

## Ders 6

Nihat Tak

2023-04-6

### BAĞLANTILI İKİ ÖRNEKLEM İÇİN İŞARET TESTİ

#### Çift Örneklem Testleri

Çift örneklem testleri, iki örneklemen aynı kütleden veya parametreleri ve dağılımları aynı olan iki ana kütleden çekilmiş rassal örneklemeler olup-olmadığını belirlemede kullanılan istatistik teknikleridir.

#### Bağlantılı Çift Örneklem Testleri

İki örneklemen belli bir değişken bakımından farklı olup olmadığı araştırılmak isteniyorsa, bu değişkenin ölçümü daima aynı konum ve koşul altında yapılmalıdır. Böylece incelenen değişken bakımından belirlenecek farklılıklar, yalnızca bu farklılığa neden olan faktöre bağlı olarak ortaya çıkmış sayılabilecektir. Bu amaçla denemelerde aynı birimin kullanılması önerilir. Ancak aynı birimden ölçüm yapma olanağı yok ise, örneğin ikizler, aynı yaşta, cinsiyette canlılar vb. denek olarak kullanılabilir. Cansız varlıklarda ise; aynı nitelikte, aynı ısıda aynı nem ortamında vb. deneyler yapılması gerekir.

#### İşaret Testi

- Yer parametresine ilişkin testtir.
- Bu test hiçbir şekilde tüm deneklerin aynı kitleden geldiği savını ortaya koymaz. Tüm denekler nerden gelirse gelsin önemli değildir. Tek gereksinim her çiftin ilgili dış değişken itibarı ile uyumlu hale sokulabilmesidir.

#### Varsayımları

- Asgari sıralayıcı ölçekle ölçülmüş verilerde kullanılır.
- İncelenen değişkenin sürekli olması gerekir
- n tane çift birbirinden bağımsız olmalı

#### Adımları

**1. Adım:** Hipotezler kurulur.  $D_i = X_A - X_B$

Çift yönlü testler:

$H_0: M_D = 0$  (Farkların yarısı “+” yarısının”-“ olması) ( $H_0: M_A = M_B$ )

$H_1: M_D \neq 0$

Tek yönlü testler:  $H_1: M_D < 0$  veya  $H_1: M_D > 0$

**2. adım:** Test istatistiği hesaplanır.

$H_0$  hipotezinin doğruluğu altında;

$y \sim b(n, 0.5)$  burada  $y$  "+" veya "-" gözlem sayısı

**not:** Eğer elde ettiğimiz farklardan ( $D_i$  değerleri) biri veya birkaçı sıfır(0) değeri alırsa o fark (o çift) çıkarılır. Bu durumda  $n$  tane çiftimiz varken  $n$  taneden daha az çiftimiz kalır.

Bu durumda test istatistiği;

$$p_{hes} = P(Y \leq y) = \sum_{i=0}^y \binom{n}{i} p^i (1-p)^{n-i} \text{ veya}$$

$$p_{hes} = P(Y \geq y) = \sum_{i=y}^n \binom{n}{i} p^i (1-p)^{n-i}$$

Veya binom tablosundan da  $p_{hes}$  değeri bulunabilir.

**3. adım:** Anlam düzeyi tablo değeri olarak alınır.

**4. adım:** Karar verilir. Eğer  $y$ = daha az gözlenen işaret sayısı ise

$$p_{hes} = P(Y \leq y) < \alpha/2 \text{ ise } H_0 \text{ reddedilir}$$

Eğer  $y$ = eksi işaretli gözlemlerin sayısı ise

$$p_{hes} = P(Y \leq y) < \alpha \text{ ise } H_0 \text{ reddedilir}$$

Eğer  $y$ = artı işaretli gözlemlerin sayısı ise

$$p_{hes} = P(Y \leq y) < \alpha \text{ ise } H_0 \text{ reddedilir}$$

**Örnek 1.** Çocuğun gelişiminde babanın yokluğunun doğurduğu etkileri ele alan bir çalışmada farklı şehirlerde ki işleri nedeniyle birbirlerinden ayrı kalan ve bu sırada bir çocukları olan evli çiftler anne ve baba ayrı ayrı olmak üzere görüşmeye tabi tutuluyor.

