

5.DERS KHİ-KARE METODU (X²)

Sayısal veri veya oranların belirli teorik bir oranla karşılaştırılmasında khi-kare (X²) testi uygulanır.

$$X^2 = \frac{(G-B)^2}{B} \quad (\text{Gözlenen-Beklenen})^2 / \text{Beklenen}$$

1. örnek: Bir genetikçi bezelyede 120 bitkide çiçek rengi üzerinde çalışmıştır. Çiçeklerin 84 tanesi kırmızı ve 36 tanesi de beyaz renklidir. Elde edilen çiçek renkleri 3:1 oranını doğrular mı? Bunu khi-kare testi ile kontrol ediniz?

HİPOTEZ= Elde edilen çiçek renkleri 3:1 oranına uymaktadır.

Tablo 1. Çiçek rengi oranları ve sayıları

Renkler	Oran	Gözlenen bitki sayıları (G)	Beklenen bitki sayıları (B)	(G-B)	(G-B) ² / B
Kırmızı	3	84	90	-6	0.40
Beyaz	1	36	30	+6	1.20
TOPLAM	4	120	120	0	1.60

Oranlar ve Gözlenen değerler tabloda verilir. Daha sonra, beklenen değerlerin bulunması için; toplam gözlenen bitki sayısı, toplam oran sayısına bölünür. (120/4=30) bulunur. Elde edilen 30 rakamı kırmızı çiçek oranı olan 3 ile çarpılarak (3x30=90) beklenen kırmızı çiçekli ve beyaz çiçek oranı olan 1 ile çarpılarak (1x30=30) beklenen beyaz çiçekli bitki sayısı belirlenir. Gözlenen ve beklenen değerler bulunduktan sonra formülde yerine konarak khi-kare değeri bulunur.

$$X^2 = \frac{(G_1-B_1)^2}{B_1} + \frac{(G_2-B_2)^2}{B_2} = \frac{(84-90)^2}{90} + \frac{(36-30)^2}{30} = 0.40+1.20= 1.60 \quad X^2=1.60$$

bulunur....

Gözlenen sayıların 3:1 oranına uyup uymadığı belirlemek için bulunan khi-kare değerinin teorik khi-kare değerleriyle karşılaştırılması gerekir.

Serbestlik derecesi= örnek sayısı-1 (n-1) den bulunur.

2 çiçek rengimiz vardır. Buna göre serbestlik derecesi=2-1=1 olarak bulunur.

Cetvelden 1 sayısının % 5 veya %1 değerine bakılır.

Khi-kare %5=3.84, %1=6.64

Hesaplanan Khi kare değerimiz 1.60 idi. Bu değer hem %5 hem de %1 değerinden küçük olduğu için HİPTEZ kabul edilir ve Elde edilen çiçek renkleri 3:1 oranına uymaktadır denir.... Şayet hesapladığımız Khi-kare değeri %5 ve %1 cetvel değerlerinden herhangi birinden büyük bulunsaydı o zaman hipotez ret edilecekti ve Elde edilen çiçek renkleri 3:1 oranına uymamaktadır denilecekti.....

Bazı Khi-kare cetvel değerleri

Serbestlik derecesi (SD)	%5	%1
1	3.84	6.64
2	5.99	9.21
3	7.81	11.34
4	9.49	13.28
5	11.07	15.09
6	12.59	16.81
7	14.07	18.48
8	15.51	20.09
9	16.92	21.67
10	18.31	23.29
11	19.68	24.73
12	21.03	26.22
13	22.36	27.69
14	23.68	29.14
15	24.99	30.58
16	26.29	32.00
17	31.41	37.57

2. örnek

Beş bitki karışımı ile bir yapay mera kurulmuştur. Bitkilerin karışıma katılma oranları ile 1m²'lik alanda sayılan bitki sayıları Tablo 1 'de verilmiştir.

HİPOTEZ: 5 yılda meradaki bitki karışımlarının oranı değişmemiştir.

