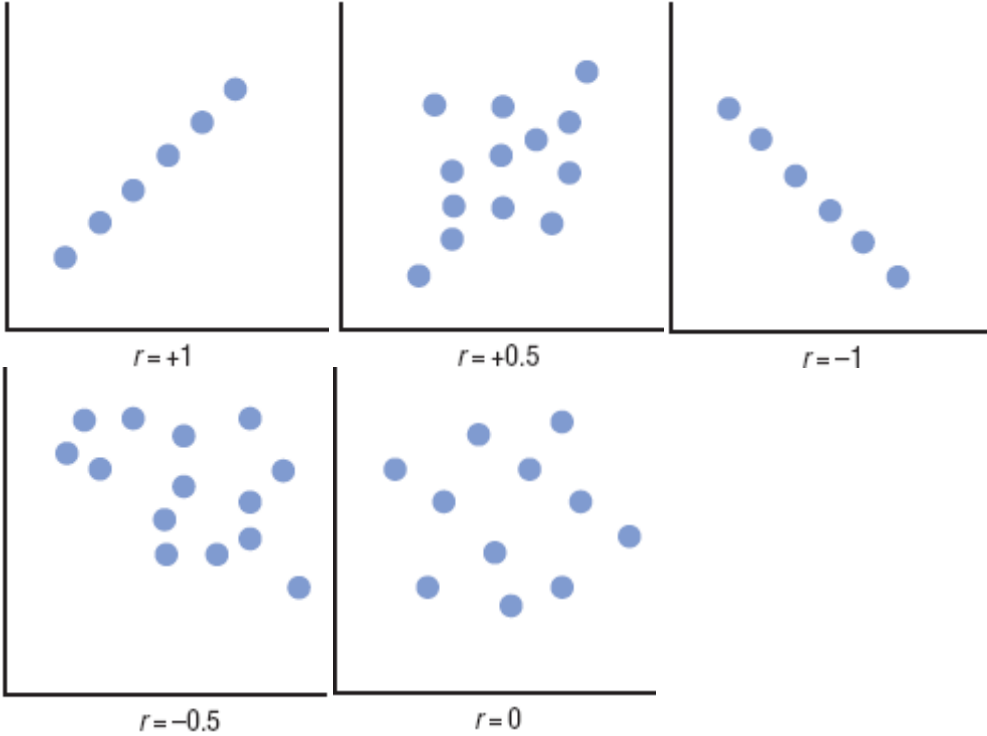
Grafikteki noktaların düz bir hat üzerinde dağılmaları halinde değişkenlerin arasında doğrusal bir ilişki olduğundan bahsederiz. **Pearson Product Moment Korelasyon Katsayısı**’nı (basitçe **korelasyon katsayısı** denir) hesaplayarak doğrusal ilişkiyi belirleyen çizgiyi bulabiliriz. Toplumdaki değeri “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 ötekini 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 örneklemdaki değişkenlerin en az ve en çok değerleri arasında geçerlidir. Örneklemden 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.

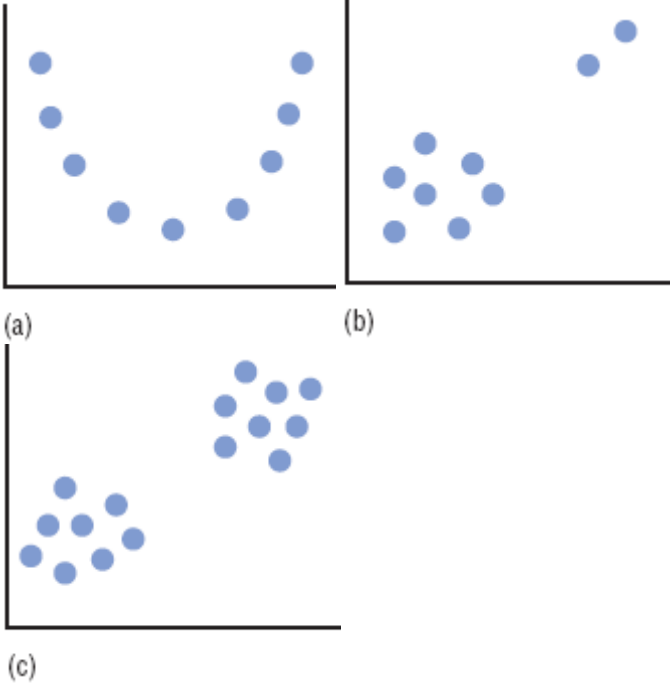


Şekil: r 'nin değişik değerlerini gösteren beş grafik.

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:

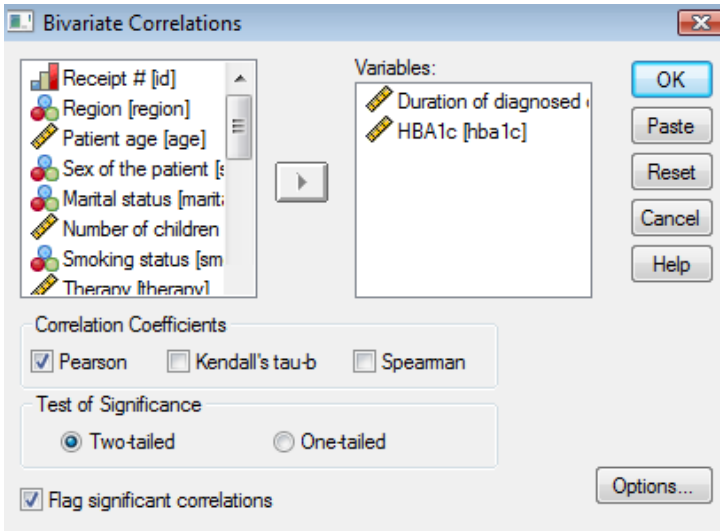
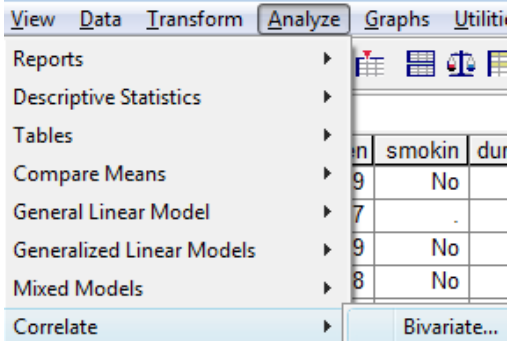
$n \leq 150$ ise r test istatistiği olarak kullanılır.

$n > 150$ ise T hesaplanır:

$$T = \frac{\sqrt{(n-2)}}{\sqrt{(1-r^2)}}$$

SPSS’te:

Aktürk Z, Acemoğlu H. Sağlık Çalışanları İçin Araştırma ve Pratik İstatistik
*Analyze>Correlate>Bivariate>[“duration” ve “hba1c” değişkenlerini
 “Variables” alanına geçirelim. “Pearson” kutucuğu işaretli]>ok*



Aşağıdaki çıktıyı elde ederiz:

Correlations

		Duration of diagnosed diabetes	HBA1c
Duration of diagnosed 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 Appendix A10'dan r bakılır.

$n > 150$ ise Appendix A2'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 Appendix A11'e bakılır
- r_s^2 hesaplanamaz.

35. Lojistik Regresyon

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 numerik 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_i} olarak hesaplanır. Hastalığın nadir olması halinde odds oranı **relatif risk** olarak 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 bilgilerimizi 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 loglikelihood**: 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 anlaşılır.
- **Sınıflandırma tablosu** (classification table): bu tablo modelin incelenen durumu ne kadar doğru sınıflandırabildiğini gösterir. Buradan duyarlık (sensitivite), özgüllük (specify) 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:

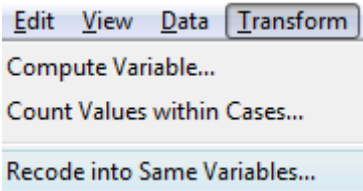
H_0 : Oluşturulan modeldeki değişkenlerin diyabet hastalarında sigara içme durumlarını tahmin etmede etkisi yoktur.

H_1 : Oluşturulan modeldeki değişkenler diyabet hastalarında sigara içme durumlarını tahmin ettirebilir.