Bu kişilerin her birinden hayatının ilk senesini babasız bir evde geçiren çocukla ilgili çeşitli konular üzerine fikirler alındı. Her çiftin farklı şehirdeki işinden dönüşü izleyen yıllarda babanın çocukla ne gibi disiplin ilişkileri kurduğu üzerinde duruldu.

Bu aileleri çok iyi bilen bir psikologdan her bir anne-baba çiftinin gösterdiği disiplinin derecelendirilmesi istendi. İleri sürülen sav şu şekilde idi; Çocukla daha uzun veya daha yakın bağlantısı olması ve farklı şehirdeki işleri nedeniyle babadan ayrı oluşun ortaya çıkardığı çeşitli diğer koşullar yüzünden anne, kocasının çocukla kurduğu disiplin ilişkilerinde babadan daha anlayışlı(daha otoriter) olacaktır.Bu çalışmada ele alınan 17 aile için aşağıdaki sonuçlar elde edilmiştir. ( $\alpha = 0.05$ )

$X_A$	4	4	5	5	3	2	5	3	1	5	5	5	4	5	5	5	5
$X_B$	2	3	3	3	3	3	3	3	2	3	2	2	5	2	5	3	1

\*\*

**1. Adım:** Hipotezler kurulur.  $D_i = X_A - X_B$

$$H_0: M_D \leq 0$$

$$H_1: M_D > 0 \text{ (Anneler daha otoriter)}$$

**2. adım:** Test istatistiği hesaplanır.

$X_A$	4	4	5	5	3	2	5	3	1	5	5	5	4	5	5	5	5
$X_B$	2	3	3	3	3	3	3	3	2	3	2	2	5	2	5	3	1
$X_A - X_B$	+	+	+	+	0	-	+	0	-	+	+	+	-	+	0	+	+

11 tane artı, 3 tane eksi gözlemlendi,  $n=17$  iken  $n=14$  oldu

$$p_{hes} = P(x \geq 11) = \sum_{x=11}^{14} \binom{14}{x} 0.5^x 0.5^{14-x}$$
$$p_{hes} = 0.0279$$

```
dbinom(11,14,0.5)+dbinom(12,14,0.5)+dbinom(13,14,0.5)+dbinom(14,14,0.5)
```

```
## [1] 0.02868652
```

```
1-pbinom(10,14,0.5)
```

```
## [1] 0.02868652
```

veya

$$p_{hes} = P(x \leq 3) = \sum_{x=0}^3 \binom{14}{x} 0.5^x 0.5^{14-x}$$

```
dbinom(0,14,0.5)+dbinom(1,14,0.5)+dbinom(2,14,0.5)+dbinom(3,14,0.5)
```

```
## [1] 0.02868652
```

```
pbinom(3,14,0.5)
```

```
## [1] 0.02868652
```

**3. adım:** Anlam düzeyi tablo değeri olarak alınır.  $p_{tablo} = \alpha = 0.05$

**4. adım:** Karar verilir. Eğer  $y = +$  işaretli sayısı ise

$p_{hes} = 0.0279 < \alpha = 0.05$  olduğundan  $H_0$  reddedilir yani çocuklara karşı, anneler babalara göre daha otoriterdir.

```
if(!require(BSDA)) install.packages("BSDA")
```

```
## Zorunlu paket yükleniyor: BSDA
## Warning: package 'BSDA' was built under R version 4.2.3
## Zorunlu paket yükleniyor: lattice
##
## Attaching package: 'BSDA'
## The following object is masked from 'package:datasets':
##
##      Orange

A<-c(4,4,5,5,3,2,5,3,1,5,5,5,4,5,5,5,5)
B<-c(2,3,3,3,3,3,3,3,2,3,2,2,5,2,5,3,1)

SIGN.test(x=B,y=A, alternative = "less")

##
##  Dependent-samples Sign-Test
##
## data:  B and A
## S = 3, p-value = 0.02869
## alternative hypothesis: true median difference is less than 0
## 95 percent confidence interval:
##  -Inf      0
## sample estimates:
## median of x-y
##           -2
##
## Achieved and Interpolated Confidence Intervals:
##
##               Conf.Level L.E.pt U.E.pt
## Lower Achieved CI      0.9283  -Inf      0
## Interpolated CI       0.9500  -Inf      0
## Upper Achieved CI      0.9755  -Inf      0
```