Tablo 1. Merada sayılan bitkiler ve oranları

Bitki türü	Oran	Gözlenen bitki sayıları (G)	Beklenen bitki sayıları (B)	(G-B)	(G-B) ² / B
A	3	116	105	+11	1.1523
B	2	78	70	+8	0.9142
C	2	74	70	+4	0.2285
D	1	32	35	-3	0.2571
E	1	15	35	-20	11.4285
TOPLAM	9	315	315	00	14.2377

Yukarıdaki tabloyu doldurunuz ve meradaki bitki kompozisyonunun değişip değişmediğini Khi-Kare yöntemiyle kontrol ediniz?

Çözüm: Beklenen değerlerin bulunması için; toplam gözlenen bitki sayısı, toplam oran sayısına bölünür. (315/9=35) bulunur. Elde edilen 35 rakamı A, B, C, D ve E bitki türünün oranları ile ayrı ayrı beklenen bitki sayıları belirlenir. Gözlenen ve beklenen değerler bulunduktan sonra formülde yerine konarak khi-kare değeri bulunur.

$$X^2 = (G_1 - B_1)^2 / B_1 + (G_2 - B_2)^2 / B_2 + (G_3 - B_3)^2 / B_3 + (G_4 - B_4)^2 / B_4 + (G_5 - B_5)^2 / B_5 =$$
$$(116 - 105)^2 / 105 + (78 - 70)^2 / 70 + (74 - 70)^2 / 70 + (32 - 35)^2 / 35 + (15 - 35)^2 / 35 = X^2 = 14.2377 \text{ bulunur.....}$$

Bu çalışmada bitki tür sayısı 5 olduğuna göre; serbestlik derecesi = (5-1) = 4 olarak bulunur. Cetvelden 4 'ün % 5 değeri: 9.49 ve %1 değeri: 13.28 olarak belirlenir....

Hesapladığımız khi-kare değeri (14.2377) hem % 5 ve hem de % 1 cetvel değerinden büyük olduğundan hipotez ret edilir ve meranın 5 yıl öncesine göre çok önemli derecede değiştiği söylenir.

Hesapladığımız khi-kare değeri % 5 den büyükse **önemli fark** vardır denir..
Hesapladığımız khi-kare değeri % 1 den büyükse **çok önemli fark** vardır denir..

Örnek 3. Tablo 3. Oranları ve sayıları

Rakamlar	Oran	(G)	Beklenen bitki sayıları (B)	(G-B)	(G-B) ² / B	X ²
A	3	20	3*5=15	5	25/15	1.67
B	2	14	2*5=10	4	16/10	1.60
C	2	10	2*5=10	0	0/10	0
D	2	8	2*5=10	-2	4/10	0.40
E	1	4	1*5=5	-1	1/5	0.20
F	1	2	1*5=5	-3	9/5	1.80
G	1	2	1*5=5	-3	9/5	1.80
TOPLAM	12	60	60	00		7.47

$$60/12=5$$

N-1=7-1=6 %5 CETCEL: 12.59..... X² H < CETVEL, DEĞİŞİM YOK...

Örnek 4. Tablo 4. Oranları ve sayıları

Rakamlar	Oran	(G)	Beklenen bitki sayıları (B)	(G-B)	(G-B) ² / B	X ²
1	10	50	10*5=50	0	-	
2	6	30	6*5=30	0	-	
3	3	10	3*5=15	-5	25/15	1.67
4	1	10	1*5=5	+5	25/5	5
Toplam	20	100		00		6.67

$$100/20=5; \quad N-1=4-1=3 \quad \%5 \text{ CETCEL: } 7.81 \dots \dots \dots X^2 \text{ H} < \text{CETVEL, DEĞİŞİM YOK...}$$

Farklı yöntemlerle Khi-kare tersti çözümleri

Örnek 1. Bir ilaç firması A hastalığına karşı yeni bir ilaç bulmuştur. Bir kısım hastayı bu ilaçla, bir kısım hastayı da eski ilaçla tedavi altına alarak kendi ilacının etkinliğini araştırmışlardır. Bulgular aşağıda gösterilmiştir. İyileşme yönünden eski ilaç ile yeni ilaç arasında fark var mıdır?

