KORELASYON ANALIZI



Kovaryans

- Korelasyon konusunu anlatmadan önce ilk olarak kovaryans terimini inceleyeceğiz.
- İki değişkenin ilişkili olup olmadığını öğrenmenin en basit yolu bu iki değişkenin birbirlerine göre değişimini gösteren kovaryans değerini hesaplamaktır.
- Kovaryansı daha iyi anlamak için betimsel istatistiklerde bahsettiğimiz varyans formülüne bakmakta fayda vardır.

Varyans

 Bir değişkenin varyansı verinin aritmetik ortalamadan ortalama uzaklığını temsil eder.

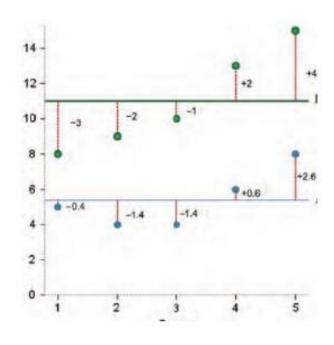
Katılımcı	1	2	3	4	5	Ar. Ortalama	St. Sapma	Varyans
İzlenen Reklam Sayısı	5	4	4	6	8	5.4	1.67	2.80
Alınan Ürün	8	9	10	13	15	11.0	2.92	8.50

variance(
$$s^2$$
) = $\frac{\sum (x_i - \overline{x})^2}{N-1}$ = $\frac{\sum (x_i - \overline{x})(x_i - \overline{x})}{N-1}$

 Yukarıdaki formülü kullanarak varyans değerleri 2.80 ve 8.50 olarak hesaplanır.

Varyans-Kovaryans

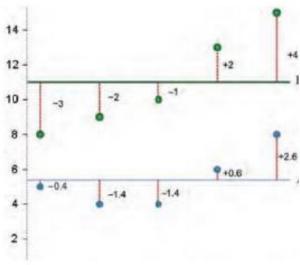
Yandaki grafikte iki değişken için de her bir değerin ortalamadan farkı (sapmalar) gösterilmektedir. Şimdi iki değişkenin birlikte değişimini bulabilmek için her bir değerin ortalamadan farkının carpımını bulacağız. Bu işleme kovaryans hesaplaması denir.



Varyans-Kovaryans

$$cov(x, y) = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{N - 1}$$

 Yandaki grafikteki sapmaları ve üstteki formülü kullanarak aşağıdaki işlemlerle kovaryans değerini hesaplarız. Sapmaların çarpımları hep pozitif olduğu için pozitif bir kovaryans değeri yani pozitif bir ilişki beklenebilir.



$$cov(x,y) = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{N - 1}$$

$$= \frac{(-0.4)(-3) + (-1.4)(-2) + (-1.4)(-1) + (0.6)(2) + (2.6)(4)}{4}$$

$$= \frac{1.2 + 2.8 + 1.4 + 1.2 + 10.4}{4}$$

$$= \frac{17}{4}$$

Kovaryans

- Kovaryans değerini hesaplayarak iki değişkenin birbirlerine göre değişimi yani ilişkisi gösterilebilir.
- Bir değişkenin değerleri ortalamanın üzerinde iken diğer değişkenin değerleri de ortalamanın üzerinde ise bu iki değişken arasında pozitif bir ilişki vardır diyebiliriz. Bu durumda iki değişken arasında pozitif bir ilişki vardır diyebiliriz.
- Biri ortalamanın altında iken diğeri de ortalamanın altında değerler gösteriyorsa genelde kovaryans negatif çıkar. Bu durumda iki değişken arasında negatif bir ilişki vardır diyebiliriz.
- Fakat kovaryans kullanmadaki problem kovaryansın değişkenlerin birimine bağlı olmasıdır. Eğer iki değişken farklı birimler ile ölçülüyorsa (kg vs. km) bu durumda kovaryansın değerini yorumlamada zorluk yaşarız. Büyük ya da küçük olmasının ne anlama geldiğini söylemek zorlaşır.

Korelasyon

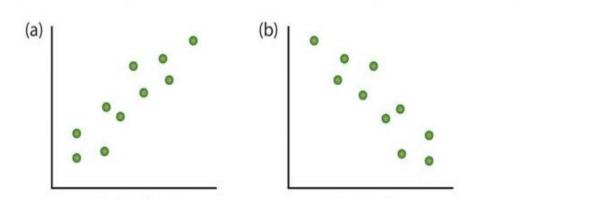
Kovaryanstaki birim probleminden kurtulmak için kovaryans değerini standartlaştırmamız gerekmektedir. Bir şekilde her türlü birimi ortak bir değere çevirebilmemiz lazım. Bunu yapabilmek için standart sapma kullanmamız gerekmektedir. Aynı standart z puanlarının hesaplanmasında yaptığımız gibi her hangi bir değerin ortalamadan sapmasını standart sapmaya böldüğümüzde standart bir ölçek elde ederiz. Kovaryans formülünü standart sapma değerleri ile böldüğümüzde elde edeceğimiz değerin adı korelasyon olacaktır.