## Büyük Örneklem Yaklaşımı

$n \geq 25$  ise binom dağılımının normale yaklaşımını kullanabiliriz.

$$\mu_y = np \text{ ve } \sigma_y^2 = npq$$

- Eğer  $y < \mu_x$  ise  $y + 0.5$  düzeltme terimi eklenir.
- Eğer  $y > \mu_x$  ise  $y - 0.5$  düzeltme terimi çıkarılır.

$$Z = \frac{(y +/-0.5) - \mu_y}{\sigma_y}$$

$$Z = \frac{(y +/-0.5) - n/2}{(1/2)\sqrt{(n)}} \sim N(0,1)$$

**Örnek 2.** Aç kalmadan ve egzersiz yapmadan kilo azaltan ürün geliştirilip ve reklamı yapılmaktadır. Bu ürünün 2 haftada kilo verdirdiği iddiası test edilmek istenmektedir. Bunun için bu ürünü kullanan 26 kişi gözlemlenmiş ve aşağıdaki sonuçlar elde edilmiştir. ( $\alpha = 0.05$ )

	Kişi	1	2	3	4	5	6	7	8	9	10	11	12	13
Ağırlık	Önce	146	175	150	190	220	157	136	146	128	187	172	138	150
	Sonra	142	178	147	187	212	160	135	138	132	187	171	135	151
	Kişi	14	15	16	17	18	19	20	21	22	23	24	25	26
Ağırlık	Önce	124	210	148	141	164	150	130	145	132	125	137	140	126
	Sonra	126	208	148	138	159	146	130	140	130	125	138	131	120

**1. Adım** Hipotezler kurulur.  $D_i = X_A - X_B$

$$H_0 = M_D \geq 0$$

$H_1 = M_D < 0$  (Yeni ürün iki hafta içinde zayıflatmaktadır)

**2. Adım** Test istatistiği hesaplanır.

	Kişi	1	2	3	4	5	6	7	8	9	10	11	12	13
Ağırlık	Önce	146	175	150	190	220	157	136	146	128	187	172	138	150
	Sonra	142	178	147	187	212	160	135	138	132	187	171	135	151
		-	+	-	-	-	+	-	-	+	0	+	-	+
	Kişi	14	15	16	17	18	19	20	21	22	23	24	25	26
Ağırlık	Önce	124	210	148	141	164	150	130	145	132	125	137	140	126
	Sonra	126	208	148	138	159	146	130	140	130	125	138	131	120
		+	-	0	-	-	-	0	-	-	0	+	-	-

15 tane eksi, 7 tane artı gözlemlendi. Ve 4 tane "0" gözlemlendiğinden yeni  $n=22$

$$Z = \frac{(y + /-0.5) - n/2}{(1/2)\sqrt{(n)}} = \frac{(7 + 0.5) - (1/2) * 22}{(1/2) * \sqrt{22}} = -1.49$$

**3. Adım** Kritik tablo değeri bulunur.

$$Z_\alpha = Z_{0.05} = 1.645$$

**4. Adım** Karar verilir.

$|Z_{hes}| = 1.49 < 1.645 = |Z_{0.05}|$  olduğundan  $H_0$  reddedilemez.

%5 anlam düzeyinde yeni ürünün iddia edildiği gibi 2 haftada zayıflatıldığına dair yeterli kanıt bulunamamıştır.

**Örnek 3.** 8 şarkı eleştirmenine yeni bestelenen bir pop şarkı dinletilmiş ve 10 üzerinden puan vermeleri istenmiştir. Aynı kişilere bir süre sonra aynı şarkı bu sefer klip görüntüleri ile birlikte yeniden dinletilerek, 10 üzerinden yeniden puan vermeleri istenmiştir. Aşağıda, klip görüntüsünden önce ve sonra verilen puanlar karşılıklı olarak yazılmıştır. Bu durumda klip çekilmesiyle şarkının beğenilirlik düzeyi daha da artmıştır denebilecek midir?