İlaç	İyileşen	iyileşmeyen	Toplam
Yeni	21 (17.5)	27 (30.5)	48
Eski	11 (14.5)	29 (25.5)	40
Toplam	32	56	88

Hipotez: iyileşme yönünden eski ilaç ile yeni ilaç arasında fark yoktur.

Çözüm: Beklenen frekansın bulunması: oran ve orantı yolu ile beklenen frekanslar bulunur. “ tedavi edilen toplam 88 hastadan 32 hasta iyileşirse, yeni ilaçla tedavi edilen 48 hastadan kaç hastanın iyileşmesi beklenir?” şeklinde orantı kurulur. Buna göre;

**88 hastadan 32 hasta iyileşirse,
48 hastadan x sayıda hasta iyileşir
 $x=(48 \times 32)/88=17.5$ bulunur.**

Diğer oranlarda benzer şekilde bulunabilir. Ancak pratik olarak yeni iyileşen beklenenini (17.5 bulundu) aynı satırdaki 48 değerinden çıkartılarak yeni iyileşmeyen beklenenini (30.5) bulunur. Aynı şekilde 30.5, 56 dan çıkarılırsa eski iyileşmeyen gözünün beklenen frekansı 25.5 bulunur. 25.5, 40'dan çıkartılarak da eski iyileşen gözünün frekansı 14.5 bulunur.... Yukarıdaki tabloda parantez içindeki değerler beklenen değerler olarak tabloya eklenir.....

$X^2 = (G1-B1)^2/B$ formülünden;

$$X^2 = (21-17.5)^2/17.5 + (27-30.5)^2/30.5 + (11-14.5)^2/14.5 + (29-25.5)^2/25.5 = X^2=2.427 \text{ bulunur.....}$$

Serbestlik derecesi: (sıra sayısı-1) x (kolon sayısı-1)= (2-1)x (2-1)= 1 %5 karşılığı=3.841

Khi-kare hesaplanan < khi-kare cetvel olduğunda; (2.427 < 3.841)

Hipotez kabul edilir. Ve : iyileşme yönünden eski ilaç ile yeni ilaç arasında fark yoktur. Denir....

Örnek 2. Bir sağlık idarecisi difteri-boğmaca karma aşısı satın alacaktır. Piyasada ayrı firmalara ait 4 aşı vardır ve idareci en etkin olanını seçmek istemektedir. Bunun için bir araştırma yaparak bütün aşıları uygulamış ve sonuçları şöyle bulunmuştur. Bulgulara göre aşılar arasında fark vardır. Bunu khi-kare yöntemiyle test ediniz?

Aşı	korunan	korunmayan	Toplam
A	82 (81)	41 (42)	123
B	70 (62)	24 (32)	94
C	45 (43)	20 (22)	65
D	48 (59)	42 (31)	90
toplam	245	127	372

Hipotez: koruyuculuk yönünden aşılar arasında fark yoktur..

Çözüm: beklenen frekanslar oran-orantı yoluyla bulunur. A aşısının beklenen frekansını bulalım;

372 kişiden 245 kişi korunursa

123 kişiden x sayıda kişi korunur

$x = (123 \times 245) / 372 = 81$. bulunan bu sayı a aşısının toplamı olan 123 den çıkartılırsa A aşısının korunmayan gözündeki beklenen frekans bulunur ($123 - 81 = 42$). Diğer aşıların beklenen frekansları da benzer yolla bulunarak parantez içersinde yukarıdaki tabloya eklenmiştir.