Korelasyon Formülü

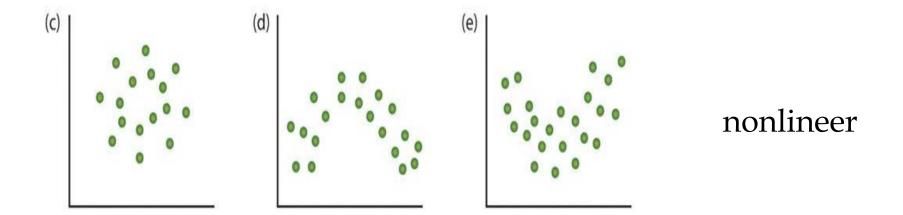
Aşağıdaki formül vasıtasıyla iki değişkene ait değerler kullanılarak hesaplanan değere «Pearson Momentler Çarpımı Korelasyon Katsayısı» denir. Karl Pearson tarafından geliştirildiği için Pearson korelasyonu adını almıştır.

$$r = \frac{\text{cov}_{xy}}{s_x s_y} = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{(N - 1)s_x s_y}$$

Doğrusal ve Eğrisel İlişkiler



lineer



korelasyon katsayısının yorumlanması

	Coefficient r				
	Positive Negative				
Strong	1 to 0.8	-0.8 to -1			
Moderate	0.8 to 0.5	-0.5 to -0.8			
Weak	0.5 to 0.3	-0.3 to -0.5			
No Correlation	0.3 to 0	0 to -0.3			

Weber, 1970 Madison vd, 1983

Grading Standards

Correlation Degree

ho=0	no correlation
$0 < \rho \le 0.19$	very week
$0.20 \le \rho \le 0.39$	weak
$0.40 \le \rho \le 0.59$	moderate
$0.60 \le \rho \le 0.79$	strong
$0.80 \le \rho \le 1.00$	very strong
1.00	monotonic correlation

Evans, 1996

Pearson Korelasyon Katsayısının yorumu; r İlişki

0,00-0,20

0,20-0,40

0,40-0,60

0,60-0,80

0,80-1,00

Çok Zayıf

Zayıf

Orta

Yüksek

Çok Yüksek

1. Pearson korelasyon katsayısı

- Her iki değişkenin de *sayısal* (kantitatif) değişken olması ve değişkenlere ilişkin verilerin normal dağılım göstermesi durumunda değişkenler arasındaki ilişki *Pearson korelasyon katsayısı* ile belirlenir (parametrik test).
- Değişkenlerin değerleri sayısal değilse:
 - Kendall's tauveya
 - Spearman korelasyon katsayısı kullanılmalıdır (non-parametrik test).
 - Eğer değişkenler arasındaki ilişki doğrusal değil ise hesaplanan korelasyon katsayısı değişkenler arsındaki ilişkiyi ölçmek için uygun bir istatistik değildir.

Korelasyon Formülü

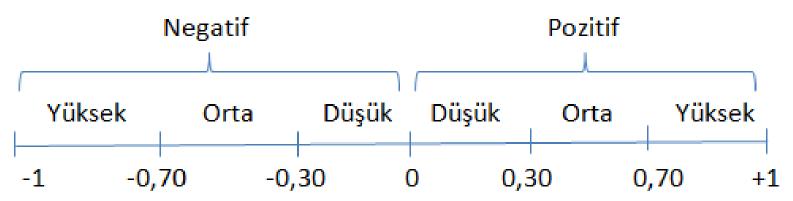
Aşağıdaki formül vasıtasıyla iki değişkene ait değerler kullanılarak hesaplanan değere «Pearson Momentler Çarpımı Korelasyon Katsayısı» denir. Karl Pearson tarafından geliştirildiği için Pearson korelasyonu adını almıştır.

$$r = \frac{\text{cov}_{xy}}{s_x s_y} = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{(N - 1)s_x s_y}$$

$$\mathbf{r}_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x}). (y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2. \sum_{i=1}^{n} (y_i - \bar{y})^2}}$$

$$-1 \le r \le +1$$

*r'nin Yorumlanması



Mükemmel Mükemmel

$$-1 \le r \le +1$$

*r'nin anlamlılığının test edilmesi

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

serbestlik derecesi= n-2

Örnek1: Aşağıda verilen vize ve final notları arasındaki korelasyon katsayısını hesaplayınız.

Vize	Final
45	83
54	78
55	80
68	72
30	45
48	26

$$\mathbf{r}_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x}). (y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2. \sum_{i=1}^{n} (y_i - \bar{y})^2}}$$