Eleştirmen	Önce	Sonra	Fark
A	8	9	-
B	6	7	-
C	8	6	+
D	7	8	-
E	5	7	-
F	10	8	+
G	9	9	0
H	8	10	-

### 1. Adım Hipotezler kurulur

$H_0$ : Klip çekilmesi şarkının beğenilmesi üzerinde herhangi bir etki yaratmamıştır.

$H_1$ : Klip çekimi nedeni ile şarkının beğenilirlik düzeyi artmıştır.

### 2. Adım Test istatistiği hesaplanır.

Tabloya bakıldığında 5 adet (-) ve 2 adet (+) ile birlikte 1 adet (0) bulunmaktadır. Bu yüzden  $n=7$  kabul edilerek (+) işaretlilerin sayısı daha az olduğundan  $x=2$  seçilir ve tablodan  $n=7$  ve  $x=2$  için  $p_{hesap}=0.227$  olarak bulunur. Yani;

$$P(x \leq 2) = 0.227 \text{ 'dir}$$

### 3. Adım Kritik değer $\alpha = 0.05$ olarak seçilir.

### 4. Adım Karar verilir.

$P(x \leq 2) = 0.227 > 0.05 = \alpha$  olduğundan  $H_0$  reddedilemez.

```
pbinom(2,7,0.5)
```

```
## [1] 0.2265625
```

**Ödev:** Latane ve Cappell birlikteliğin kalp atış oranında etkili olup olmadığını araştırmaktadır. Rassal olarak 10 farenin birlikte ve yalnızken kal atış oranları kaydedilmiştir, ve araştırmacılar birlikteliğin kalp atışını hızlandırıldığını iddia etmektedir.  $\alpha=0.05$  için bu iddianın doğruluğu söylenebilir mi ?

Fare	1	2	3	4	5	6	7	8	9	10
Tek	463	462	462	456	450	426	418	415	409	402
Birlikte	523	494	461	535	476	454	448	408	470	437

## Wilcoxon Eşleştirilmiş Çiftlerin İşaret Testi

- Wilcoxon bağımlı iki örneklem testi, gözlem çiftlerinin sadece birbirlerinden farklı olup olmadığına değil aynı zamanda bu farklılığın büyüklüğünü de baz alan bir testtir.
- Diğer bir deyişle, Wilcoxon bağımlı iki örneklem testi  $X_i$  ve  $Y_i$  gözlem çiftlerinin farkının yönünü ortaya koyduğu gibi aynı zamanda farkın miktarını da belirleyebilmek için uygun bir testtir.
- Bu testin ana ilkesi, değerlere sıra dönüşümü uygulanmasıdır. Bu testte bağımlı iki grubun ortalamaları değil, ortancaları arasındaki farkın önemli olup olmadığı test edilir. Yani kitle medyan farkı hakkındaki hipotezi test eder.
- Genel olarak, normal dağılım göstermeyen değerler için Wilcoxon testi, t testine göre daha güçlüdür, yani önce ve sonra değerleri arasında fark varsa, daha doğru olarak saptar. **Normal dağılım gösteren değerler için her iki testin gücü aynıdır.**

### Varsayımları

- Veriler n tane gözlem değerlerinin farkından oluşur.  $D_i = X_i - Y_i$ . Her bir  $(X_i, Y_i)$  çifti rastsaldır.
- Bu farklar, sürekli değişkenlerden oluşan gözlemleri temsil eder.
- Farkların kitle dağılımı, MD median etrafında simetrik dağılır.
- Farklar birbirinden bağımsızdır.
- Farklar en az sıralı ölçekte ölçülmüştür.

### Adımları

#### 1. Adım Hipotezler kurulur

Çift yönlü hipotez

$$H_1: M_D \neq 0 \text{ (A)}$$

Tek yönlü alternatif hipotezler:

$$H_1: M_D > 0 \text{ (B)} \text{ veya } H_1: M_D < 0 \text{ (C)}$$

#### 2. Adım Test istatistiği hesaplanır.

1. İşaretlenmiş farklar elde edilir:

$$D_i = Y_i - X_i$$

2. Farkların mutlak değeri en küçükten en büyüğe sıralanarak numaralandırılır (ranklanır);

$$|D_i| = |Y_i| - |X_i|$$

3. Her bir ranka hesaplanan farkın işareti atanır.

4. Pozitif ve negatif işaretli rankların toplamı hesaplanır.

T+ : pozitif işaretli rankların toplamı

T- : negatif işaretli rankların toplamı

Burada, T+ ve T- alternatif hipoteze bağlı olarak test istatistikleridir.

