

TEMEL İSTATİSTİK

Korelasyon II

Prof. Dr. Ezel Tavşancıl

1. Pearson Momentler arpım Korelasyon Katsayısı

- En az eşit aralıklı ölçek düzeyinde ölçülen iki sürekli deęişken arasındaki doğrusal ilişkinin derecesinin belirlenmesinde kullanılır.
- A ve B deęişkenleri arasında manidar bir ilişki var mıdır? sorusunun cevabı aranır.
- Sayıltılar:
 - İki deęişken arasındaki ilişkinin doğrusal olduğu varsayılır.
 - İki deęişken de sürekli olmalı (eşit aralık ya da oran ölçeęi düzeyinde)
 - Her iki deęişken de normal dağılımalıdır.

1. Pearson Momentler Çarpım Korelasyon Katsayısı

- Oran ölçeği düzeyindeki iki sürekli değişken arasındaki doğrusal ilişkiyi açıklamak üzere kullanılır.
- AMAÇ: Her bir Y değerinin her bir X değeri ile doğrusal biçimde ne kadar tutarlı bir eşleşme gösterdiğini açıklamak (Büyüköztürk ve diğerleri, 2012).

$$r = \frac{n \sum XY - \sum X \sum Y}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}}$$

n: eşleşmiş ölçüm çiftinin sayısı

Korelasyon Katsayısının Yorumlanması

1. BÜYÜKLÜK

Korelasyon katsayısının mutlak değeri «İlişkinin Gücü» sayfasında belirtilen sınır değerlerine bağlı yorumlanır [0.01-0.30-0.70 > düşük (low), orta (moderate), yüksek (high)]

2. YÖN

r değeri negatifse, değişkenlerden bir artarken diğeri azalmakta demektir. Pozitif ise iki değişkenin birlikte değişim yönünün aynı olduğunu (iki değişkenin birlikte artıp azaldığını) göstermektedir.

3. AÇIKLANAN VARYANS

Değişkenlerden birindeki değişimin bir kısmının diğeri değişken tarafından açıklandığını yüzde olarak ifade etmektedir. Korelasyon katsayısının karesidir.

Açıklanan Varyans

- Değişkenlerden birindeki değişimin ne kadarının diğer değişken tarafından açıklandığının yüzde olarak ifade edilebilmesini sağlar.
- Değişkenlerin birbirlerinde açıkladıkları varyans miktarı korelasyon katsayısının karesine eşittir ve buna determinasyon katsayısı denir.
- Örneğin Türkçe başarısı ile okuma hızı arasındaki korelasyon $r=0.80$ olsun. Buna göre determinasyon katsayısı $r^2 = 0.64$ 'dür. Buna göre, öğrencilerin okuma hızındaki toplam değişkenliğin %64'ü Türkçe dersindeki başarılarından kaynaklanmaktadır. Ya da öğrencilerin Türkçe dersindeki başarılarının %64'ü okuma hızlarındaki değişimden kaynaklanmaktadır; biçiminde yorumlanabilir. Ayrıca geriye kalan %36'lık varyansın da başka değişkenlerden kaynaklandığı söylenebilir

Örnek

Okul Öncesi Dönem Çocuklarında Fiziksel Ve İlişkisel Saldırganlığın Ebeveyn Tutumları Açısından İncelenmesi

Bulgular

Okul öncesi dönem çocuklarında görülen Fiziksel Saldırganlık (FS) ve İlişkisel Saldırganlık (İS) ile Anne Demokratik Tutumu (ADT), Baba Demokratik Tutumu (BDT), Anne Otoriter Tutumu (AOT), Baba Otoriter Tutumu (BOT), Anne Aşırı Koruyucu Tutumu (AAKT), Baba Aşırı Koruyucu Tutumu (BAKT) ve Anne İzin Verici Tutumu (AİVT), Baba İzin Verici Tutumu (BİVT) alt boyutları arasındaki ilişkiyi belirlemek için yapılan pearson korelasyon testi sonuçları Tablo 2. de belirtilmiştir.