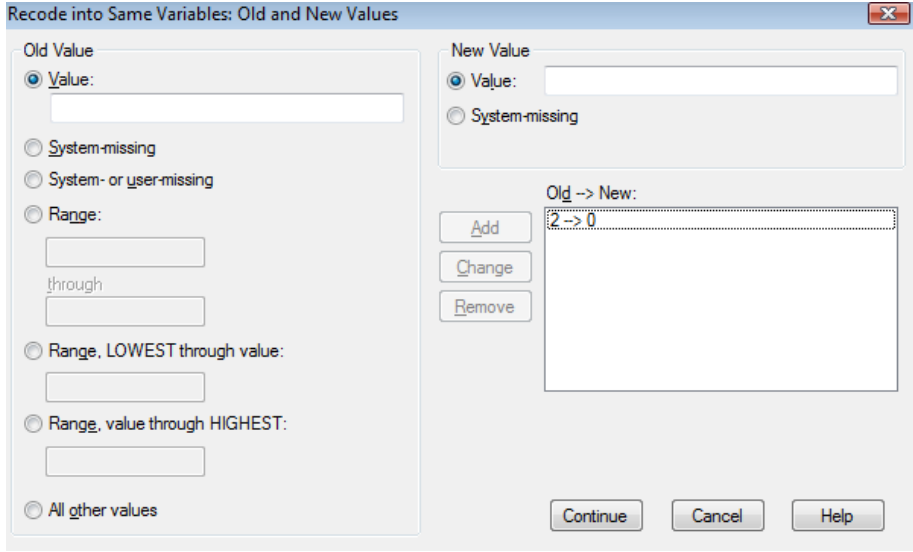
2. Verilerin toplanması:

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.

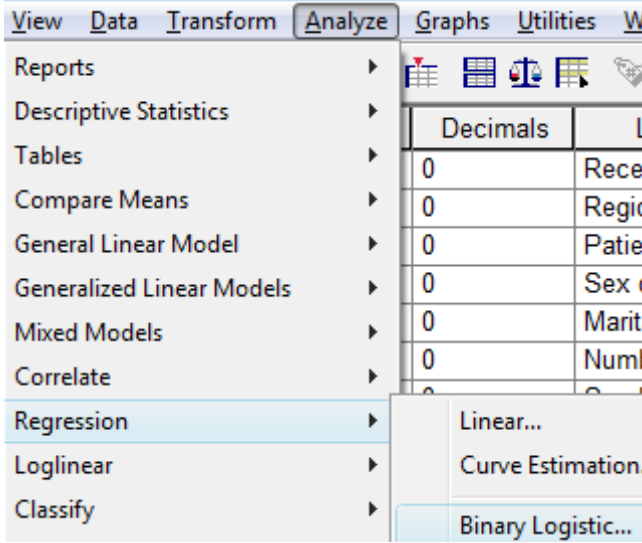
Name	Type	Width	Decimals	Label	Values
id	Numeric	8	0	Receipt #	None
region	Numeric	8	0	Region	{1, South}...
age	Numeric	8	0	Patient age	None
sex	Numeric	8	0	Sex of the pati	{1, Male}...
marital	Numeric	8	0	Marital status	{1, Single}...
children	Numeric	8	0	Number of chil	None
smoking	Numeric	8	0	Smoking statu	{1, Yes}... ..
duration	Numeric				
therapy	Numeric				
hba1c	Numeric				
glucose	Numeric				
systbp	Numeric				
diastbp	Numeric				
totchol	Numeric				

3. İlgili sıfır hipotezi için test istatistiğinin hesaplanması:

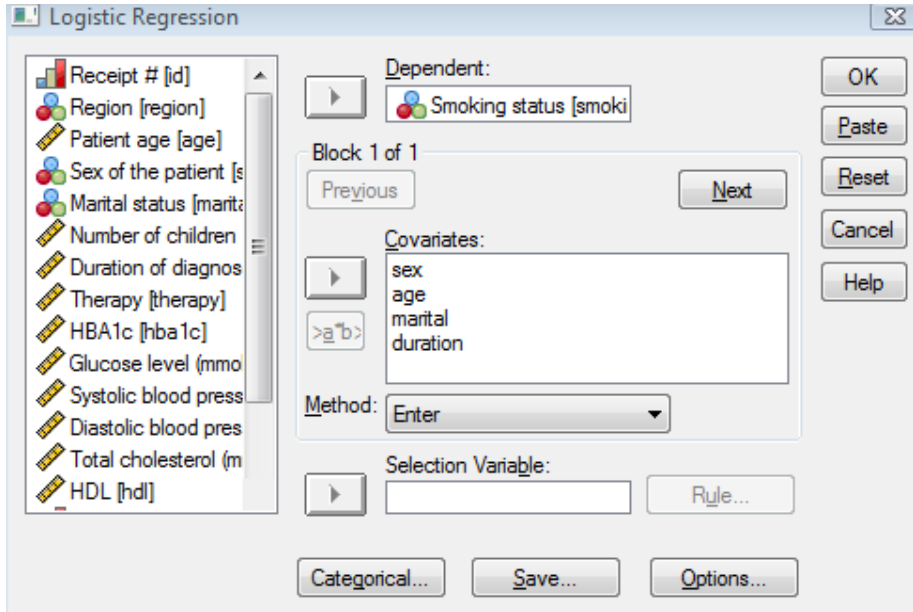
Öncelikle modelimizi belirlemeliyiz. Verisetimizdeki değişkenler içerisinde “age”, “sex”, “marital” ve “duration” değişkenlerinin sığra içme durumu üzerinde

Aktürk Z, Acemoğlu H. Sağlık Çalışanları İçin Araştırma ve Pratik İstatistik belirleyizi olabileceğini varsayıyoruz (modeli doğru seçmemizin önemini burada bir kez daha vurgulamakta yarar var).

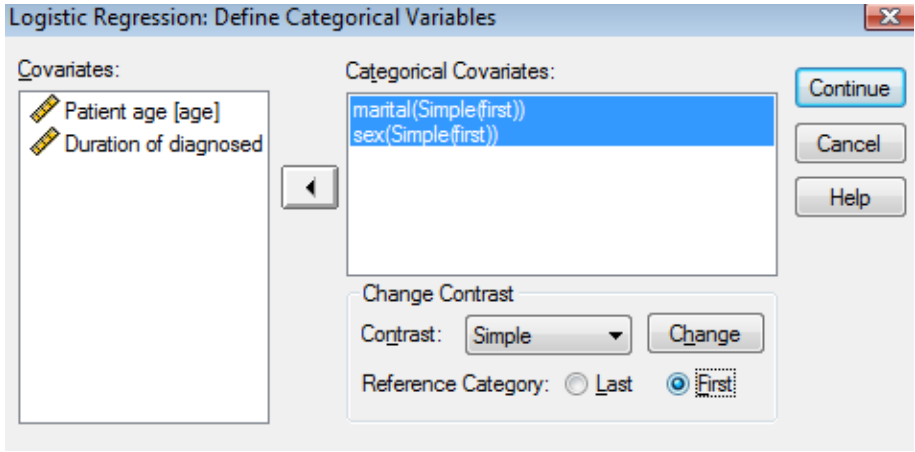
Analyze>Regression>Binary Logistic>



[“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 değilkenlerden herhangi birinde veri eksikliği olması haline o birey analize dahil 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 loglikelihood 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 Tablô

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 oranı %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 birarada değerlendirilmelidir.

50 yaşında, 10 yıldır diyabet hastası, evli bir bayan için lojistik regresyon eşitliğini yazacak olursak:

$$\text{logit}(p) = -8,7 - 0,05 \times 50 + 0,027 \times 10 - 0,569 \times 2 \text{ (evli)} - 2,594 \times 2 \text{ (erkek)} = -16,12$$

$$p = (e^{-16,12}) / (1 + e^{-16,12}) = 0,0000009312$$

Modelimize göre bu bireyin sigara içme olasılığı 100 milyonda 9'dur.

5. P değerinin ve sonuçların yorumlanması

Modelimizin uygunluğu gerek -2 loglikelihood, gerekse sınıflandırma tablosuna göre iyi değildir. Bununla birlikte modele itibar edilip kullanılması düşünülürse dahil edilen değişkenler içerisinde sigara içmeye anlamlı olarak etki eden değişkenin sadece cinsiyet olduğu görülmektedir. Bayan olmak sigara içimi açısından koruyucudur.

Alıştırma 11

Lojistik regresyon konusunu pekiştirmek için aşağıda bağlantısı verilen makaleyi okuyunuz ve soruları cevaplandırınız.

