# Numerik Veri İki Bağımlı Grup

Amaç: Bu konu sonunda okuyucunun iki bağımlı gruptan elde edilen numerik verilerin ortalamalarının karşılaştırılmasında kullanılan bağımlı örneklemlerde *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:

* Bağımlı örneklemlerde *t*-testi varsayımlarını açıklayabilmek
* SPSS ile bağımlı örneklemlerde *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 uyguladığımızı varsayalım. Program sonunda katılımcıların kilo verip vermediklerini araştırmak istiyoruz (*H0*: Bireylerin uygulanan beslenme ve egzersiz programı öncesi ve sonrasındaki vücut ağırlıkları arasında fark yoktur). Ölçmek istediğimiz değişken numeriktir. 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 ötekinin de buna yakın bir değer olacağını tahmin edebiliriz.

Bizim örneğimizde bağımlı iki numerik ölçüm söz konusudur. Kitabın başında da verilen istatistik testlerin uygulama yerleri akış şemalarına baktığımızda seçeceğimiz testin bağımlı örneklemlerde *t*-testi veya Wilcoxon işaretli sıra testi olacağını görürüz.

## Bağımlı örneklemlerde t-testi varsayımları

* Değişken, toplumda normal dağılmalı
* Örneklem sayısı yeterli olmalı.

## Testin mantığı

Eğer tedavi öncesi ve sonrasında bireylerin değerleri arasında fark yoksa, her bireyin önceki ve sonraki ölçümleri arasındaki farkların ortalamasını aldığımızda sıfır olmalıdır. Bu durumda testimiz, daha önce gördüğümüz tek örneklemde t-testine indirgenebilir. Değişkenimiz, farkların ortalaması, toplum ortalaması ise sıfırdır.

## Ek açıklama

Veriler bağımlı (eşleştirilmiş) olduğundan her iki ölçümde de örneklem sayısı eşit olmalıdır. *n* adet bireydeki iki ölçüm arasında *n* adet fark olacaktır. Bunların aritmetik ortalaması  ve tahmin edilen standart sapması *sd*’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 dağılıma uyduğunu görüyoruz:



|  |  |  |  |
| --- | --- | --- | --- |
|  | N | Ortalama | Standart Sapma |
| Diyet öncesi ağırlık | 174 | 74,65 | 8,65 |



|  |  |  |  |
| --- | --- | --- | --- |
|  | N | Ortalama | Standart Sapma |
| Diyet sonrası ağırlık | 174 | 71,03 | 8,10 |

Hipotez testleri için 5 basamaklı genel yaklaşımımızı uygulayacak olursak:

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

*H0*: Bireylerin uygulanan beslenme ve egzersiz programı öncesi ve sonrasındaki vücut ağırlıkları arasında fark yoktur

*H1*: Bireylerin uygulanan beslenme ve egzersiz programı öncesi ve sonrasındaki vücut ağırlıkları arasında fark vardır.

1. **Verilerin toplanması:**

Verilerimizi topladık ve SPSS’e uygun kodlama yaparak girdik. Verilerde hata ayıklamasını yaptık.

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

Bağımlı örneklemlerde *t* testi için kullanacağımız formül tek örneklemde *t* testindekine benzemektedir:

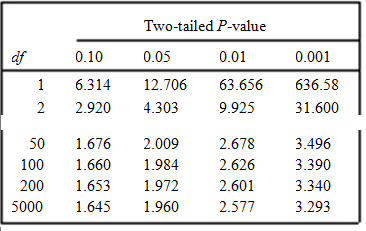


|  |  |
| --- | --- |
| = öncesi ve sonrasında yapılan ölçümler arasındaki farkın aritmetik ortalaması | 3,63 |
| 0 = toplum ortalaması (aradaki farkın sıfır olduğunu varsayıyoruz) | 0 |
| sd = öncesi ve sonrasında yapılan ölçümler arasındaki farkın standart sapması | 1,77 |
| *n* = örmeklem sayısı. (Bu değişken için mevcut veriler) | 174 |

Buradan ***t* = 26,98** olarak hesaplanır.