Tablo 2: Fiziksel ve İlişkisel Saldırganlık ile Demokratik, Otoriter, Aşırı Koruyucu ve İzin Verici Anne-Baba Tutumlarına Ait Korelasyon Değerleri

		ADT	AOT	AAKT	AİVT	BDT	BOT	BAKT	BİVT
FS	r	-0,032	0,175**	0,027	-0,119*	-0,068	0,216**	0,004	-0,075
	p	0,577	0,002	0,645	0,040	0,244	0,000	0,952	0,195
İS	r	0,086	0,181**	0,008	0,004	-0,051	0,209**	0,025	-0,067
	p	0,135	0,002	0,887	0,948	0,380	0,000	0,661	0,249

*p< 0,05 **p< 0,01

Tablo 2 incelendiğinde, fiziksel saldırganlık ile annenin otoriter tutumu ($r=0,175$, $p<0,01$) ve babanın otoriter tutumu ($r= 0,216$, $p<0,01$) arasında pozitif ve istatistiksel açıdan anlamlı bir ilişki olduğu görülmektedir. Buna göre otoriter anne ve baba tutumuna sahip olan çocuklarda fiziksel saldırganlığın yüksek olduğu belirlenmiştir. Ayrıca fiziksel saldırganlık ile annenin izin verici tutumu arasında

negatif yönde ve istatistiksel açıdan anlamlı bir ilişki bulunmuştur ($r= -0,119$, $p<0,05$).

Fiziksel saldırganlık ile annenin demokratik tutumu ($r= -0,032$, $p>0,05$), aşırı koruyucu tutumu ($r=0,027$, $p>0,05$), babanın demokratik tutumu ($r= -0,068$, $p>0,05$), aşırı koruyucu tutumu ($r=0,004$, $p>0,05$) ve izin verici tutumu ($r= -0,075$, $p>0,05$) arasında istatistiksel açıdan anlamlı bir ilişkiye rastlanmamıştır.

İlişkisel saldırganlık ile annenin otoriter tutumu ($r= 0,181$, $p<0,01$) ve babanın otoriter tutumu ($r=0,209$, $p<0,01$) arasında pozitif yönde ve istatistiksel açıdan anlamlı bir ilişki olduğu bulunmuştur. İlişkisel saldırganlık davranışı yüksek olan çocukların, anne ve babalarının otoriter tutumu da nispeten daha yüksektir.

İlişkisel saldırganlık ile annenin demokratik ($r= 0,086$, $p>0,05$), aşırı koruyucu ($r=0,008$, $p>0,05$), izin verici tutumu ($r=0,004$, $p>0,05$), babanın demokratik ($r= -0,051$, $p>0,05$), aşırı koruyucu ($r=0,025$, $p>0,05$) ve izin verici ($r= -0,067$, $p>0,05$) tutumu arasında istatistiksel açıdan anlamlı bir ilişkiye rastlanmamıştır.

Bu sonuçlar değerlendirildiğinde; anne ve babanın sergilediği otoriter tutumun fiziksel saldırganlığı anlamlı düzeyde artırdığı, annenin izin verici tutumunun ise fiziksel saldırganlığı azalttığı, diğer anne baba tutumları ile fiziksel ve ilişkisel saldırganlık arasında anlamlı ilişki olmadığı şeklinde yorumlanabilir.

Örnek

Annelerin Kabul Red Düzeyi İle Çocuklarının Empati Becerisi Arasındaki İlişkinin İncelenmesi

2007 yılında İstanbuldaki ilköğretim okullarının 4., 5. ve 6. sınıflarına devam eden 387 öğrenci ile çalışılmış.

İşlem

Çocukların Kişisel Bilgi Formları ve Empati ölçekleri çocuklar tarafından doldurulmuştur. Ebeveyn Kabul-red Ölçeği (Anne Formu) çocukların anneleri tarafından doldurulmuştur.

Verilerin Analizi

Demografik veriler için, kişisel bilgilerin frekans ve yüzde dağılımları hesaplanmıştır. Ebeveyn kabul- red düzeyi ile çocukların empati düzeyi arasındaki ilişkiyi ölçmek amacıyla Pearson Momentler Çarpımı Korelasyon tekniği uygulanmıştır.