	<u>Korunan</u>	<u>Korunmayan</u>
B aşısı:	$(245 \times 94) / 372 = 62$	$94 - 62 = 32$
C aşısı:	$(245 \times 65) / 372 = 43$	$65 - 43 = 22$
D aşısı:	$(245 \times 90) / 372 = 59$	$90 - 59 = 31$

$X^2 = (G1 - B1)^2 / B$ formülünden;

A aşısı için; $X^2 = (82 - 81)^2 / 81 + (41 - 42)^2 / 42 = 0.036$

B aşısı için; $X^2 = (70 - 62)^2 / 62 + (24 - 32)^2 / 32 = 3.032$

C aşısı için; $X^2 = (45 - 43)^2 / 43 + (20 - 22)^2 / 22 = 0.275$

D aşısı için; $X^2 = (48 - 59)^2 / 59 + (42 - 31)^2 / 31 = 5.954$

Toplam $X^2 =$ 9.297

Serbestlik derecesi: (sıra sayısı-1) x (kolon sayısı-1) = (4-1) x (2-1) = 3 %5 karşılığı = 7.815

Khi-kare hesaplanan > khi-kare cetvel olduğunda; (9.297 > 7.815)

Hipotez RED edilir. Ve : koruyuculuk yönünden aşılar arasında fark VARDIR denir...

Hangi aşının daha iyi olduğu daha sonraki konularda F testi ile kontrol edilecektir....

Örnek 3. Bir bölgede görülen kızamık vakalarının mevsimlere göre dağılımı aşağıda verilmiştir. Kızamık görülme bakımından mevsimler arasında fark var mıdır. Bunu Khi-kara yöntemiyle test ediniz?.

Mevsimler	Kızamık vaka sayısı
İlkbahar	210 (225)
Yaz	165 (225)
Sonbahar	225 (225)
Kış	300 (225)
toplam	900

Hipotez: kızamık görülme sayısı bakımından mevsimler arasında fark yoktur.

Çözüm: Beklenen frekans şu şekilde bulunur. Hipotezimiz: “kızamık görülme sayısı bakımından mevsimler arasında fark yoktur.” Olduğuna göre, kızamık vakası her mevsimde aynı sayıda görülmesi gerekir. Bu nedenle toplam vaka sayısı 900 vaka 4’e(grup sayınsa) bölünerek beklenen frekanslar (900/4=225) belirlenir. Ve yukarıdaki tabloda parantez içersinde yerine konur....

$X^2 = (G1-B1)^2/B$ formülünden;

$$X^2 = (210-225)^2/225 + (165-225)^2/225 + (225-225)^2/225 + (300-225)^2/225 = 42$$

Serbestlik derecesi: grup sayısı-1= 4-1=3

Tabloda % 5 için khi-kara değer= 7.82

Khi-kare hesaplanan > khi-kare cetvel olduğunda; (42 > 7.82)

Hipotez RED edilir. Ve : : kızamık görülme sayısı bakımından mevsimler arasında fark VARDIR denir...

Örnek 4. 70 öğrencinin istatistik dersine ait başarı durumları aşağıda verilmiştir. Başarı durumu cinsiyete bağımlı mıdır?

Başarı durumu	kız	erkek	Toplam
Başarılı	20 (17.86)	30 (32.14)	50
başarısız	5 (7.14)	15 (12.86)	20
Toplam	25	45	70

Hipotez: başarı durumu cinsiyete bağımlı değildir.

Çözüm: Beklenen frekansın bulunması: oran ve orantı yolu ile beklenen frekanslar bulunur. “toplam 70 kişiden 50 kişisi başarılı ise, 25 kızdan kaç kişi başarılıdır? şeklinde orantı kurulur. Buna göre;

**70 kişiden 50 kişi başarılı ise,
25 kızdan x kişi başarılıdır
 $x=(25 \times 50)/70=17.86$ bulunur.**

Diğer oranlarda benzer şekilde bulunabilir. Ancak pratik olarak başarılı kız beklenenini (17.86 bulundu) aynı satırdaki 50 toplam değerinden çıkartılarak başarılı erkek beklenenini (32.14) bulunur..... Yukarıdaki tabloda parantez içindeki değerler beklenen değerler olarak tabloya eklenir.....

$X^2 = (G1-B1)^2/B$ formülünden;

$X^2 = (20-17.86)^2/17.86 + (30-32.14)^2/32.14 + (5-7.14)^2/7.14 + (15-12.86)^2/12.86 = X^2=4.59$ bulunur.....