X	Y	$(x-\overline{x})$	$(x-\overline{x})^2$	$(y-\overline{y})$	$(y-\overline{y})^2$	$(x-\overline{x})(.y-\overline{y})$
45	83	-5	25	19	361	-95
54	78	4	16	14	194	56
55	80	5	25	16	256	80
68	72	18	324	8	64	144
30	45	-20	400	-19	361	380
48	26	-2	4	-38	1444	76
			$\sum_{x} (x - \bar{x})^2$ $= 794$		$\sum (y - \bar{y})^2$	$\sum (x - \bar{x})(.y - \bar{y}) = 641$
			= 794		= 2682	$-\bar{y}) = 641$

$$\overline{x} = 50 \quad \overline{y} = 64$$

$$\sum_{y=794} (x - \bar{x})^2 = 2682$$
$$\sum_{y=2682} (x - \bar{y})^2 = 2641$$

$$\mathbf{r}_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x}). (y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2. \sum_{i=1}^{n} (y_i - \bar{y})^2}}$$

$$=\frac{641}{\sqrt{794.2682}}$$

$$=\frac{641}{1459,2838}$$

$$= 0,4392 = 0,44$$

Evans, 1996

→ Pozitif yönde orta düzeyde

	Grading Standards	Correlation Degree
	$\rho = 0$	no correlation
_	$0 < \rho \le 0.19$	very week
2	$0.20 \le \rho \le 0.39$	weak
	$0.40 \le \rho \le 0.59$	moderate
	$0.60 \le \rho \le 0.79$	strong
	$0.80 \le \rho \le 1.00$	very strong
	1.00	monotonic correlation

Anlamlılık testi için;

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

$$t = \frac{0,44}{\sqrt{\frac{1 - (0,44)^2}{6 - 2}}} = \frac{0,44}{\sqrt{\frac{0,8064}{4}}} = \frac{0,44}{\sqrt{0,2016}} = \frac{0,44}{0,4489 \dots}$$

$$= 0.9799 = 0.98$$

Örnekte;

 H_0 : Öğrencilerin vize ve final sın

 H_1 : Öğrencilerin vize ve final sın



Tek kuyruk? Çift Kuyruk?

Tablo: t dağılımı tablosu

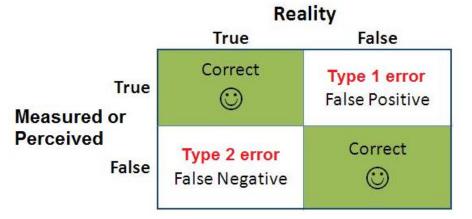
						Lasa
Serbestlik		A/E	Olası	lık (α)	81	
derecesi	0,100	0,050	0,025	0,010	0,005	0,001
1	3,078	6,314	12,706	31,821	63,657	318,309
2	1,886	2,920	4,303	6,965	9,925	22,327
3	1,639	2,353	3,192	1,511	5,941	10,215
4	1,533	2,132	2,776,	3,747	4,604	7,173
5	1,4/6	2,015	2,571	3,305	4,032	5,893
6	1,440	1,943	2,447	3,143	3,707	5,208
7	1,415	1,895	2,365	2,998	3,499	4,785
8	1,397	1,860	2,306	2,896	3,355	4,501
9	1,383	1,833	2,262	2,821	3,250	4,297
10	1,372	1,812	2,228	2,764	3,169	4,144
11	1,363	1,796	2,201	2,718	3,106	4,025
12	1,356	1,782	2,179	2,681	3,055	3,930
13	1,350	1,771	2,160	2,650	3,012	3,852
14	1,345	1,761	2,145	2,624	2,977	3,787
15	1,341	1,753	2,131	2,602	2,947	3,733

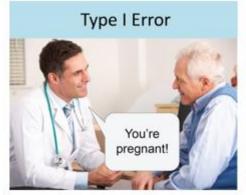
0,98 < 2,132

 $t_{hesaplanan < t_{tablo}}$ olduğundan H_0 :Kabul olur. Yani, vize ve final sınavlarına ilişkin notlar arasında anlamlı bir ilişki yoktur.

Hatırlatma...

	H₀:Kabul	H₀:Red
Doğru	1 – α	2.tip hata eta
Yanlış	1.tip hata α	1-β







	Null hypothesis is TRUE	Null hypothesis is FALSE
Reject null hypothesis	Type I Error (False positive)	Correct outcome! (True positive)
Fail to reject null hypothesis	Correct outcome! (True negative)	Type II Error (False negative)

ÖRNEK2: Araştırma sorusu: Saçı uzun olanın aklı kısa mıdır?

 H_0 : Saç uzunluğu ile akıl arasında bir ilişki yoktur. H_1 : Saç uzunluğu ile akıl arasında ilişki vardur.

Sınıftaki öğrencilere ait saç uzunluğu ile GANO notları aşağıda verilmiştir.

Sıra	Saç Uzunluğu	GANO
1	40	3,00
2	30	2,50
3	3	2,40
4	7	2,00
5	50	2,80
6	20	2,50
7	25	3,70
8	10	2,90

ÇÖZÜM:

X	У	$(x-\bar{x})$	$(y-\overline{y})$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(.y-\bar{y})$
40	3,00	16,875	0,275			
30	2,50	6,875	-0,225			
3	2,40	-20,125	-0,325			
7	2,00	-16,125	-0,725			
50	2,80	27,875	0,075			
20	2,50	-3,125	-0,225			
25	3,70	1,875	0,975			
10	2,90	-13,125	0,175			

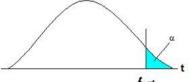
siz devam ediniz...