Bulgular

Tablo 2. Annelerin kabul- red düzeyi ile çocukların empati düzeyleri arasındaki ilişki

	N	X	ss	r
Annelerin Kabul Red Düzeyi	387	67.09	8.34	.042
Çocukların Empati Düzeyi	387	82.51	14.28	

Tablo 2 incelendiğinde, annelerin kabul red puanları ile çocukların empati puanları arasında, çok düşük ve anlamlı olmayan bir ilişki olduğu saptanmıştır ($r = .042$ $p > .01$). Bu sonuca göre, annelerin kabul red düzeyi ile çocukların empati düzeyi değişkenleri arasında bir ilişkinin olmadığı söylenebilir.

Pearson
Momentler
Çarpımı
Korelasyon
Katsayısı
Hesaplama
Örneği

Tablo Çalışma Saati ve Başarı Puanları

Birey	X	X ²	Y	Y ²	XY
1	2	4	3	9	6
2	4	16	5	25	20
3	1	1	1	1	1
4	5	25	3	9	15
n=4	ΣX=12 (ΣX)²=144	ΣX²=46	ΣY=12 (ΣY)²=144	ΣY²=44	ΣXY=42

$$r = \frac{n \sum XY - \sum X \sum Y}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}}$$

$$r = \frac{4(42) - (12)(12)}{\sqrt{[4(46) - (12)^2][4(44) - (12)^2]}} = \frac{24}{\sqrt{(40)(32)}} = \frac{24}{35.8} = 0.67$$

2. Spearman Brown Sıra Farkları Korelasyon Katsayısı (ρ , r_s)

- 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 Momentler Çapımı Korelasyon Katsayısı yerine kullanılır.

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

2. Spearman Brown Sıra Farkları Korelasyon Katsayısı (ρ , r_s)

- Üç durumda kullanılır:
 - Nesne ya da bireye ilişkin yargılar sıra değerleri ile gösterilmiş ise
 - Değişkenlerden biri sıralı diğeri eşit aralıklı ya da oranlı ölçek düzeyinde ise
 - En az eşit aralık düzeyinde olan ancak **normallik varsayımının karşılanmadığı** değişkenler var ise Spearman Brown Sıra Farkları Korelasyon Katsayısı kullanılır.

Spearman Brown Sıra Farkları Korelasyon Katsayısı Hesaplama Örneği II

Örnek Soru: İki resim öğretmenin sekiz öğrencinin yaptığı resimleri sıralamaları aşağıda verilmiştir. Öğretmenlerin sekiz öğrencinin resimlerini sıralamaları arasındaki ilişkiyi hesaplayınız.

Tablo 4.3. Örnek: Pearson sıralama farkları korelasyon katsayısının hesaplanması

Resim	A Öğretmenin Sıralaması	B Öğretmenin Sıralaması	F Sıralamalar arasındaki Fark	F ² Sıralama Far- kının Karesi
A	2	1	1	1
B	1	2	-1	1
C	4	5	-1	1
D	3	6	-3	9
E	6	4	2	4
F	5	3	2	4
G	7	8	-1	1
H	8	7	1	1
Toplam			0	22

$$r_s = 1 - \frac{6(22)}{8(8^2 - 1)} = 0.74$$

3. Nokta Çift Serili Korelasyon Katsayısı

- En az eşit aralık ölçek düzeyinde olan ve sadece iki kategorisi olan gerçek süreksiz değişken puanları arasındaki ilişkiyi incelemek için kullanılır.
- Başka bir ifade ile nokta çift serili korelasyon katsayısı biri sürekli diğeri iki kategorili olan gerçek bir süreksiz değişken arasındaki doğrusal ilişkiyi açıklama olanağı verir.

- $$r_{pb} = \left(\frac{\bar{Y}_p - \bar{Y}_q}{S_y} \right) \sqrt{pq}$$

p, süreksiz değişkenin oluşturduğu gruplardan birinde yer alan deneklerin toplam denek sayısı içindeki oranını, \bar{Y}_p , ise bu denekler için Y değişkenine ilişkin ortalamayı; q, ikinci grup için aynı oranı ve \bar{Y}_q , ise bu gruptaki deneklerin Y değişkeni ortalamasını gösterir.