Serbestlik derecesi: (satır sayısı-1) x (kolon sayısı-1)= (2-1)x (2-1)= 1 %5 karşılığı=3.841

**Khi-kare hesaplanan > khi-kare cetvel olduğunda; (4.59 > 3.841)
Hipotez RET edilir. Ve : başarı durumu cinsiyete bağımlıdır denir.....**

Örnek 5. Çocuklarda doğum ağırlığı ile ilk diş çıkarma yaşı değerleri aşağıdaki tabloda verilmiştir. İlk diş çıkarma yaşı doğum ağırlığına bağımlıdır? Bunu khi-kare yöntemiyle test ediniz?

Doğum ağırlığı (kg)	İlk diş çıkarma yaşı (hafta)			
	İlk 25 hafta	26-30 hafta	31 ve yukarısı	toplam
1.5 ve aşağısı	4 (10)	11 (7.5)	5 (2.5)	20
1.5-2.5	16 (17.5)	15 (13.1)	4 (4.4)	35
2.6-3.5	20 (12.5)	4 (9.4)	1 (3.1)	25
toplam	40	30	10	80

Hipotez: Diş çıkarma yaşı doğum ağırlığına bağımlı değildir...

Çözüm: beklenen frekanslar oran-orantı yoluyla bulunur. 1. ağırlığın beklenen frekansını bulalım;

80 kişiden 40 kişi ilk 25 haftada diş çıkartıyorsa

20 kişiden x sayıda kişi diş çıkartır

$x = (20 \times 40) / 80 = 10$ diğerleri de aynı mantıkla bulunarak veriler elde edilir....

İlk 25 hafta= $X^2 = (4-10)^2/10 + (16-17.5)^2/17.5 + (20-12.5)^2/12.5 = 3.6+0.13+4.5=8.23$

26-30 hafta arası= $X^2 = (11-7.5)^2/7.5 + (15-13.1)^2/13.1 + (4-9.4)^2/9.4 = 1.6+0.28+3.1=4.98$

31 hafta ve yukarısı arası= $X^2 = (5-2.5)^2/2.5 + (4-4.4)^2/4.4 + (1-3.1)^2/3.1 = 2.5+0.04+1.4=3.96$

Genel khi kare toplamı= $8.23+4.98+3.96=17.17$ olarak bulunur...

Serbestlik derecesi: (sıra sayısı-1) x (kolon sayısı-1) = (3-1)x (3-1) = 4 %5 karşılığı=9.49

Khi-kare hesaplanan > khi-kare cetvel olduğunda; (17.17 > 7.815)

Hipotez RED edilir. Ve : Diş çıkarma yaşı doğum ağırlığına bağımlıdır denir.....

6.DERS

t - KONTROL METODU

Varyansın bilinmediği durumlarda populasyon ortalaması ile örnek ortalamasının karşılaştırılmasında (t) kontrol metodu kullanılır.

$t = \frac{d-P}{S_d}$ formülünden t değeri bulunur...

d= iki örnek ortalaması arasındaki fark

P= İki populasyon ortalaması arasındaki fark

S_d= ortalama farkın standart hatası

Ortalama farkın standart hatası (S_d) dört farklı durumda değişir ve farklı farklı hesaplanması gerekir.

1. Örnek sayıları ve varyanslar eşit ise: ($n_1=n_2$ ve $S^2_1=S^2_2$)

$S_d = \sqrt{(S^2_1+S^2_2)/n}$ veya $S_d = \sqrt{(2 \times S^2_1)/n}$

Serbestlik derecesi: $SD = 2 \times (n-1)$

2. Örnek sayıları farklı ve varyanslar eşit ise: ($n_1 \neq n_2$ ve $S^2_1=S^2_2$)

$S_d = \sqrt{(S^2 \times (n_1+n_2) / n_1 \times n_2)}$

Serbestlik derecesi: $SD = (n_1-1) + (n_2-1)$

3. Örnek sayıları eşit ve varyanslar farklı ise: ($n_1=n_2$ ve $S^2_1 \neq S^2_2$)