$$\mathbf{r}_{xy} = \frac{\sum_{i=1}^{n} (x_i - \overline{x}). (y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \overline{x})^2. \sum_{i=1}^{n} (y_i - \overline{y})^2}}$$

$$=\frac{23,575}{\sqrt{1904,90\times1795}}$$

$$= 0.4031 \dots = 0.40$$

Tablo: t dağılımı tablosu



Serbestlik derecesi	Olasılık (α)						
	0,100	0,050	0,025	0,010	0,005	0,001	
1	3,078	6,314	12,706	31,821	63,657	318,309	
2	1,886	2,920	4,303	6,965	9,925	22,327	
3	1,638	2,353	3,182	4,541	5,841	10,215	
4	1,533	2,132	2,776,	3,747	4,604	7,173	
5	1,476	2,015	2,571	3,365	4,032	5,893	
6	1,440	1,943	2,447	3,143	3,707	5,208	
7	1,415	1,895	2,365	2,998	3,499	4,785	
8	1,397	1,860	2,306	2,896	3,355	4,501	
9	1,383	1,833	2,262	2,821	3,250	4,297	
10	1,372	1,812	2,228	2,764	3,169	4,144	
11	1,363	1,796	2,201	2,718	3,106	4,025	
12	1,356	1,782	2,179	2,681	3,055	3,930	
13	1,350	1,771	2,160	2,650	3,012	3,852	
14	1,345	1,761	2,145	2,624	2,977	3,787	
15	1,341	1,753	2,131	2,602	2,947	3,733	

$$t = \frac{r}{\sqrt{1 - r^2}} = \frac{0.40}{\sqrt{\frac{1 - 0.40^2}{8 - 2}}} = \frac{0.40}{\sqrt{\frac{0.84}{6}}} = \frac{0.40}{\sqrt{0.14}} = \frac{0.40}{0.3741657} = 1.069045 = 1.07$$

 $t_{besaplanan} < t_{tablo} \rightarrow H_0$: Kabul

2. Spearman Brown Sıra Farkları Korelasyon Katsayısı (rho, r_s) (Köklü ve diğ., 2006)

- Sıralı puanlar kullanılarak ölçülen iki değişken arasındaki doğrusal ilişkiyi açıklar.
- Eşit aralıklı ya da eşit oran düzeyindeki ölçümler sıralı ölçeğe dönüştürülerek analiz edilmesi gerektiğinde de ölçümler sıralı ölçeğe dönüştürülerek uygulanabilir.
- Ayrıca Pearson korelasyonundaki değişkenler, normal dağılım varsayımını sağlamadığında da Pearson korelasyon yerine kullanılır.

*r'nin formülü ve t-değerinin hesaplanması

$$\boldsymbol{r_s} = 1 - \frac{6\sum \boldsymbol{d}^2}{\boldsymbol{n}.\left(\boldsymbol{n}^2 - 1\right)}$$

$$t = \frac{r_s}{\sqrt{\frac{1 - r_s^2}{n - 2}}}$$

SD=n-1

Örnek3: Aşağıda matematik bölümünde öğrenim gören öğrencilerin iki farklı derse ilişkin başarı sıraları gösterilmektedir. Buna göre iki farklı dersten elde edilen sıralamalar arasındaki korelasyon katsayısını ve anlamlılığını hesaplayınız.

Öğrenciler	Müjdat Hoca	Serhan Hoca
Esma	2	1
Kadir	1	4
Ayşe	3	2
Cemile	4	6
Rana	5	3
Begüm	7	5
Kübra	6	7

Çözüm;

Öğrenciler	Müjdat	Serhan	D	D^2
Esma	2	1	1	1
Kadir	1	4	-3	9
Ayşe	3	2	1	1
Cemile	4	6	-2	4
Rana	5	3	2	4
Begüm	7	5	2	4
Kübra	6	7	-1	1

$$\boldsymbol{r_s} = 1 - \frac{6\sum \boldsymbol{d}^2}{\boldsymbol{n}.\left(\boldsymbol{n}^2 - 1\right)}$$

$$\sum D^2 = 24$$

$$r_s = 1 - \frac{6\sum d^2}{n.(n^2 - 1)}$$
 $1 - \frac{6x24}{7x(49 - 1)} = 0.57$

$$1 - \frac{6x24}{7x(49-1)} = 0,57$$

*r'nin formülü ve t-değerinin hesaplanması

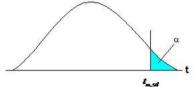
$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

$$=\frac{0,57}{\sqrt{\frac{1-0,57^2}{5}}}$$

$$=\frac{0,57}{\sqrt{0,13502}}$$

$$=\frac{0.57}{0.36745}=1.36$$

Tablo: t dağılımı tablosu



-						£ _{m,Sd}
Serbestlik	Olas _ι lik (α)					
derecesi	0,100	0,050	0,025	0,010	0,005	0,001
1	3,078	6,314	12,706	31,821	63,657	318,309
2	1,886	2,920	4,303	6,965	9,925	22,327
3	1,638	2,353	3,182	4,541	5,841	10,215
4	1,533	2,132	2,776、	3,747	4,604	7,173
5	1,476	2,015	2,571	3,365	4,032	5,893
6	1,440	1,943	2,447	3,143	3,707	5,208
7	1,415	1,895	2,365	2,998	3,499	4,785
8	1,397	1,860	2,306	2,896	3,355	4,501
9	1,383	1,833	2,262	2,821	3,250	4,297
10	1,372	1,812	2,228	2,764	3,169	4,144
11	1,363	1,796	2,201	2,718	3,106	4,025
12	1,356	1,782	2,179	2,681	3,055	3,930
13	1,350	1,771	2,160	2,650	3,012	3,852
14	1,345	1,761	45 [ر 2	2,624	2,977	3,787
15	1,341	1,753	2,131	2,602	2,947	3,733