İyit 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 sensitivite ve spesifitesi (duyarlık ve özgüllüğü) kaçtır? Doğru sınıflandırma oranı kaçtır?
4. Modelin -2 log likelihoodu kaçtır?
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 kaçtır? Bunu klinik olarak nasıl yorumlarız?
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. 299

Alıştırma 12

Lineer regresyon ve Lojistik regresyon konuları, günlük pratiğimizde sık karşılaştığımızdan 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 dahil 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. 300

36. Örneklem Hesabı

Amaç: Bu konu sonunda katılımcıların 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, önemlilik derecesini açıklayabilmek
- Çeşitli istatistik analizler için gerekli örneklem hesabını yapabilmek
 - Independent samples t test (Bakınız sayfa 149)
 - Ki kare
 - Tek yönlü ANOVA

Örneklem sayısı ve güç hesabında birbirine etki eden faktörlerden daha önce bahsedilmişti (Bakınız S. 147). 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 4 faktör vardır:

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. **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.

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/>.

Tanımlayıcı istatistikler için örneklem büyüklüğü formülleri

Sonuç ölçütümüz kategorik ise:

Hedef toplumdaki birey sayısını bilip bilmemize göre farklı formüller kullanırız.

$$\text{Hedef kitledeki birey sayısı bilinmiyorsa: } n = t^2 * p * q / d^2$$

$$\text{Hedef kitledeki birey sayısı biliniyorsa: } n = (N * t^2 * p * q) / [d^2 * (N - 1) + t^2 * p * q]$$

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 örneklem hatasıdır.

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 güvenirlilik aralığında (alfa = 0.05), d = 0.05 örneklem hatası içerebileceğini kabul ediyorsak, kaç kadına anket uygulaymamız gerekir?

Verilerimizi formüle uygulayalım:

$$p = 0.20$$

$$q = 0.80 (1-0.20)$$

t = 1,96 (alfa = 0.05 de serbestlik derecesine göre tablodan bulunmuştur.

d = 0.05 (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 numerik ise:

$$n = (t^2 * s^2) / d^2$$

n : Örnekleme alınacak birey sayısı

s : İncelenen numerik değişkenin tahmini standart sapması

t : Belirli bir anlamlılık düzeyinde, t tablosuna göre bulunan teorik değer

d : Etki büyüklüğü (farkların önemlilik derecesi)

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?

$$n = (1.98^2 * 4^2) / 0.5^2 = \mathbf{251 \text{ kişi}}$$

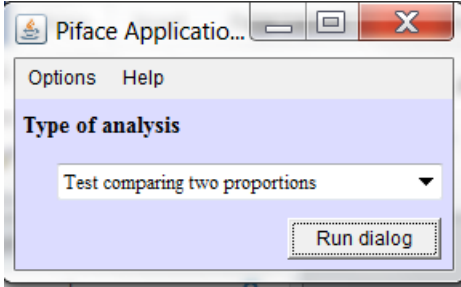
Hipotez testleri için örneklem büyüklüğünün hesaplanması

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.

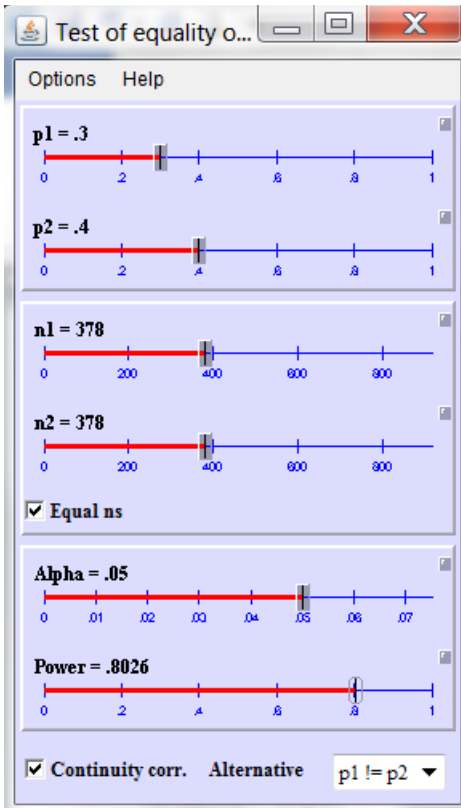
[Araştırmamızın çıktısı sigara içme durumu (evet/hayır) kategorik bir değişken.
Diabeti olan ve olmayan olmak üzere iki grubumuz var]

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.



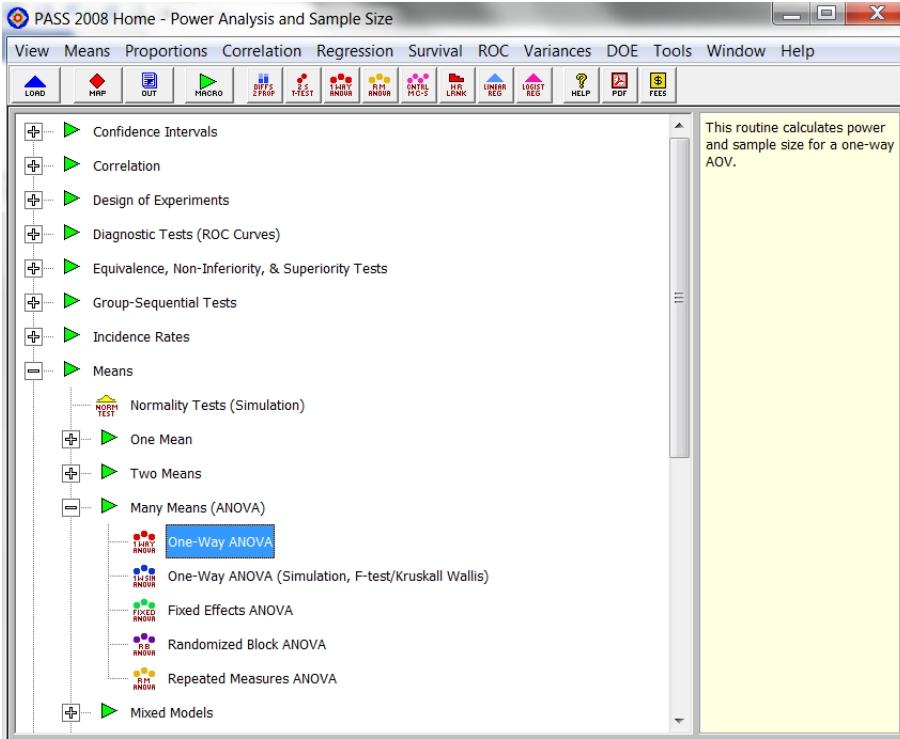
Görüldüğü gibi alfa = %5, iki yönlü p (alternatif hipotez = p_1 ile p_2 eşittir) ve her iki grupta eşit sayıda birey olacak şekilde gruplar arasındaki %10'lu bir farkı saptamak için %80'lik güç düzeyinde her iki grupta da 378'er kişilik bir örneklem gerekmektedir.

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 yanılma payı ve %80 güç ile gösterebilmek için araştırmamızın her bir kolunda kaç bireyin olması gerektiğini hesaplamak istiyoruz.

[Araştırmanın çıktısı kan şekeri değerleri (numerik). Eğitim durumlarına göre 4 grubumuz var (okur yazar / ilköğretim / üniversite / doktora)]

PASS yazılımını¹ kullanarak:

Means > Many Means (ANOVA) > One-Way ANOVA



Açılan pencerede örnekleme etki edecek faktörleri belirtiniz (alfa yanılma 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>

PASS: One-Way Analysis of Variance

File Run Means Proportions Correlation Regression Survival ROC Variances DOE Tools Window Help

Plot Type Symbols/Background Iterations Template
Data Reports Axes/Legend/Grid Plot Text

Load the PASS Help System.