- Eğer hipotez çift yönlü ise test istatistiği olarak  $\min(T-, T+)$
- Eğer hipotez sağ kuyrukta ise test istatistiği olarak T- (negatif işaretlilerin sayısı)
- Eğer hipotez sol kuyrukta ise test istatistiği olarak T+ (pozitif işaretlilerin sayısı)

**Not:** İki özel durum bulunmaktadır;

- Herhangi bir çift için  $X_i = Y_i$  olduğunda, farkın "0" olduğu o gözlemi analizden çıkarırız.
- İki veya daha fazla  $|D_i|$  eşit olduğunda,  $|D_i|$  o değere atanan rankların ortalamasını alır.

**3. Adım** Kritik tablo değeri bulunur.

$W_{\alpha,n}$  veya  $W_{\alpha/2,n}$  tablo değeri bulunur

**4. Adım** Karar verilir.

- Eğer  $H_0$  doğru ise, pozitif işaretli rankların toplamı kadar negatif işaretli ranklar olmasını bekleriz.
- Hem T+'nın yeterince küçük değeri için hem de T-'nin yeterince küçük değeri için  $H_0$  hipotezi yani farkların medyanının 0 olduğu hipotez reddedilir. Bu durumda test istatistiği yeterince küçüktür.
- **Eğer hesaplanan T istatistiği,  $\alpha$  anlamlılık seviyesinde W-tablo değerinden küçük ise  $H_0$  hipotezi reddedilir.**
- $H_0$  hipotezini reddedersek, farkların kitlesinin medyanı 0 değildir deriz.

### Büyük Örneklem İçin

$n > 20$  olduğunda örnekleme dağılımı normale yakınsar ve test istatistiği z-dönüşümü yardımıyla aşağıdaki gibi hesaplanır.

$$z_{hes} = \frac{T - \mu_T}{\sqrt{\sigma_T^2}}$$

$$\mu_T = \frac{n(n+1)}{4}$$

$$\sigma_T^2 = \frac{n(n+1)(2n+1)}{24}$$

\*\*Eğer  $|z_{hes}| > |z_{tablo}|$  ise  $H_0$  reddedilir.



**Örnek** Bir araştırmada hastaların akciğer trombositlerindeki hemodinamik değişimleri araştırılmaktadır. Tabloda terapide 24 saat önce ve sonra 9 hasta için ölçülen akciğer arterlerindeki ortalama basınç verilmektedir.

Bu veriden hareketle, bu terapinin akciğer arter basıncını düşürdüğü söylenebilir mi?

Akciğer arterindeki ortalama basınç (mm)									
Hasta	1	2	3	4	5	6	7	8	9
0 saat (X)	33	17	30	25	36	25	31	20	18
24 saat (Y)	21	17	22	13	33	20	19	13	9

**1. Adım** Hipotezler kurulur.  $M_D = X - Y$

$$H_0: M_D \geq 0$$

$$H_1: M_D < 0$$

**2. Adım** Test istatistiği hesaplanır.

Terapiden Önce (X)	24 sa. sonra (Y)	$D_i = Y_i - X_i$	İşaretlenmiş Rank
33	21	-12	-7
17	17	0	----
30	22	-8	-4
25	13	-12	-7
36	33	-3	-1
25	20	-5	-2
31	19	-12	-7
20	13	-7	-3
18	9	-9	-5

Hipotez sol kuyruk tek yön olduğu için T+ kullanılır.  $T^+ = 0$ , n=8 olarak bulunur.

**3. Adım** Kritik tablo değeri bulunur.  $W_5 = 0.039$

**4. Adım** Karar verilir.

$T^+ = 0 < W_{0.05,5} = 0.0039$  olduğundan  $H_0$  reddedilir ve farkların medyanı 0'dan küçüktür denir. 0.05 önem seviyesinde terapi akciğer arteri basıncını azaltmaktadır denir.