Tablo : Cinsiyete Göre Kişilik Testi Puanları

Denek	Cinsiyet X	Kişilik Testi Puanı Y		Y ²
1	Erkek	50	$\sum Y_p = 182$	2500
2	1	38		1444
3	1	41		1681
4	1	53		2809
	Kız		$\sum Y_q = 321$	
5	2	60		3600
6	2	50		2500
7	2	44		1936
8	2	68		4624
9	2	53		2809
10	2	46	2116	
n=10	p= 4/10=0.40 q=6/10=0.60	$\sum Y = 503$		$\sum Y^2 = 26019$

Nokta Çift Serili Korelasyon Katsayısı Hesaplama Örneği

$$S_y = \sqrt{\frac{n \sum Y^2 - (\sum Y)^2}{n(n-1)}} = \sqrt{\frac{10(26019) - (503)^2}{10(10-1)}}$$

$$S_y = \sqrt{\frac{(260190) - (253009)}{90}} = \sqrt{\frac{7181}{90}} = 8.93$$

Hesaplanan standart sapma değerini formüldeki diğer terimlerle birlikte yerine koyalım.

$$r_{pb} = \left(\frac{45.50 - 53.50}{8.93} \right) \sqrt{(0.40)(0.60)} = \frac{-8}{8.93} \sqrt{0.24}$$

$$r_{pb} = (-0.90) (0.49) = -0.44$$

4. ift Serili Korelasyon

- ift serili korelasyon katsayısı, srekli bir deęişken ile gerekte srekli ancak yapay olarak sreksiz hale getirilen iki kategorili bir deęişken arasındaki ilişki miktarını hesaplamada kullanılır.
- Eęitimde başarı testlerinin geliştirilmesi sürecinde madde ayırt edicilięinin belirlenmesinde sıka kullanılan bir tekniktir.

$$r_b = \left(\frac{\overline{Y}_p - \overline{Y}_q}{S_Y} \right) \frac{pq}{y}$$

p= sresiz deęişkenin birinci kategorisindeki ölçm sayısının toplam içindeki oranı

q= sresiz deęişkenin ikinci kategorisindeki ölçm sayısının toplam içindeki oranı **y** normal dağılım eğrisi altında kalan alanda p ve q'yu

ayırır ordinat yüksekliğidir (Kitap Ek-D).

\overline{Y}_p ve \overline{Y}_q yapay sresiz deęişkenlerin srekli deęişkene ait

ortalamları S_Y srekli deęişkenlere ait tüm ortalamların standart sapmasını belirtir.

Kitaptaki Ek-D: Standart Normal Dağılım Eğrisindeki P ve Q Alanlarını Ayıran Ordinatlar

STANDART NORMAL DAĞILIM EĞRİSİNDE P VE Q ALANLARINI AYIRAN
ORDİNATLAR

p	q	y	pq/y	p	q	y	pq/y
.01	.99	.027	.372	.26	.74	.324	.593
.02	.98	.048	.405	.27	.73	.331	.596
.03	.97	.068	.428	.28	.72	.337	.599
.04	.96	.086	.446	.29	.71	.342	.602
.05	.95	.103	.461	.30	.70	.348	.604
.06	.94	.119	.474	.31	.69	.353	.606
.07	.93	.134	.485	.32	.68	.358	.609
.08	.92	.149	.495	.33	.67	.362	.612
.09	.91	.162	.504	.34	.66	.366	.612
.10	.90	.176	.513	.35	.65	.370	.614
.11	.89	.188	.521	.36	.64	.374	.616
.12	.88	.200	.528	.37	.63	.378	.617
.13	.87	.212	.535	.38	.62	.381	.619
.14	.86	.223	.541	.39	.61	.384	.620
.15	.85	.233	.547	.40	.60	.386	.621
.16	.84	.243	.552	.41	.59	.389	.622
.17	.83	.253	.558	.42	.58	.391	.623
.18	.82	.262	.563	.43	.57	.393	.624
.19	.81	.271	.567	.44	.56	.394	.625
.20	.80	.280	.572	.45	.55	.396	.625
.21	.79	.288	.576	.46	.54	.397	.626
.22	.78	.296	.580	.47	.53	.398	.626
.23	.77	.304	.583	.48	.52	.398	.626
.24	.76	.311	.587	.49	.51	.399	.627
.25	.75	.318	.590	.50	.50	.399	.627