Solve For
Find (Solve For):
Power and Beta

Error Rates
Power (1-Beta):
0.80
Alpha (Significance Level):
.05

Effect Size
Means
Hypothesized Means:
90 100 110 115
Planned Comparisons
Contrast Coefficients (Optional):
None
Standard Deviation
S (Standard Deviation of Subjects):
15
Standard Deviation Estimator

Sample Size / Groups
Sample Size Multiplier:
n (Sample Size Multiplier):
1 to 20 by 1
Groups:
k (Number of Groups):
4
Group Sample Size Pattern:
Equal

Template Id: Reset Guide Me

Aşağıdaki çıktıyı elde ederiz:

One Way ANOVA Power Analysis
Page/Date/Time 1 16.05.2010 10:22:05

Numeric Results

Power	Average n	Total k	Total N	Alpha	Beta	Std Dev of Means (Sm)	Standard Deviation (S)	Effect Size
0,00000	1,00	4	4	0,05000	1,00000	9,60	15,00	0,6401
0,14689	2,00	4	8	0,05000	0,85311	9,60	15,00	0,6401
0,28306	3,00	4	12	0,05000	0,71694	9,60	15,00	0,6401
0,42217	4,00	4	16	0,05000	0,57783	9,60	15,00	0,6401
0,54999	5,00	4	20	0,05000	0,45001	9,60	15,00	0,6401
0,65962	6,00	4	24	0,05000	0,34038	9,60	15,00	0,6401
0,74895	7,00	4	28	0,05000	0,25105	9,60	15,00	0,6401
0,81888	8,00	4	32	0,05000	0,18112	9,60	15,00	0,6401
0,87185	9,00	4	36	0,05000	0,12815	9,60	15,00	0,6401
0,91089	10,00	4	40	0,05000	0,08911	9,60	15,00	0,6401
0,93899	11,00	4	44	0,05000	0,06101	9,60	15,00	0,6401
0,95882	12,00	4	48	0,05000	0,04118	9,60	15,00	0,6401
0,97255	13,00	4	52	0,05000	0,02745	9,60	15,00	0,6401

Görüldüğü üzere %80' lik bir güce ulaşmak için araştırmanın her bir kolunda yaklaşık 8 kişi olması (toplam 32 kişi) yeterlidir.

Alıştırma 13

Örneklem ve araştırma tasarımı problemleri. Örneklem hesabı için Russ Lenth'in modülünü kullanabilirsiniz: <http://www.stat.uiowa.edu/~rlenth/Power/>.

1

Bulduğunuz 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 toplumunkiyile 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

Bulduğunuz 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?
- Kaç kişilik örneklem gerekir?

4

Bulduğunuz 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 önerimli 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. 301

37. Kendini Değerlendirme Çalışması

Sağlık Çalışanları için Araştırma Yöntemleri ve İstatistik

Kendini Değerlendirme Sınavı

Değerli Okuyucumuz,

Bu sınavımızın amacı öğrendiklerimizi tekrar etmek, bilgilerimizi pekiştirmek ve eksiklerimizi farkedip 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 <http://www.aile.net/agep/istat/sinav.sav> adresindeki SPSS veri setini kullanınız.

Sınav süresi sonunda sayfa 269'daki "Kendini Değerlendirme Çalışması Çözümleri" bölümüne bakarak sınavınızı puanlandırınız. Arşiv ve araştırma amaçlı kullanılmak üzere sınavınızı ve puan durumunuzu zekeriya.akturk@gmail.com adresine gönderirseniz seviniriz.

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 eğrisi nde;
 - 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. Onbirinci 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
- Randomize kontrollü deney
 - Kohort
 - Vaka kontrol çalışması
 - 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

38. Kendini Değerlendirme Çalışması Çözümleri

1. Verileri nasıl sınıflandırırsınız? Ordinal veriye örnek veriniz.
3 puan

Numerik (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

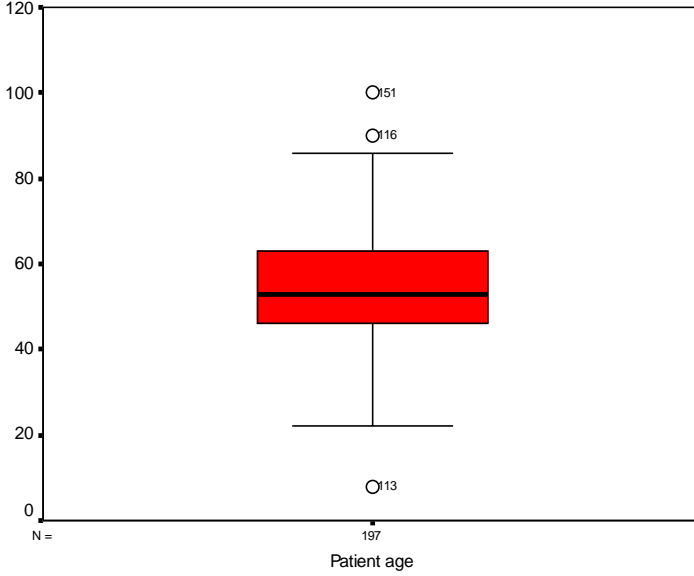
Oran ve hızın farkı hızın zamanı içermesidir. Erkek ve kızların sigara içme oranı (sigara içen erkekler/sigara içen kızlar). **(1 puan)** Bebek ölüm hızı (Bir yılda 0-30 gün arası ölen bebekler/o yıldaki tüm canlı doğumlar) **(1 puan)**

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)**
(Boxplot çizene ekstra 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) uygularız? 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:

$$(200 \times 40 + 100 \times 10 + 600 \times 80 + 50 \times 10) / (200 + 100 + 600 + 50) = 60,5$$

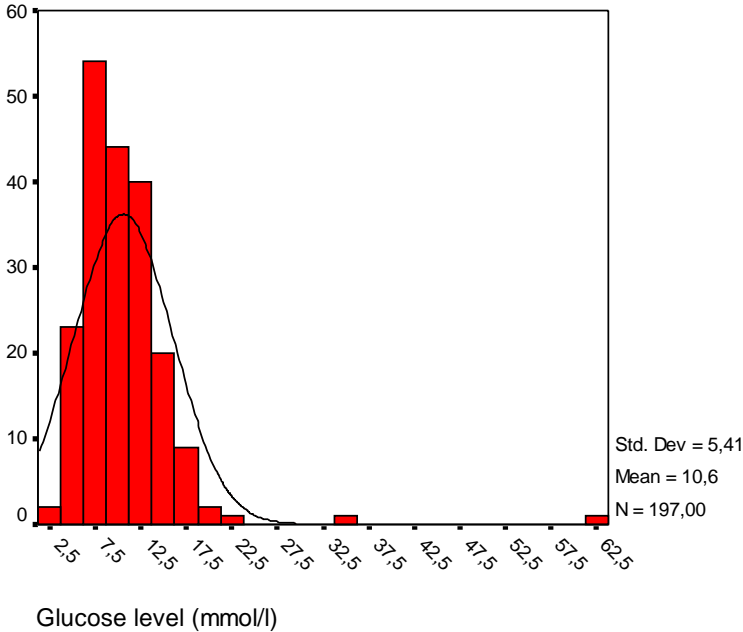
Genel ağırlıklı ortalamayı (60,5) bulana 3 puan

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

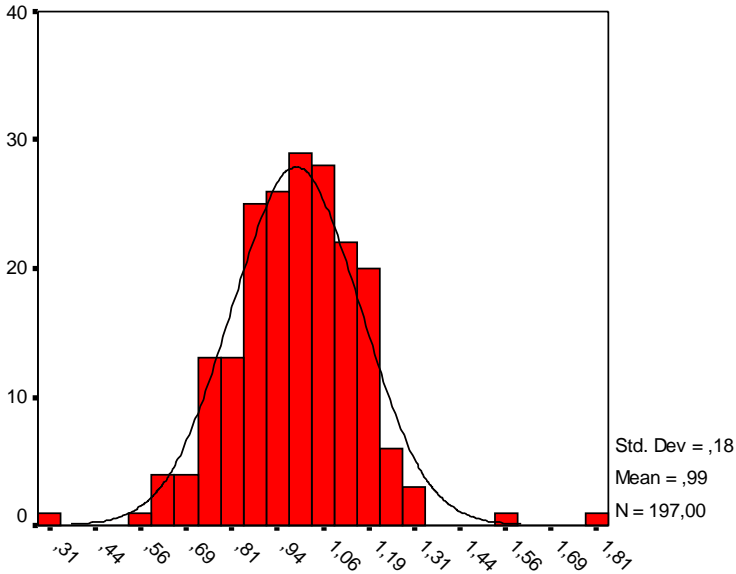
katından büyük olduğundan veri normal dağılmamaktadır. Değeri + olduğundan sağa eğimli olduğunu anlıyoruz. **(yukarıdakilerden herhangi birisini belirtene 3 puan)**

(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 uygulanacağı belirtilmesi 3 puan

Logaritmik dönüşümün yapılarak histogram veya skewness değerinin verilmesi 5 puan.

13. Bir örneklemede 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

Toplumun 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.

Toplumun 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 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 halukarda 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.)

