# Numerik Veri İki Bağımsız Grup

Amaç: Bu konu sonunda okuyucunun iki bağımsız gruptan elde edilen numerik verilerin ortalamalarının karşılaştırılmasında kullanılan bağımsız örneklemlerde *t*-testi (independent samples *t*-test veya Student *t* testi) ve bunun nonparametrik alternati olan Mann-Whitney U testi hakkında bilgi sahibi olması ve bu testleri yapabilmesi amaçlanmıştır.

Hedefler:

* Numerik veri, iki bağımsız grubu tanımlayabilmeli
* Student *t* testi varsayımlarını sayabilmeli
* Boy ve ağırlık değişkenlerini kullanarak SPSS ile vücut kitle endeksini hesaplayabilmeli
* 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 istediğimizi varsayalım (*H0*: Diyabet hastası erkekler ve bayanlar arasında vücut kitle indeksleri açısından fark yoktur). Ölçmek istediğimiz değişken (VKİ) numerik 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 Mann-Whitney U testini yapabiliriz.

## Bağımsız örneklermlerde t testi

### Varsayımları

* Değişken, toplumda normal dağılmalı
* Değişkenin varyansları her iki grupta eşit olmalı
* Örneklem sayısı normal dağılımı ve varyansların eşitliğini test edebilecek kadar yeterli olmalı.

### Mantığı

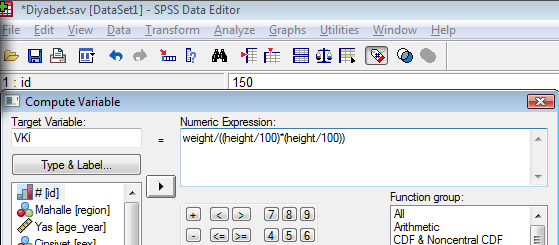
İki grupta yapılan ölçümlerin ortalamasını karşılaştırıyoruz. Sıfır hipotezimize göre iki grubun ortalamaları arasında fark olmaması gerekir (Ortalama1 – Ortalama2 = 0). Student *t* testi daha önce gördüğümüz tek örneklemde *t* testi ve bağımlı örneklemlerde *t* testi gibi ortalamalar arasındaki farka dayanmaktadır. Bu testin verileri teorik dağılımlarından *t* dağılımına benzemektedir.

### Not

İki grubumuzun örneklemleri *n1* ve *n2*, ortalamaları  ve , standart sapmaları ise *s1* ve *s2*’dir.

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

*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.

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.

1. **Verilerin toplanması:**

Verilerimizi topladık ve SPSS’e girdik.

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

Student *t* testi için kullanacağımız formül de önceki *t* testi uygulamalarına benzemektedir:



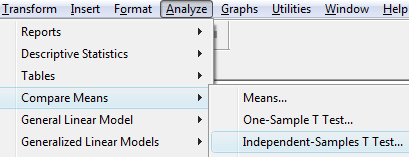
Burada *s1* ve *s2*’den “pooled standard deviation” hesaplanması gerekmektedir:



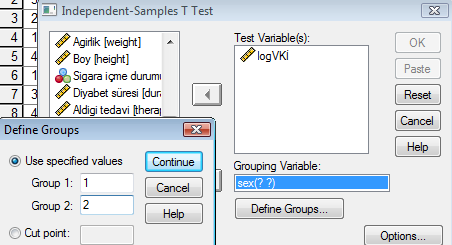
|  |  |
| --- | --- |
| (Erkek) | 27,5 |
|  | 30,4 |
| *s1* | 3,3 |
| *s2* | 4,0 |
| *n1* | 107 |
| *n2* | 65 |

Yukarıdaki formülü hesap makinesiyle yapmak zordur. Doğrudan SPSS ile uygulamaya geçilecektir:

*Analyze>Compare Means>Independent-Samples T Test>[“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:

**Test istatistiğinden elde edilen değerin bilinen bir olasılık dağılımı ile karşılaştırılması:**

İkinci tablodan görüldüğü gibi Levene testine göre varyanslar arasında fark yoktur (Sig. = 0,054; >%5). *t* istatistiği **-5,12** olarak hesaplanmıştır. Serbestlik derecesi (*n1*+*n2*-2) = 170 olarak alınmıştır. İki yönlü *p* değeri sıfıra yakındır (p<0.001).

1. ***P değerinin* ve sonuçların yorumlanması**

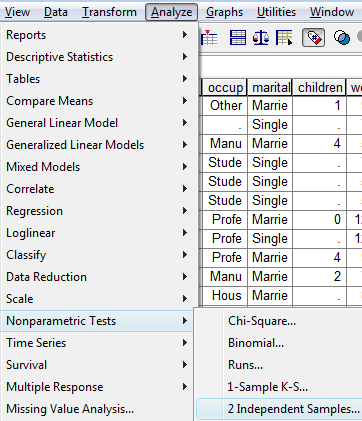
*P* değeri 0,05’ten küçük olduğu için *H0* hipotezini reddederiz. Dolayısıyla diyabetik erkeklerin VKİ değerlerinin diyabetik bayanlarınkinden daha düşük olduğu sonucuna varabiliriz.

## 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 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:

Mann-Whitney U testi sonucunda da *z* testi değeri **-4,44** ve *p* değeri **<0,001** olduğundan H0 hipotezini reddederiz.

**NPar Tests**

**Mann-Whitney Test**

**Ranks**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Sex of the patient | N | Mean Rank | Sum of Ranks |
| VKİ | Male | 107 | 73,37 | 7850,50 |
| Female | 65 | 108,12 | 7027,50 |
| Total | 172 |  |  |

**Test Statistics(a)**

|  |  |
| --- | --- |
|  | VKİ |
| Mann-Whitney U | 2072,500 |
| Wilcoxon W | 7850,500 |
| Z | -4,437 |
| Asymp. Sig. (2-tailed) | ,000 |

a Grouping Variable: Sex of the patient