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

Sonuçlarımız (n-1) serbestlik derecesinde *t* dağılımı gösterir. *t* tablosu ile karlışaştıracak olursak *t* = 26,98ve serbestlik derecesi = 173 için p değerinin 0,01’den küçük olduğunu görürüz.



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

*P* değeri 0,05’ten küçük olduğu için *H0* hipotezini reddederiz. Dolayısıyla bireylerin tedavi öncesi ve sonrasında yapılan ağırlık ölçümleri arasında istatistiksel olarak anlamlı bir fark vardır.

## Örneğin SPSS ile yapılması

SPSS ile yapacak olursak:

*Analyze>Compare Means >Paired-Samples T Test>[“Weight” ve “Weight2” değişkenlerini birlikte işaretleyiğ “Paired Variables” alanına geçirelim]>ok*

**T-Test**

**Paired Samples Statistics**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | Mean | N | Std. Deviation | Std. Error Mean |
| Pair 1 | Diyet öncesi ağırlık | 74,656 | 174 | 8,6509 | ,6558 |
| Diyet sonrası ağırlık | 71,0264 | 174 | 8,09910 | ,61399 |

**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 | Diyet öncesi ağırlık - Diyet sonrası ağırlık | 3,6299 | 1,77452 | ,13453 | 3,3644 | 3,8954 | 26,983 | 173 | ,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 testinin mantığı

Bir önceki konuda gördüğümüz gibi, işaret testi (sign test) örneklemimizden elde ettiğimiz ortalamayı belli bir değerle (toplum ortalaması) karşılaştırmamızı sağlıyordu. İşaret testini burada da kullanabiliriz. Eşleştirilmiş örneklerdeki önce ve sonraki ölçümler arasındaki farkın sıfır olduğunu varsayarsak, hesabımız tek örneklemdeki t-testine benzeyecektir.

İşaret testine göre Wilcoxon işaretli sıra testi daha güçlü bir testtir. Bu test sadece farklar arasındaki işarete (artı mı eksi mi olduğuna) bakmayıp, farkın miktarını da dikkate alır.

Her iki eşleşmiş veri arasındaki fark hesaplanır. Farkı sıfır olan veriler dikkate alınmaz. Daha sonra bu farklar artı (+) veya eksi (-) olmalarına göre gruplandırılır. Ayrıca, elde edilen veriler işaretine bakılmaksızın (mutlak değerine göre) küçükten büyüğe doğru sıraya konur. En küçük farka 1, sıradakine 2 vs. numara verilir. En büyük farkın sırasına n' denir. İki veya daha fazla farkın eşit olması halinde sıralarının ortalaması alınır (örn. 0,4 – 1,2 – 1,2 – 1,2 – 3,5 için sıralama 1 – 3 – 3 – 3 – 5 şeklinde olur). Sıfır hipotezinin doğru olabilmesi için artı ve eksi farkların sıralamaların toplamı sıfır olmalıdır.

## Bir örnek üzerinden Wilcoxon işaretli sıra testi

Yukarıdaki örneğe benzer bir çalışma yaptığımızı ama daha küçük bir örneklemden veri topladığımızı düşünelim:

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

*H0*: Bireylerin uygulanan beslenme ve egzersiz programı öncesi ve sonrasındaki vücut ağırlıkları arasında fark yoktur

*H1*: Bireylerin uygulanan beslenme ve egzersiz programı öncesi ve sonrasındaki vücut ağırlıkları arasında fark vardır.

1. **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 |

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

Örneklemimizde 32 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 | 76,0 | 64,60 | 11,4 |
| 28 | 94,0 | 79,90 | 14,1 |
| 29 | 91,5 | 77,78 | 13,72 |
| 30 | 61,5 | 52,28 | 9,22 |
| 31 | 54 | 65 | **-11** |
| 32 | 88,0 | 74,80 | 13,2 |

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

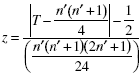
|  |  |
| --- | --- |
| **Öncesi-Sonrası** | **Sıra no** |
| **-0,5** | 1 |
| **-1** | 3 |
| **-1** | 3 |
| **-1** | 3 |
| 9,22 | 5 |
| 9,75 | 6,5 |
| 9,75 | 6,5 |
| 10 | 8 |
| 10,51 | 9 |
| **-10,8** | 10 |
| **-11** | 11,5 |
| 10,95 | 11,5 |
| 10,95 | 13,5 |
| 11 | 13,5 |
| 11,17 | 15 |
| 11,32 | 16 |
| 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 | n'=28 |