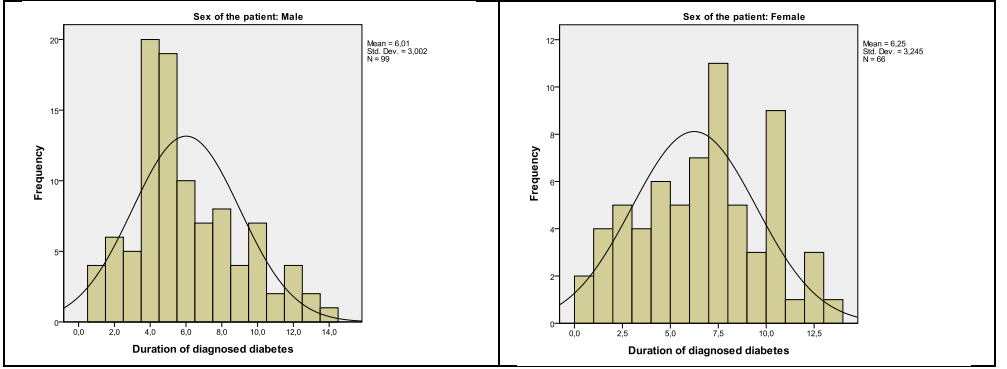
Mutually exclusive kuralını açıklayanlara ek 2 puan.

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 numerik değişkenimiz histogram grafiğinde her iki grupta da normal dağılıma yakındır:



Bu durumda numerik 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 dağılmadığına dikkat ediniz).

7 puan

Numerik 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 binomial 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.

(H₀'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

df	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

df	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

Paydanın SD	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 – Alıřtırmaların çözümleri

Not: alıřtırmaların çözümleri istatistik kurslarında katılımcıların verdikleri cevaplardan derlenmiřtir. Bu cevaplar düzeltme ve eleřtiriye açıktır.

Alıřtırma 1 (Sayfa 39)

Arařtırma anketinin ařağıdaki řekilde deęiřtirilmesi önerilebilir:

Diyabet Arařtırma Anketi

Hasta no:

S1. Doğum tarihiniz:.....

S2. Cinsiyetiniz:

1 () Kadın

2 () Erkek

S3. Kaç yıldır diyabet hastasısınız?:.....

S4. En son HbA1c deęeriniz kaçtı?..... 99 () Bilmiyorum

S5. Son iki hafta içerisinde hastalıęınız nedeniyle günlük iřlerinizi yapmakta ne kadar zorlandınız?

1 () Hiç zorlanmadım

2 () Biraz zorlandım

3 () Epey zorlandım

4 () Çok zorlandım

5 () Hiçbir iř yapamadım

S6. Diyabet için hangi ilaçları kullanıyorsunuz? Yazınız.

1.....

2.....

3.....

4.....

5.....

6.....

S7. Muayenenizde diyabete baęlı herhangi bir komplikasyon saptandı mı?

0 () Hayır

1 () Evet

Yapılabilecek yorumlar:

- 1) Anket numarası / irey numarası her veri setinde mutlaka bulunmalıdır.
- 2) Kiřilerin kaç yaşında olduğunu deęil 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-diamicron, 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?

Alıştırma 2 (Sayfa 48)

SPSS'te Variable View'de

- 1- Anket ve soru numarası verilebilirdi.
- 2- Yaş numerik bir değişkendir. Kategorize edilmemeliydi.
- 3- Income numerik bir değişkendir. Kategorize edilmemeliydi.
- 4- Age yerine dogum 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) numerik bir değişkendir. Kategorize edilmemeliydi.

10- Arabistan’da geçirilen zamana ordinal yerine nominal yazılmış

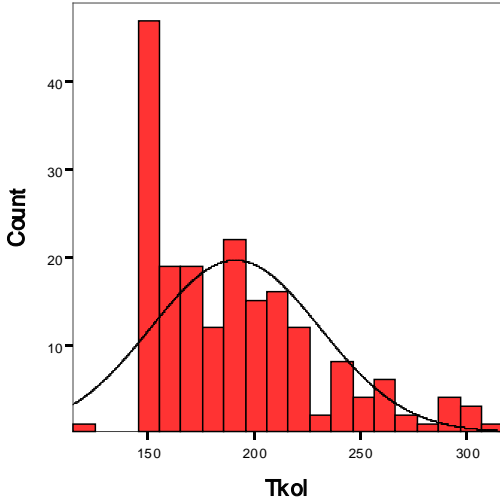
11- Din (Religion) seçeneğine “4=Diğer” değeri eklenebilirdi.

SPSS’te Data View’de

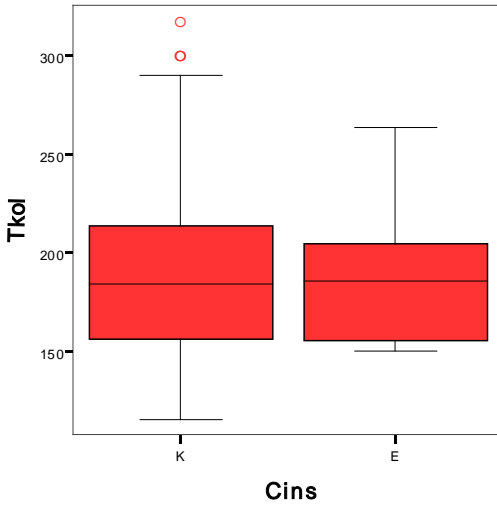
- 1- Education’da bir adet 11 değeri var
- 2- Income’da üç adet 4 değeri girilmiş
- 3- Life’da iki adet 4 değeri girilmiş
- 4- Knowledge’da bir adet 11 değeri girilmiş
- 5- Real age başlığında 16 tane missing değer var. Bunların dışında yaş dağılımı ise uç değer içermiyor.

Alıştırma 3 (Sayfa 61)

- 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.

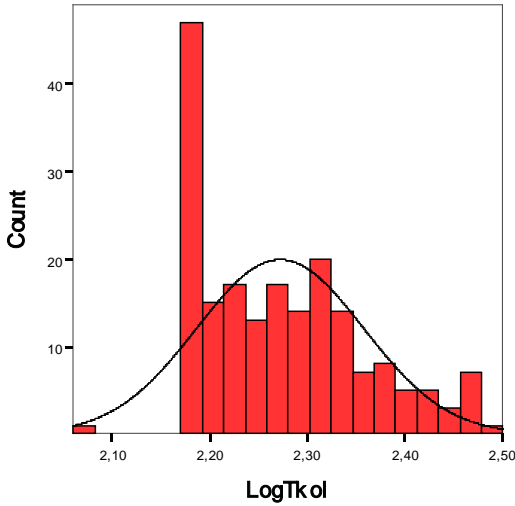
En küçük değer 115, en büyük değer 290 civarında

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-Yaş AO=45.93

Medyan=46.00

Mod=45 (birden fazla mod değeri olduğu ancak en küçük değerin gösterildiği belirtilmiş)

Yaş

N	Valid	189
	Missing	5
Mean		45,93
Median		46,00
Mode		45(a)

a Multiple modes exist. The smallest value is shown

8-Yaşta 5 adet missing veri var. En küçük değer 16, en büyüğü 78.

9-E=25, K=169,

Boy AO=157.54cm,

Boy cinsiyete göre ağırlıklandırılmış AO=158.73cm

Alıştırma 4 (Sayfa 74)

- 1- Bu durumda 20 depresyonlu, 20 psikozlu, 60 başka tanıli 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.

Alıştırma 5 (Sayfa 94)

1-

- a. Yaş, boy, kilo, BKİ, B/K oranı, triseps deri altı kalınlığı, kan basıncı, ve kanda bakılan tüm değişkenler numerik sürekli;
- b. kaybedilen akraba, anne, kardeş numerik 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 yüzdeler belirtilmiş (hepsi numerik türetilmiş veriler).

2- Sistolik kan basıncı ve BKİ'yi için saçılma grafiği kullanılmış. İki numerik 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.

Alıştırma 6 (Sayfa 108)**Soru 1-**

Kızlar için: AO: 18.38 St sapma: 1.37 Serbestlik derecesi: 1563 n: 1564 SEM: $1.37/\sqrt{1564} = 0.035$ $\%95GA = 18.38 \pm 1.965 \times 0.035$ $\%95GA = 18.38 \pm 0.069 = 18.31-18.45$	Erkekler için: AO: 18.25 St sapma: 1.37 Ser derecesi: 1545 n: 1546 SEM: $1.30/\sqrt{1545} = 0.033$ $\%95GA = 18.25 \pm 1.965 \times 0.033$ $\%95GA = 18.25 \pm 0.065 = 18.185-18.315$
---	--

	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 kesiştiğinden arada fark olmadığı sonucunu çıkarabiliriz.