 $t_{hesaplanan} < t_{tablo}
ightarrow H_0$: Kabul (Yani, ilişki anlamlı değil)

Örnek4: Aşağıda matematik bölümünde öğrenim gören öğrencilerin iki farklı dersin final sınavından aldıkları notlar görünmektedir. Buna göre iki farklı dersten elde edilen sıralamalar arasındaki korelasyon katsayısını ve anlamlılığını hesaplayınız.

Öğrenciler	İstatistik	Topoloji
Ayşe	40	60
Hatice	70	20
Mert	50	10
Kağan	30	30
Hakan	60	20
Sıla	55	40
Mahmut	80	70

Çözüm;

Öğrenciler	İstatistik	Topoloji	D	D^2
Ayşe	40 (6)	60 (2)	4	16
Hatice	70 (2)	20 (5.5)	-3,5	12,25
Mert	50 (5)	10 (7)	-2	4
Kağan	30 (7)	30 (4)	3	9
Hakan	60 (3)	20 (5.5)	-2,5	6,25
Sıla	55 (4)	40 (3)	1	1
Mahmut	80 (1)	70 (1)	0	0
				$\sum D^2$
				= 48,5

$$r_s = 1 - \frac{6\sum d^2}{n.(n^2 - 1)} = 1 - \frac{6 \times 48,5}{7x(49 - 1)} = 1 - 0,87 = 0,13$$

3) Çift Serili Korelasyon Katsayısı (Biserial Correlation Coeficient)

Aralarında ilişki olduğu düşünülen iki değişkenden birinin sürekli, diğerinin yapay olarak süreksiz hale getirilmesi (iki kategorili) durumunda kullanılır.

$$r_{\varsigma} = \frac{\bar{x}_p - \bar{x}_q}{S_{\chi}} \cdot \frac{p \cdot q}{y}$$

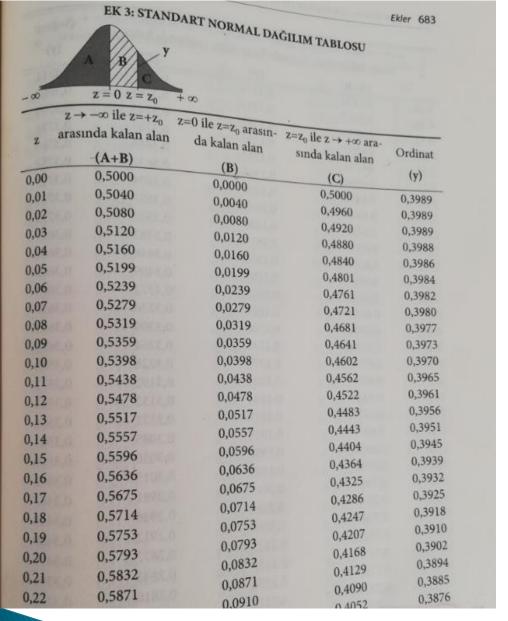
$$t = \frac{r_{\varsigma}.\sqrt{n-2}}{\sqrt{1-r_{\varsigma}^2}}$$

SD = n - 2

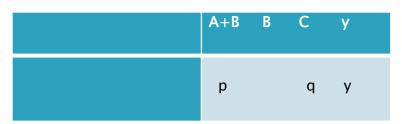
Y: ordinat değeri

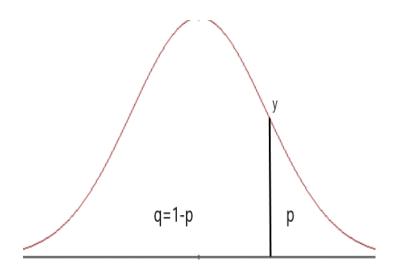
P: gözleme oranı

Q: gözlenmeme oranı



Kaynak: Baykul ve Güzeller (2014)





$$r_{\varsigma} = \frac{\bar{x}_p - \bar{x}_q}{S_x} \cdot \frac{p \cdot q}{y}$$

4) Nokta- Çift Serili Korelasyon Katsayısı

(Point-biserial Correlation Coeficient)

Aralarında ilişki olduğu düşünülen iki değişkenden birinin sürekli, diğerinin doğal iki kategorili olması durumunda kullanılır.