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

|  |  |  |
| --- | --- | --- |
| T+ = | 5+6,5+6,5+8+9+11,5+11,5+13,5+15+16+17+18,5+18,5+  20+21+22+23+24+25+26+27+28 | =**372,5** |
| T-= | 1+3+3+3+10+13,5 | = **33,5** |

n' ≤ 25 olması durumunda T testi değeri T+ veya T-’den hangisi küçükse o olur.

n' > 25 olması durumunda *z* hesabı yapılır



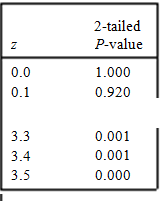
Bizim örneğimizde n' = 28 değeri 10’dan büyük olduğuna göre *z* formülünü uygulamalıyız:

***z*** = [│33,5- ((28x29)/4)│-(1/2)] / [((28x29)x(56+1))/24] = **-3,86.**

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

n' ≤ 25 olması durumunda T testi değeri Wilcoxon istatistik değerler tablosuyla karşılaştırılır

n' > 25 olması durumunda *z* tablo değeri ile karşılaştırılır.



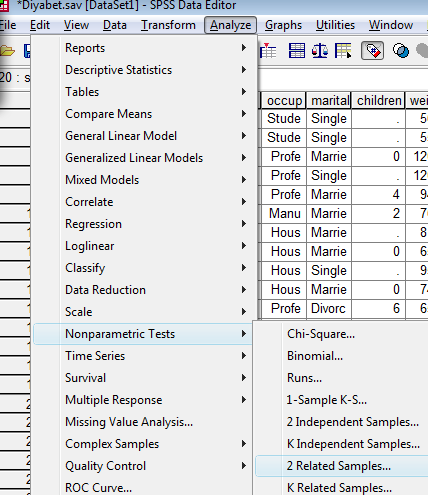
*z* tablosundan *z* = -3,86 için iki yönlü p değerinin 0,001 den küçük olduğunu görüyoruz.

1. ***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>[“weight3” ve “weight4” değişkenlerini işaretleyip “Test Pair(s) List” alanına geçirelim>ok.*



SPSS çıktısı aşağıda

**NPar Tests**

**Wilcoxon Signed Ranks Test**

**Ranks**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | N | Mean Rank | Sum of Ranks |
| VAR00002 - VAR00001 | Negative Ranks | 22(a) | 16,93 | 372,50 |
| Positive Ranks | 6(b) | 5,58 | 33,50 |
| Ties | 3(c) |  |  |
| Total | 31 |  |  |

a VAR00002 < VAR00001

b VAR00002 > VAR00001

c VAR00002 = VAR00001

**Test Statistics(b)**

|  |  |
| --- | --- |
|  | VAR00002 - VAR00001 |
| Z | -3,861(a) |
| Asymp. Sig. (2-tailed) | ,000 |

a Based on positive ranks.

b Wilcoxon Signed Ranks Test

**NPar Tests**

**Wilcoxon Signed Ranks Test**

**Ranks**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | N | Mean Rank | Sum of Ranks |
| Diyet sonrası ağırlık - Diyet öncesi ağırlık | Negative Ranks | 22(a) | 16,93 | 372,50 |
| Positive Ranks | 6(b) | 5,58 | 33,50 |
| Ties | 3(c) |  |  |
| Total | 31 |  |  |

a Diyet sonrası ağırlık < Diyet öncesi ağırlık

b Diyet sonrası ağırlık > Diyet öncesi ağırlık

c Diyet sonrası ağırlık = Diyet öncesi ağırlık

**Test Statistics(b)**

|  |  |
| --- | --- |
|  | Diyet sonrası ağırlık - Diyet öncesi ağırlık |
| Z | -3,861(a) |
| Asymp. Sig. (2-tailed) | ,000 |

a Based on positive ranks.

b Wilcoxon Signed Ranks Test