NOT: Bu problemi daha sonra göreceğimiz farkların güven aralığı hesabı ile de çözüyoruz.

Aliştirma 7 (Sayfa 125)

1.

Araştırmadaki kümeler;5 sağlık hizmeti organizasyonu ;3 medikal grup,2 medikal plans Kümeye içindeki bireylerin birbirine benzeme ihtimali nedeniyle vaka sayısı görünenden az olacağından küme örnekleme yapılan çalışmalarda örnekleme sayısını büyük tutmak zorundayız. Örnekleme hesabında küme katsayısı dikkate alınmalıdır.

2.

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.

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 örnekleme 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 örnekleme 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.

Aliřtırma 8 (Sayfa 139)

Arařtırma 1: Bir kohort alıřmanın sonuları.

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

Arařtırma 2: Bir vaka kontrol alıřmanın sonuları

		Risk		Toplam
		Var	Yok	
Hastalık	Var	600	400	1000
	Yok	250	1750	2000
	Toplam	650	2350	3000

Sorular

1. Arařtırma 1 iin relatif riski hesaplayınız.
2. Arařtırma 2 iin odds oranını hesaplayınız
3. Arařtırma trleri belirtilmemiř olsaydı Arařtırma 1'in kohort, Arařtırma 2'nin ise vaka kontrol alıřması olduėunu nasıl ayırt ederdiniz?
4. Arařtırma 2 iin odds oranının %95 gven aralıėı **[8,73 – 12,60]** olarak hesaplanmıřtır. Bunu yorumlayınız. Gven aralıėı **[0,02-12,60]** olsaydı yorumunuz nasıl olurdu?

1.

Arařtırma 1 iin rlatif risk = $(100/8100)/(10/15010) = 18,5$

2.

Arařtırma 2 iin odds oranı = $600*1750/400*250 = 10,5$

3.

Birinci arařtırma ile ikinci arařtırmanın vaka sayılarına bakarsak;

I.Arařtırma 23110 vakadan 110 hasta vaka

II.Arařtırma 3000 vakadan 1000 hasta vaka

Vaka-kontrol çalışmalarını nadir görülen hastalıklarda yapıldığı için 23bin gibi sayıca çok vakayı 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.

4.

Güven aralığını şu şekilde yorumlayabiliriz: Bu araştırma 100 kez tekrarlırsa 95’inde odds oranı 8,7 ile 12,6 arasında çıkacaktır.

Güven aralığı 0,02-12,60 aralığında olsaydı 1 den küçük rakam olduğu için “riskler açısından fark yoktur, anlamlı değil” derdik. Bir hastalık için riskin anlamlı olması için birden büyük olması gerekir.

Not: Ek bilgi olarak odds oranı için güven aralığı formülünü aşağıda veriyoruz:

OR Güven Aralığı Formülü:

$$\ln(\text{OR}) \text{ için \%95 GA} = \ln(\text{OR}) \pm 1.96\left(\frac{1}{A} + \frac{1}{B} + \frac{1}{C} + \frac{1}{D}\right)^{0.5}$$

$$= \ln(\text{OR}) \pm 1.96\left(\frac{1}{600} + \frac{1}{400} + \frac{1}{250} + \frac{1}{1750}\right)^{0.5}$$

$$= \ln(\text{OR}) \pm 1.96(0,0087)^{0.5}$$

$$= \ln(\text{OR}) \pm 1.96 \times 0,0934$$

$$= \ln(10,5) \pm 0,1832$$

$$= 2,3513 \pm 0,1832$$

$$= [2,1681 ; 2,5345]$$

$$\begin{aligned} \text{OR için \%95 GA} &= e^{2,1681} ; e^{2,5345} \\ &= 2,718^{2,1681} ; 2,718^{2,5345} \end{aligned}$$

$$\text{OR için \%95 GA} = [8,73 - 12,60]$$

Alıştırma 9 (Sayfa 155)

1.

İstatistiksel anlamlılık düzeyi p olarak gösterilen alfa hatası genelde kaç olarak alınır ve neyi ifade eder?

P değeri genelde %5 olarak alınır ve tip 1 hatayı yani H_0 'ın doğru olduğu halde reddedilme olasılığını ifade eder.

2.

Tip 1 ve Tip 2 hata arasında nasıl bir ilişki vardır?

Tip 1 hata ihtimalini azaltırsak, tip 2 hata yapma ihtimalini artırırız yani ters ilişki vardır.

3.

Bir araştırmanın gücüne neler etki eder?

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.

Örneklem sayısını gerektiğinden çok yüksek tutarak yanıltıcı bir sonuç elde edebilir miyiz?

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.

Alıştırma 10 (Sayfa 164)

Bu alıřtırmada daha önceki konuda iřlediđimiz arařtırmanın gücü (power) ve buna etki eden faktörleri tekrar gözden geçirmek istiyoruz.

Bu durumda arařtırmanın gücünü řu řekilde tanımlayabiliriz: “*arařtırmanın gücü, dođru olduđu durumda sıfır hipotezini kabul edebilme olasılıđımızdır*”.

Ařađıdaki bađlantıda verilen “Quality of life of people with diabetes attending primary care health centres in Riyadh: bad control - good quality?” bařlıklı çalıřmayı inceleyiniz:

<http://www.ingentaconnect.com/content/rmp/qpc/2007/00000015/00000005/art00009;jsessionid=1bxrilhxp9s.victoria>

Makalenin yöntem kısmı ařađıya çıkarılmıřtır:

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.

1. Araştırmanın ana sonuç ölçütü nedir?

Quality of life of people with Diabetes = Diyabetli insanların yaşam kalitesi

2. Araştırmanın hipotezi nedir? H_0 ve H_1 'i kurunuz.

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 örnekleme kaç kişidir?

Araştırmanın örnekleme 440 kişidir. (Data could be obtained for 440 patients)

4. Ana sonuç ölçütüyle ilgili gözlemlerin variabilitesi nedir?

Bu dağılımın variabilitesi ortalamasının standart sapması veya varyansı ile hesaplanır.

5. Effect of interest (comparison of interest, true difference of means, effect of interest veya treatment effect (önemlilik derecesi, saptamak istediğimiz en küçük fark, etki genişliği, kaç olarak alınmıştır?)

Bu araştırma için **effect of interest %7** alınmıştır.

6. İstatistiksel anlamlılık düzeyi kaç olarak alınmıştır?

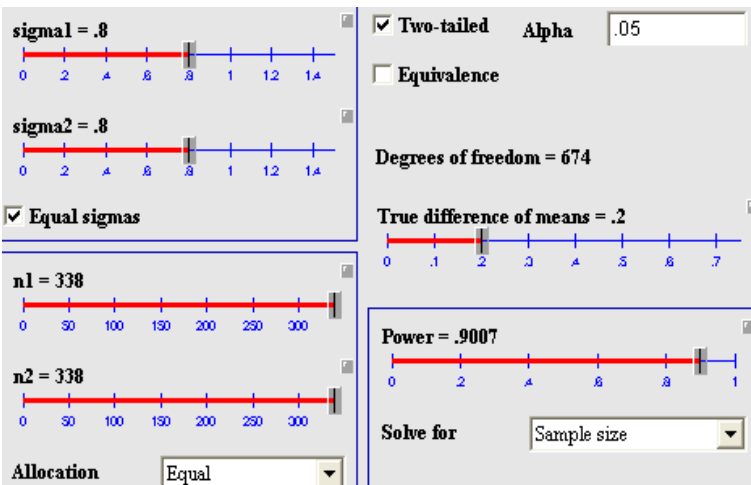
A P value of less than **0.05** was regarded as statistically significant.

7. Araştırmanın gücü kaçtır?

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. 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). Yöntemde belirtilen, eski araştırmalara bakarak "mean physical fitness score" 3 alındığına ve bunun %7'sinin true difference of means'ı oluşturacağına dikkat ediniz!



Alıştırma 11 (Sayfa 255)

1.
Bağımlı değişken menapoz 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.
Sensitivitesi: %88.9
Spesivitesi: %94.9
Doğru sınıflandırma oranı:(32+75)/115*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. Menapozu 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 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 yaşı bağımlı değişkene etkisinin istatistiksel olarak anlamlı olmadığı sonucunu çıkarabiliriz.

Alıştırma 12 (Sayfa 256)

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 bir çok 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 dahil edilmemesi 19-60 yaş grubunda kadın ve erkeklerde ayrı ayrı modeller oluşturulmasıyla açıklanabilir.