Or. Cinsiyet ile IQ düzeyi arasındaki ilişkinin belirlenmesi

$$r_{n\varsigma} = (\frac{\bar{x}_p - \bar{x}_q}{S_x}).\sqrt{p.q}$$
 $t = \frac{r_{\varsigma}.\sqrt{n-2}}{\sqrt{1 - r_{\varsigma}^2}}$ $SD = n-2$

$$t = \frac{r_{\varsigma}}{\sqrt{\frac{1 - r_{\varsigma}^2}{n - 2}}}$$
 Y: ordinat değeri P: gözleme oranı

Q: gözlenmeme oranı

Örnek5: Aşağıda evli ve bekar adayların KPSS sınavından elde ettikleri puanlar gösterilmektedir. Buna göre iki değişken arasındaki korelasyon katsayısını hesaplayınız.

Medeni		
Durum		
Bekâr		
Evli		
Evli		
Evli		
Bekâr		
Bekâr		
Bekâr		
Evli		
Bekâr		
Bekâr		

$$H_0$$
: $r_{n\varsigma}$ =0

$$H_0: r_{n\varsigma} = 0$$
 $H_1: r_{n\varsigma} \neq 0$

KPSS	Medeni		
	Durum		
70	Bekâr		
60	Evli		
65	Evli		
70	Evli		
10	Bekâr		
40	Bekâr		
35	Bekâr		
55	Evli		
40	Bekâr		
50	Bekâr		

Öncelikle evli ve bekar adaylara ilişkin aritmetik ortalamalar hesaplanacaktır.

(evli):
$$\bar{x}_p = 62.5$$

(bekar):
$$\bar{x}_q$$
= 40,833

Sonrasında tüm adaylara (evli ve bekar) ilişkin ortalama ve standart sapma hesaplanacaktır.

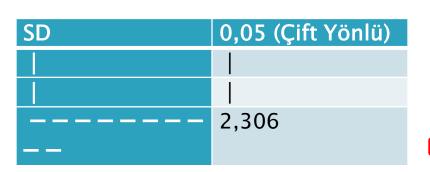
$$\bar{x} = 49.5$$
 p= 0.40

$$Sx = 18,774$$
 $q = 0,60$

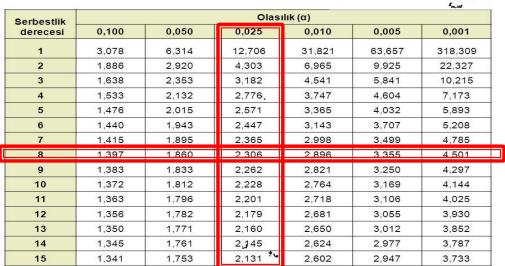
$$r_{n\varsigma} = (\frac{\bar{x}_p - \bar{x}_q}{S_x}).\sqrt{p.q}$$

$$r_{n\varsigma} = (\frac{\bar{x}_p - \bar{x}_q}{S_x}).\sqrt{p.q} = (\frac{62,5-40,833}{18,774}).\sqrt{0,40.0,60} = 0,565 = 0,57 \text{(orta düzeyde)}$$

$$t = \frac{r_{n\varsigma}.\sqrt{n-2}}{\sqrt{1-(r_{n\varsigma})^2}} = \frac{0.57.\sqrt{8}}{\sqrt{1-0.57x0.57}} = 0.57.\sqrt{\frac{8}{0.6751}} = 0.57.\sqrt{11.8501} = 1.96$$



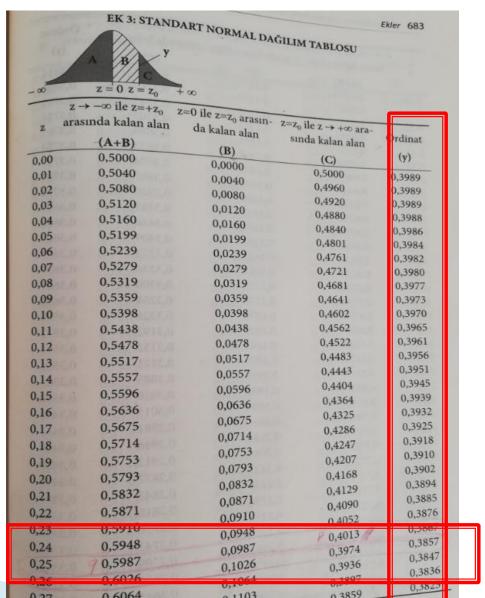
Tablo: t dağılımı tablosu



Çözüm;
$$r_{n\varsigma} = (\frac{\bar{x}_p - \bar{x}_q}{S_x}).\sqrt{p.q} \quad \text{yerine}$$

Eğer çift serili olsaydı yandaki tablo kullanılacaktı!!!

$$r_{\varsigma} = \frac{\bar{x}_p - \bar{x}_q}{S_x} \cdot \frac{p \cdot q}{y}$$



Örnek6: Aşağıda bir sınıftaki kız ve erkek öğrencilerin İstatistik sınavından elde ettikleri puanlar gösterilmektedir. Buna göre iki değişken arasındaki korelasyon katsayısını

hesaplayınız.