Tablo . : Matematik Testi ve Testteki İkinci Maddeye Ait Puanlar

Madde Puanı	Test Puanı	Madde Puanı	Test Puanı	Madde Puanı	Test Puanı	Madde Puanı	Test Puanı
1	80	0	50	1	40	1	70
1	60	1	70	1	55	0	60
0	40	1	50	0	40	1	70
1	60	1	55	0	50	0	40
1	55	0	60	1	65	1	70

$$S = \sqrt{\frac{n \sum Y^2 - (\sum Y)^2}{n(n-1)}} = \sqrt{\frac{20(67600) - (1140)^2}{20(20-1)}}$$

$$S = \sqrt{\frac{1352000 - 1299600}{380}} = \sqrt{\frac{52400}{380}} = \sqrt{137.89} = 11.74$$

Şimdi bulunan değerleri Formül . yerine koyarak r_b 'yi hesaplayalım.

$$\frac{pq}{y} = 0.614 \quad (p = 0.65 \text{ değeri için tablodan } y = 0.37)$$

$$r_b = \left(\frac{\bar{Y}_p - \bar{Y}_q}{S_Y} \right) \frac{pq}{y} = \frac{61.54 - 48.57}{11.74} \cdot 0.614 = 0.67$$

Madde Puanı	Test Puanı		Madde Puanı	Test Puanı	
1	80		0	40	
1	60		0	50	
1	60	$p = \frac{13}{20} = 0.65$	0	60	$q = \frac{7}{20} = 0.35$
1	55		0	40	
1	70		0	50	
1	50	$\bar{Y}_p = \frac{800}{13} = 61.54$	0	60	
1	55		0	40	$\bar{Y}_q = \frac{340}{7} = 48.57$
1	40				
1	55				
1	65				
1	70				
1	70				
1	70				
		$\sum Y = 1140$			$\sum Y^2 = 67600$

5. Dörtlü (Phi) Korelasyon

- Dörtlü korelasyon katsayısı, sınıflama ölçeğinde ölçülmüş iki kategorili süreksiz iki değişken arasındaki ilişkinin incelenmesinde kullanılır.
- Başka bir ifade ile 2x2'lik bir veri matrisi olduğunda kullanılan bir tekniktir.

	Y1	Y2	
X1	a	b	a+b
X2	c	d	c+d
	a+c	b+d	n

$$\phi = \frac{cb - ad}{\sqrt{(a+c)(b+d)(a+b)(c+d)}}$$

Dörtlü (Phi) Korelasyon Katsayısı Hesaplama Örneği

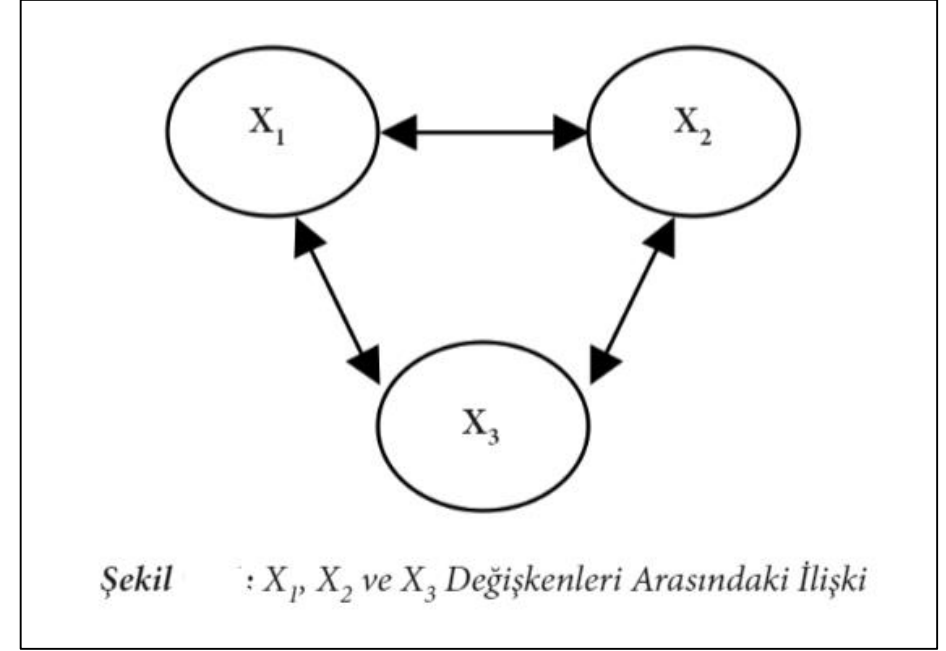
Kişilik Tipi	Kaygı Bozukluğu		Satır Toplamları
	Yaşıyor	Yaşamıyor	
A Tipi	70 (a)	30 (b)	a+b=100
B Tipi	20 (c)	100 (d)	c+d=120
Sütun Toplamları	a+c=90	b+d=130	Genel Toplam=220