Aksi halde cinsiyet PEFR değerlerine etki edecek önemli bir değişken olduğundan dahil 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.

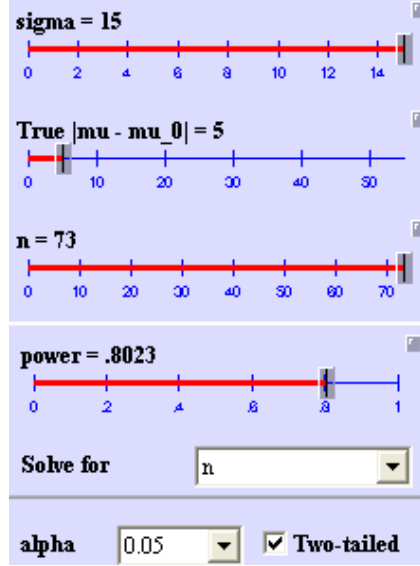
5.

PEFR ye en çok etki eden değişken yaştır. Bu modele göre yaştaki bir birimlik değişim PEFR ye 9.7 kat etki edecektir.

Alıştırma 13 (Sayfa 263)

1

- H_0 Erişkin bireylerin sistolik kan basıncı ortalamaları toplumunkinden farklıdır.
- Kesitsel araştırma (Bakınız s. 104)
- Tek örneklemede t testi (Bakınız s. 157)
- 73 kişi



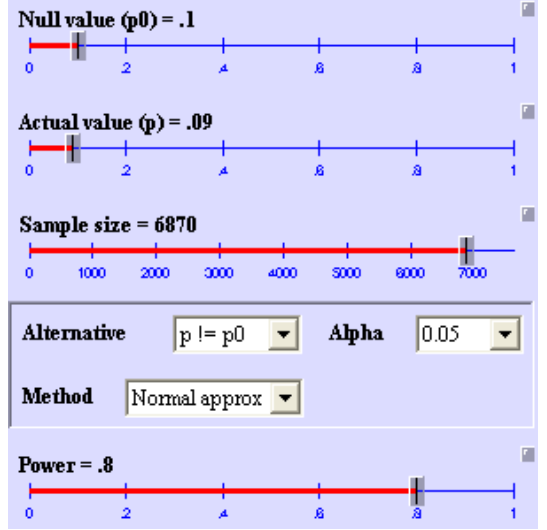
2

- H_0 : Mahalledeki erişkin erkek ve kadınların kan basınçları arasında fark yoktur.
- Kesitsel araştırma (Bakınız s. 104)
- Bağımsız örneklemlerde t testi (Bakınız s. 179)
- Toplam 164 kişi. 82 kadın, 82 erkek.



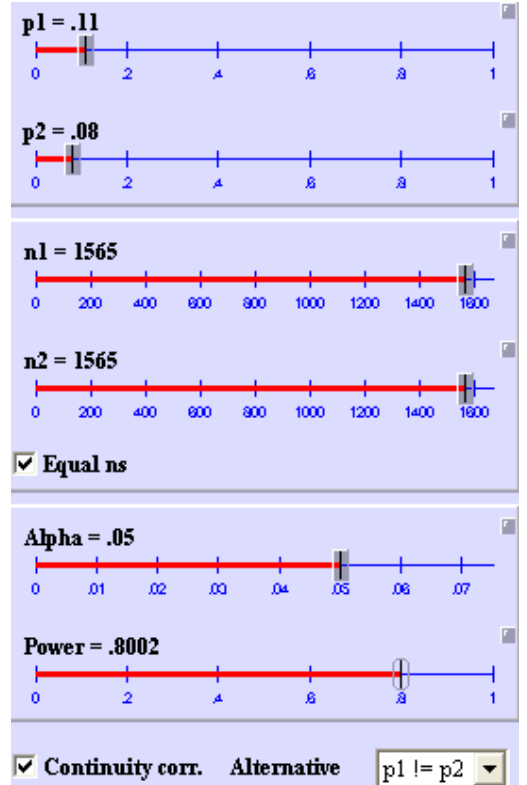
3

- H_0 : Mahallemizdeki şeker hastalığı yaygınlığı toplumdakinden farklı değildir.
- Kesitsel araştırma (Bakınız s. 104)
- z testi (Bakınız s. 209)
- 6870 kişi (Toplumdaki birey sayısı (evren) bilindiğinde daha düşük olacaktır. Bu durumda örneklem hesabı formülü değişecektir. Bakınız s. 258)



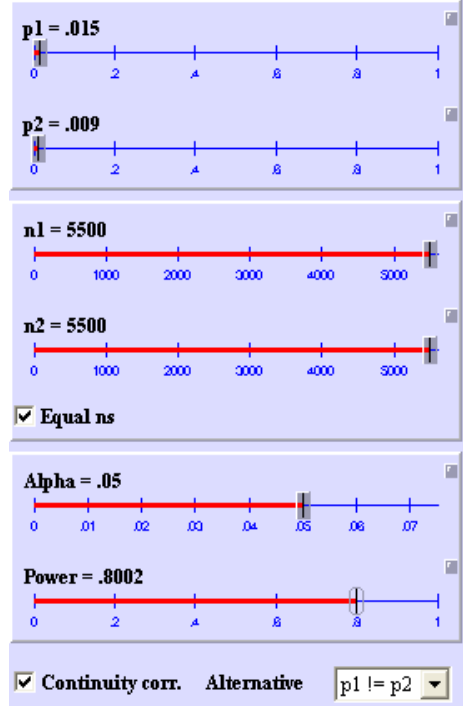
4

- H_0 : Mahalledeki erkekler ve kadınlar arasında şeker hastalığı yaygınlığı eşittir.
- Kesitsel araştırma (Bakınız s. 104)
- Ki kare testi (Bakınız s. 219)
- 3130 kişi. 1565 kadın, 1565 erkek.



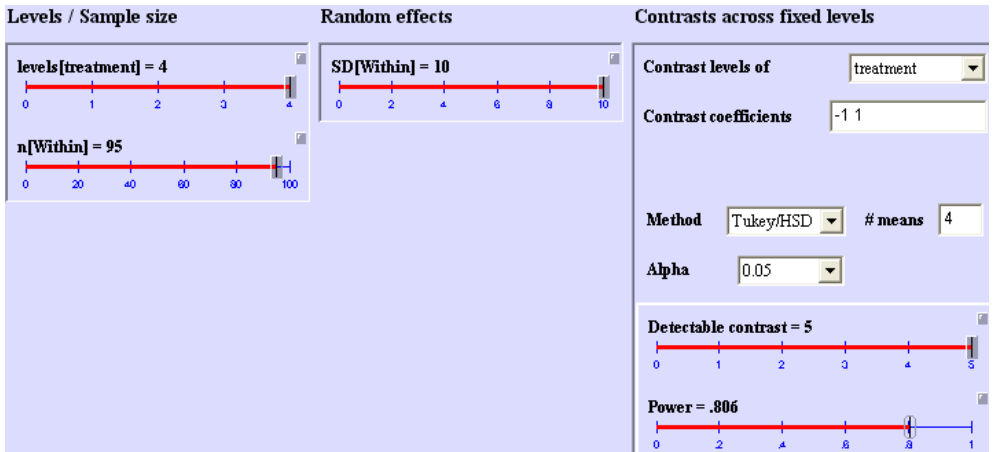
5

- H_0 : İnme riski açısından A ve B ilaçları arasında fark yoktur.
- Vaka kontrol çalışması (Bakınız s. 135)
- Ki kare testi (Bakınız s. 219)
- 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. 115)
- Tek yönlü varyans analizi (ANOVA) (Bakınız s. 199)
- 380 kişi. Her grupta 95 kişi.



Kaynaklar

1. SPSS. Training. http://www.spss.com/training/?source=homepage&hpzone=nav_bar
Elektronik Erişim 20.07.2010
2. SPSS Türkiye. <http://www.spss.com.tr/> Elektronik Erişim 20.06.2010
3. Lenth R. Java applets for power and sample size.
<http://www.stat.uiowa.edu/~rlenth/Power/> Elektronik Erişim 20.06.2010
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. Blackwell Publishing, London 2006.
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. Kaan Kitabevi, Eskişehir 1997
12. Sümbüloğlu K, Sümbüloğlu V. Biyoistatistik Uygulama Kitabı. Hatiboğlu Yayınevi, Ankara 1991.