$H_0: r_{n\varsigma} = 0$

$$H_1: r_{n\varsigma} \neq 0$$

Notlar	Cinsiyet
75	Kadın
82	Kadın
69	Erkek
57	Erkek
69	Kadın
67	Kadın
57	Kadın
88	Erkek
92	Kadın
84	Kadın
91	Kadın
80	Kadın
72	Kadın
75	Erkek
70	Kadın
52	Erkek
68	Erkek
79	Kadın
62	Erkek
59	Erkek

Notlar	Cinsiyet
75	Kadın
82	Kadın
69	Erkek
57	Erkek
69	Kadın
67	Kadın
57	Kadın
88	Erkek
92	Kadın
84	Kadın
91	Kadın
80	Kadın
72	Kadın
75	Erkek
70	Kadın
52	Erkek
68	Erkek
79	Kadın
62	Erkek
59	Erkek

Öncelikle kız ve erkek adaylara ilişkin aritmetik ortalamalar hesaplanacaktır.

(kiz):
$$\bar{x}_p = \frac{75 + 82 + 69 + \dots + 79}{12} = 76,50$$

(erkek):
$$\bar{x}_q = \frac{69 + 57 + 88 + \dots + 59}{8} = 66,25$$

Sonrasında tüm adaylara (kız ve erkek) ilişkin standart sapma hesaplanacaktır.

$$\bar{x}_{genel}$$
=72,40 p= 0,40 Sx= 11,63 q= 0,60

$$r_{n\varsigma} = \left(\frac{\bar{x}_{p} - \bar{x}_{q}}{S_{x}}\right) \cdot \sqrt{p \cdot q} = \left(\frac{76,5 - 66,25}{11,63}\right) \cdot \sqrt{0,40.0,60}$$
$$= \left(\frac{10,25}{11.63}\right) \cdot (0,49) = 0,4318 = 0,43$$

$$t = \frac{r_{n\varsigma}.\sqrt{n-2}}{\sqrt{1-(r_{n\varsigma})^2}}$$

$$= \frac{0,43.\sqrt{20-2}}{\sqrt{1-(0,43)^2}} = 0,43.\sqrt{\frac{18}{1-0,0559}}$$
$$= 0,43.\sqrt{\frac{18}{0,9441}}$$
$$= 2,02$$

 $t_{hesaplanan} < t_{tablo}
ightarrow H_0$: Kabul (Yani, ilişki anlamlı değil)

5) Dörtlü Korelasyon Katsayısı

Aralarında ilişki olduğu düşünülen iki değişkenin de iki kategorili olması durumunda kullanılan korelasyon katsayısıdır.

Bu korelasyon türüne medeni durum ile sigara içme arasındaki ilişki, makyaj yapma ile atanma durumu arasındaki ilişki ve gözlük kullanma ile sınıfı geçme arasındaki ilişki gibi her ikisi de iki düzeyden oluşan iki durumun olduğu olaylar örnek verilebilir.

	1. DEĞİŞKEN			
		1	0	
2. DEĞİŞKEN	1	A	В	
	0	C	D	

$$\varphi = \frac{BC - AD}{\sqrt{(A+C)(B+D)(A+B)(C+D)}}$$

Örnek7: Aşağıda bir grup kız öğrenciden elde edilen verilere göre katılımcıların makyaj yapma durumları ile lisan düzeyinde mezun olup olmama durumları gösterilmiştir. Buna göre iki değişken arasındaki korelasyon katsayısını hesaplayınız.

 $H_0: r_{n\varsigma} = 0$

 $H_1: r_{n\varsigma} \neq 0$

	Makyaj yapıyor	Makyaj yapmıyor
Lisan mezunu değil	52	60
Lise mezunu	47	41

	Makyaj yapıyor	Makyaj yapmıyor	Toplam
Lisan mezunu	52	60	112
Lise mezunu	47	41	88
Toplam	99	101	200

$$\varphi = \frac{BC - AD}{\sqrt{(A+C)(B+D)(A+B)(C+D)}}$$

$$=\frac{47x60-52x41}{\sqrt{99x101x112x88}}$$

$$=\frac{688}{9927.2425} = 0,069302 \cong 0,07$$

İlişkinin anlamlı olup olmadığını tabloya bakılarak yorumlarız.

6) Tetrakorik Korelasyon Katsayısı

Normal dağılıma sahip sürekli iki değişkenin yapay olarak iki kategorili hale getirilmesi sonucunda elde edilen değişkenler arası ilişkinin belirlenmesinde kullanılır. (Genelde başarı testlerinde kullanılır.)

$$r_{t=} \cos \frac{180^{\circ}}{1 + \sqrt{\frac{AB}{CD}}}$$

Ornek8: Aşağıda bir testte yer alan iki dersten geçme ve kalma durumları gösterilmiştir. Buna göre iki dersten geçme durumları arasında anlamlı bir ilişki var mıdır?