$$\phi = \frac{cb - ad}{\sqrt{(a+c)(b+d)(a+b)(c+d)}} \quad \phi = \frac{(20)(30) - (70)(100)}{\sqrt{(90)(130)(100)(120)}} = -0.54$$

6. Kısmi Korelasyon

Bir ya da daha çok deęişkenin yol açtığı varyansların kontrol edilmesiyle iki deęişken arasındaki doğrusal ilişkiye denir. Kısmi korelasyonun iki temel varsayımı bulunmaktadır.

- Aralarındaki ilişki araştırılan deęişkenler ve bu deęişkenlerinin varyanslarının sabitleneceęi kontrol deęişkeni süreklidir.
- Analize dahil edilen tüm deęişkenlerin normal dağılım göstermesi gerekmektedir.



$$r_{123} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{(1 - r_{13}^2)(1 - r_{23}^2)}}$$

Kısmî Korelasyon Katsayısı Hesaplama Örneği

IQ değeri sabit tutulduğunda (kontrol edildiğinde) gelir ve başarı arasındaki korelasyon kaçtır?

Gelir (X_1)	Başarı(X_2)	IQ (X_3)
4	2	4
3	4	1
6	6	5
1	1	2
5	3	3

$r_{12}=0.757$ (gelir ve başarı),

$r_{13}=0.740$ (gelir ve IQ) ve

$r_{23}=0.411$ (başarı ve IQ)

Bu değerler Formül $r_{123} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{(1 - r_{13}^2)(1 - r_{23}^2)}}$ yerlerine konulduğunda,

$$r_{12.3} = \frac{0.757 - (0.740)(0.411)}{\sqrt{(1 - 0.740^2)(1 - 0.411^2)}} = 0.739 \text{ bulunur.}$$

$$r_{123} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{(1 - r_{13}^2)(1 - r_{23}^2)}}$$

7. Çoklu Korelasyon

- Çoklu korelasyon k tane bağımsız değişkenin doğrusal bir kombinasyonu ile bir bağımlı değişken arasındaki ilişkinin incelenmesinde kullanılan bir tekniktir. «R» sembolü ile gösterilir.

$$R_{Y.x_1x_2} = \frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1}r_{yx_2}r_{x_1x_2}}{1 - r_{x_1x_2}^2}$$

Tablo : Çoklu Korelasyon İçin Hipotetik Veriler

Hastalığın Şiddeti (X ₁)	Sağlık Sigortası Ödemesi (X ₂)	Tedavi Süresi (Y)
5	25	20
3	0	10
7	150	120
5	20	80
4	0	2
10	10	60
6	20	60
8	10	30
5	15	60
8	100	120
$\bar{X}_1 = 6.10$	$\bar{X}_2 = 35.00$	$\bar{Y} = 56.20$
$S_{X_1} = 2.02$	$S_{X_2} = 47.01$	$S_Y = 39.77$
	$r_{x_1Y} = .51$	
	$r_{x_2Y} = .83$	
	$r_{x_1x_2} = .33$	

$$R_{Y.x_1x_2} = \sqrt{\frac{.51^2 + .83^2 - 2(.51)(.83)(.33)}{1 - .33^2}}$$

$$R_{Y.x_1x_2} = .87 \text{ dir.}$$

KAYNAKLAR

- B y k zt rk, Ő.,  okluk,  . ve K kl , N. (2018). *Sosyal bilimler i in istatistik*. Ankara: Pegem Akademi.