İndeks

A

Ağırlıklı ortalama, 58
Akış şemaları, 29
Alternatif hipotez, 141
Ampirik frekans dağılımı, 49, 71
Anlamlılık düzeyi (p), 143
ANOVA, 15, 23, 78, 147, 150, 199, 200, 201, 202, 203, 204, 205
Aralık (range), 63
Araştırma fikri, 18
Araştırma panosu, 19
Araştırma protokolü, 117
Araştırma Tasarımı, 103, 111
Araştırma yöntemleri, 21
Araştırmanın çıktıları, 116
Aritmetik ortalama, 55
Asansör grafiği.*Bakınız* Saplı kutu
ASCII, 35
Aşırı değer, 46
Ayırma hataları, 116

B

Bağımlı örneklerde t-testi, 165
Bağımsız örneklerde t testi, 147, 151, 179
Bar grafik, 49
Binomiyal dağılım, 81
Birden çok hipotezi test etme, 150
Birey içi farklılık, 67
Bireyler arası farklılık, 67
Blokla, 111, 116
Bonferroni, 150, 204
Boş verilerin kodlanması, 38
Boxplot.*Bakınız* saplı kutu

C

Cevaplanabilir bir klinik soru sormak.*Bakınız* PICO
Cochran's Q, 233, 234
SPSS ile yapılması, 233
varsayımları, 233
Coefficient of variation.*See* varyans katsayısı
Confounding.*Bakınız* Karıştırıcı faktörler
CONSORT bildirisi, 118

Ç

Çıkarılma kriterleri, 117
Çok seçenekli soruların SPSS'e girilmesi, 37, *Bakın* SPSS'e veri girişi

D

Dahil edilme kriterleri, 117
Değişken, 31
Deneysel araştırmalar, 104
Dikotom, 32, 209
Dot plot, 49, 53
Dörtlük çeyrek değerler genişliği, 46, 64
Durdurulmuş (censored) veri, 34

E

Eksik veriler, 44
Endpoints.*Bakınız* Araştırmanın çıktıları
Eşdeğerlik (non inferiority), 142
Etik, 117
Etki genişliği, 147, 257, 298
Evren, 31

F

Faktöriyel tasarımlar, 112
Faz, 115
Frekans dağılımı, 49
Frekans dağılımının şekli, 54

G

Gauss dağılımı.*Bakınız* Normal dağılım
Geometrik ortalama, 56
Gözlemsel araştırmalar, 104
Güç ve ilişkili faktörler, 149
Güven aralığı
orantılar için, 97
Güven Aralıkları, 95

H

Hata ayıklama, 43, 53
Hatalar (bias), 107
Hız, 34

Hipotez, 24

Hipotez kurma, 17, 20

Hipotez Testleri, 141

Hipotez Testlerinde Hatalar, 147

Histogram, 49, 52

İ

Intention-to-treat, 117

İki yönlü ve tek yönlü hipotez, 142

İstatistiksel testler ve kullanım yerleri, 29

İstatistik tablolar, 279

işaret testi, 157, 162, 172

K

Kare dönüşümü, 90

Karekök dönüşümü, 88

Karıştırıcı faktörler, 107

Kategorik, 31

Kategorik Veri İki Bağımsız Grup, 3, 220

Kategorik Veri İki Bağımlı Grup, 3

Kategorik Veri İki Bağımlı Kategori, 3, 233, 237

Kategorik Veri İki Bağımsız Grup, 225

Kategorik Veri İki Bağımlı Kategori, 239

Kategorik Veri İki Bağımlı Grup, 3

Kategorik veriler, 36, 124

Kendini Değerlendirme, 14, 265

Ki Kare, 207, 215, 216, 218, 220, 221, 222, 223, 225, 226, 230, 231, 235

Klinik Araştırmalar, 115

Klinik kohortlar, 130

Kohort Araştırmaları, 22, 127

Kolmogorov Smirnov, 85, 87, 182

Kontrol grupları, 107

Korelasyon analizi, 3, 241

Körleme, 117

Kruskal Wallis Testi, 204

L

Levene, 88, 185, 202, 205

Literatür tarama, 20

Logaritmik dönüşüm, 84

Logit (logistic) dönüşüm, 90

Lojistik regresyon analizi, 3, 37, 247

Longitudinal araştırmalar, 106

M

Mann-Whitney U testi, 23, 179, 186, 187

Marginal homogeneity test, 237

SPSS ile yapılması, 237
varsayımları, 237

McNemar testi, 215

MedCalc™, 211, 258

Merkezi dağılım ölçütleri, 55

MINITAB, 150, 154, 211, 258

Mod, 56

N

Nedensellik, 106

Nokta prevalansı, 106

Nominal, 32

Nonparametrik testler, 23

Normal Dağılım, 69

Numerik, 31

Numerik Veri İki Bağımlı Grup, 3, 165

Numerik Veri İki Bağımsız Grup, 3, 179

Numerik Veri İki Bağımlı Grup, 3, 189

Numerik Veri Tek Grup, 3, 157

Numerik veriler, 37, 100, 124

O

Odds oranı, 101, 137, 248, 255

Odds oranı için güven aralığı, 101, 295

Olasılık kuralları

Çarpma kuralı, 70

Toplama kuralı, 70

Olasılık yoğunluk fonksiyonu, 142, 143

On-treatment, 117

Oran, 33

Orantı, 34

Orantının örneklem dağılımı, 93

Ordinal, 32

Ortalamanın örneklem dağılımı, 92

Ortalamanın standart hatası, 92, 95, 124

Ortanca, 56

Outcomes.*Bakınız* Araştırmanın çıktıları

Ö

Öğrenim hedefleri, 14

Önemlilik testleri, 23

Örneklem Hesabı, 257

Örneklem varyasyonu, 91

Örnekleme, 91

kota örnekleme, 91

küme örnekleme, 91

sistemik, 91

Standart sapma, 64

Student t test.*Bakınız* Bağımsız örneklemlerde t testi

Sütun grafikleri, 49

P

p.*Bakınız* Anlamlılık düzeyi

Paralel ve çapraz çalışmalar, 112

Parametrik olmayan testler, 144

Parametrik testler, 23

varsayımları, 23

PASS, 258, 261

Pasta grafiği, 49, 51

Pearson korelasyon katsayısı, 241

Persantil aralıkları, 63

PICO, 20, 103, 116, 124

Pie graph.*Bakın* Pasta grafiği

Poisson dağılımı, 81

Power, 111, 148, 150, 164, 297, 298

R

Randomize Kontrollü Çalışma, 22

Referans aralığı, 64

Relatif risk, 248

Relative Deviate, 95, 102

Replikasyon, 111

Resiprok dönüşüm, 90

Retrospektif kohort, 127

Russ Lenth, 151, 164, 258, 298

S

Saçılma grafiği, 49, 53, 291

Sağa eğimli, 54

Sağlıklı birey etkisi, 127

Saplı kutu grafiği, 49, 53

Saplı kutu grafikleri, 45, 266

Scatter diagram.*Bakınız* saçılma grafiği

SEM.*Bakınız* Ortalamanın standart hatası

Serbestlik derecesi, 98

Sıfır hipotezi, 141

Skewness, 85, 87, 97, 158, 179

Sola eğimli, 54

Spearman korelasyon katsayısı, 245

SPSS'e veri girişi, 35

Standart Normal Dağılım, 73

Standart Normal Sapma, 73

T

Tabakalı randomizasyon, 116

Tek orantı testi, 209

Tek örnekleme t-testi, 157

Tekrarlayan kesitsel araştırmalar, 106

Teorik dağılımlar

F dağılımı, 78

Ki kare dağılımı, 78

Lognormal dağılım, 78

t-dağılımı, 77

Tip I hata, 148

Tip II hata, 148

Toplum, 31, 64, 74, 91, 92, 93, 96, 104, 121, 157, 159, 160, 163, 166, 169, 171, 172, 209, 211

U

Uç değerler, 43, 45, 46, 48, 53, 122, 266

V

Vaka Kontrol, 22, 135

Varyans, 64

Varyans katsayısı, 65

Varyasyon, 111

Veri çeşitleri, 23, 31

Veri Dönüştürme, 83

Veri giriş yöntemleri, 35

Veri girişi.*Bakın* SPSS'e veri girişi

Verilerde bulunması gereken özellikler, 23

Verilerin grafiklerle gösterilmesi, 49

W

Wilcoxon işaretli sıra testi, 165, 172, 173

Y

Yaygınlık ölçütleri, 63

Yöntem, 24

Yüzde, 33

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.

İstatistikten korkmanıza gerek yok!