 $H_0: r_{n\varsigma} = 0$ $H_1: r_{n\varsigma} \neq 0$

		İstatistik		
		1 (geçti)	0 (kaldı)	
Pataloji	1 (geçti)	40	15	
	0 (kaldı)	60	20	

$$r_{t=} \cos \frac{180^{\circ}}{1 + \sqrt{\frac{AB}{CD}}}$$
 $r_{t} = \cos \frac{180^{\circ}}{1 + \sqrt{\frac{40x15}{60x20}}} = \cos \left(\frac{180^{\circ}}{1,7071}\right)$

Cos (105,44°)= 0,27 (düşük ilişki)

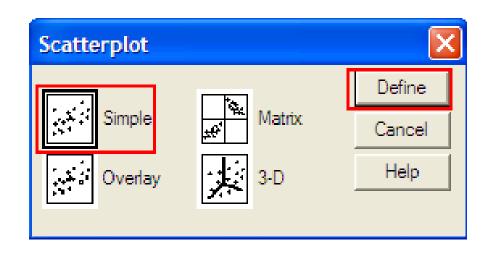
İstatistik 2 Dersinde Bununla İlgili Uygulama SPSS Programı üzerinden yapılacaktır..

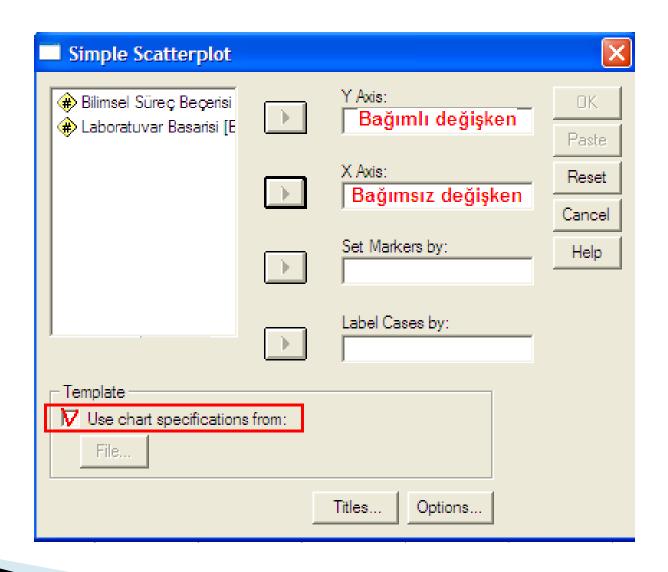
Ör. Fen bilgisi laboratuvar dersini alan bir grup öğrencinin bilimsel süreç becerisi (BSB) puanları ile ders başarı puanları arasında anlamlı bir ilişki var mıdır? Şeklindeki bir araştırma sorusunun cevabını araştıralım.

iii korel	asyon - SPSS	Data Editor			
File Edit	View Data	Transform A	nalyze Graph	s Utilities W	/indow Help
	a 🖳 🖂		#4 *	<u>* # 4 </u>	<u>F</u> 😽 🚳
22:					
	BSB	BBASARI	var	var	var
1	15	23	Val	vai	vai
2	18	27			
3	24	36			
4	22	40			
5	25	44			
6	29	49			
7	30	56			
8	32	54			
9	35	60			
10	38	68			
11	30	45			
12	36	54			
13	48	77			
14	44	80			
15	50	87			
16		97			
17	60	113			
18	64	108			
19	70	121			
20	76	135			

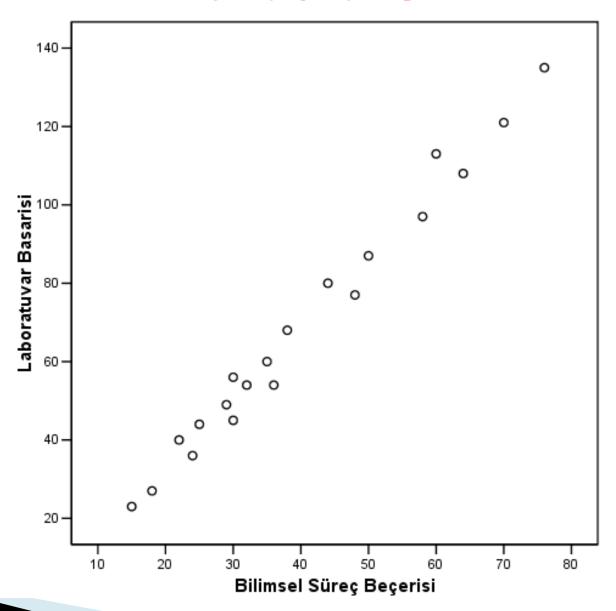
- Drneğimize ait veriler SPSS te girildikten sonra verileri özelliklerine bakılır. Verilerin sürekli olduğu görülmektedir. Normal dağılıma bakılır;
- Veriler normal dağılıma uygunsa Pearson Korelasyon Katsayısı,
- Normal dağılım göstermiyorsa Sperman Sıra Korelasyonu hesaplanır. Ayrıca doğrusal ilişki olup olmadığını görmek için serpilme diyagramı çizilir.







Serpilme (Saçılma) Grafiği



Saçılma grafiği incelendiğinde değişkenler arasında pozitif bir doğrusal ilişki olduğu görülmektedir. Veriler kontrol edildikten sonra korelasyon analizine geçilir.



Dinlediğiniz İçin Teşekkür Ederim...