**Örnek** Bir süre İnsülin tedavisi gören 12 hastada açlık kan şekeri düzeyleri(mg) aşağıdaki gibi verilmiştir.

Öncesi	Sonrası
280	120
260	90
180	110
130	140
250	160
190	100
220	100
270	130
160	100
180	90
140	160
150	150

**1. Adım** Hipotezler kurulur.  $M_D = X - Y$

$H_0: M_D \geq 0$  (İlaç etkili olmamıştır)

$H_1: M_D < 0$  (İlaç etkili olmuştur ve insülin miktarını azaltmıştır.)

**2. Adım** Test istatistiği hesaplanır.

İnsülin Önce (X)	İnsülin sonra (Y)	$D_i = Y_i - X_i$	İşaretlenmiş Rank
280	120	-160	10
260	90	-170	11
180	110	-70	4
130	140	10	1
250	160	-90	6
190	100	-90	6
220	100	-120	8
270	130	-140	9
160	100	-60	3
180	90	-90	6
140	160	20	2
150	150	0	----

$T_+ = 3$   $n=11$  ,  $\alpha = 0.05$

**3. Adım** Kritik tablo değeri bulunur.

$$T_{tablo} = T_{0.05,12} = 13$$

**4. Adım** Karar verilir.

$T_+ = 3 < 13 = T_{tablo}$  olduğundan  $H_0$  reddedilir.

Hipotez çift taraflı hipotez olsaydı;

$$\alpha = 0.05, \alpha/2 = 0.025, T_+ = 3, n = 11, T_{tablo} = 10$$

$T_+ = 3 < T_{tablo} = 10$  olduğundan  $H_0$  reddedilir.

```
if(!require(exactRankTests)) install.packages("exactRankTests")
## Zorunlu paket yükleniyor: exactRankTests
## Warning: package 'exactRankTests' was built under R version 4.2.3
## Package 'exactRankTests' is no longer under development.
## Please consider using package 'coin' instead.

A<-c(33,17,30,25,36,25,31,20,18)
B<-c(21,17,22,13,33,20,19,13,9)

wilcox.exact(x=B,y=A, paired = T, alternative = "less")

##
## Exact Wilcoxon signed rank test
##
## data: B and A
## V = 0, p-value = 0.003906
## alternative hypothesis: true mu is less than 0

A<-c(280,260,180,130,250,190,220,270,160,180,140,150)
B<-c(120,90,110,140,160,100,100,130,100,90,160,150)

wilcox.exact(x=B,y=A, paired = T, alternative = "less")

##
## Exact Wilcoxon signed rank test
##
## data: B and A
## V = 3, p-value = 0.002441
## alternative hypothesis: true mu is less than 0
```

**Table A5 Table of Critical  $T$  Values for Wilcoxon's Signed-Ranks and Matched-Pairs Signed-Ranks Test**

One-tailed level of significance					One-tailed level of significance				
.05 .025 .01 .005					.05 .025 .01 .005				
Two-tailed level of significance					Two-tailed level of significance				
.10 .05 .02 .01					.10 .05 .02 .01				
$n$					$n$				
5	0	–	–	–	28	130	116	101	91
6	2	0	–	–	29	140	126	110	100
7	3	2	0	–	30	151	137	120	109
8	5	3	1	0	31	163	147	130	118
9	8	5	3	1	32	175	159	140	128
10	10	8	5	3	33	187	170	151	138
11	13	10	7	5	34	200	182	162	148
12	17	13	9	7	35	213	195	173	159
13	21	17	12	9	36	227	208	185	171
14	25	21	15	12	37	241	221	198	182
15	30	25	19	15	38	256	235	211	194
16	35	29	23	19	39	271	249	224	207
17	41	34	27	23	40	286	264	238	220
18	47	40	32	27	41	302	279	252	233
19	53	46	37	32	42	319	294	266	247
20	60	52	43	37	43	336	310	281	261
21	67	58	49	42	44	353	327	296	276
22	75	65	55	48	45	371	343	312	291
23	83	73	62	54	46	389	361	328	307
24	91	81	69	61	47	407	378	345	322
25	100	89	76	68	48	426	396	362	339
26	110	98	84	75	49	446	415	379	355
27	119	107	92	83	50	466	434	397	373