$S_d = \sqrt{(S^2_1+S^2_2)/n}$ veya $S_d = \sqrt{(2 \times S^2_1)/n}$

Serbestlik derecesi: $SD = (n-1)$

4. Örnek sayıları ve varyanslar farklı ise: ($n_1 \neq n_2$ ve $S^2_1 \neq S^2_2$)

$S_d = \sqrt{(S^2_1/n_1 + S^2_2/n_2)}$

Serbestlik derecesi: $SD = (n_1-1) + (n_2-1)$

1. örnek: özel hastane ve devlet hastanesinde doğan çocukların sarılık hastalığına yakalanma oranlarını karşılaştırmak üzere bir araştırma yapılmış ve sonuçlar aşağıdaki tabloda verilmiştir.

Tablo 1. hastalar için hesaplamalar.

Civciv ağırlıkları		İstatistik hesaplamalar	
Özel hastane (X ₁)	Devlet hastanesi (X ₂)	Özel hastane (X ₁)	Devlet hastanesi (X ₂)
9	8	$\sum X_1 = 130$	$\sum X_2 = 110$
17	15	$n_1 = 10$	$n_2 = 10$
14	11	$X_{1ort} = 13$	$X_{2ort} = 11$
13	11	$\sum X_1^2 = 1744$	$\sum X_2^2 = 1254$
15	9	$(\sum X_1)^2 / n = 1690$	$(\sum X_2)^2 / n = 1210$
10	12	$(\sum X_1 - X_{ort.1})^2 = 54$	$(\sum X_2 - X_{ort.2})^2 = 44$
11	11	$d = X_{1ort} - X_{2ort} = 2$	
13	10		
13	9		
15	14		

özel hastane ve devlet hastanesinde doğan çocukların sarılık hastalığına yakalanma oranları arasında fark olup olmadığını t testi ile kontrol ediniz? (örnek sayıları ve varyanslar eşit).

1. adım. X1 ve X2' nin ayrı ayrı varyansları bulunur.....

$S^2_1 = (\sum X_1 - X_{ort.1})^2 / n - 1 = 54 / 9 = 6$ veya $S^2_1 = (\sum X_1^2 - (\sum X_1)^2 / n) / n - 1 = (1744 - 1690) / 9 = 6$ bulunur.
(Varyansın 2 tane farklı yöntemle bulunduğunu hatırlayınız!!!)

$S^2_2 = (\sum X_2 - X_{ort.2})^2 / n - 1 = 44 / 9 = 4.89$ veya $S^2_2 = (\sum X_2^2 - (\sum X_2)^2 / n) / n - 1 = (1254 - 1210) / 9 = 4.89$ bulunur...
(Varyansın 2 tane farklı yöntemle bulunduğunu hatırlayınız!!!)

2. adım. d yani iki örnek ortalaması arasındaki fark bulunur.

$d = X_{1ort} - X_{2ort} = 13 - 11 = 2$ olarak bulunur..

3. adım. P yani İki populasyon ortalaması arasındaki fark bulunur (bunu soruda verilmedikçe daima 0 (sıfır) olarak alacaksınız.....)

4. adım. Sd yani ortalama farkın standart hatası hesaplanır.

Örnek sayıları ve varyanslar eşit ($n_1 = n_2$ ve $S^2_1 = S^2_2$) olarak verildiğinden aşağıdaki formül kullanılır.

$Sd = \sqrt{(S^2_1 + S^2_2) / n}$ veya $Sd = \sqrt{(2 \times S^2_1) / n}$ Serbestlik derecesi: $SD = 2 \times (n - 1)$

Buna göre; $Sd = \sqrt{(6 + 4.89) / 10} = \sqrt{10.89 / 10} = 1.04$ bulunur. Bulunan değerler formülde yerine konarak t değerimiz hesaplanır.

$t = d - P / Sd = 2 - 0 / 1.04 = 1.92$ bulunur.

5. adım. Hesaplanan t değeri serbestlik derecesi $2 \times (n - 1)$ 'den $2 \times (10 - 1) = 18$ olan cetvel değeriyle karşılaştırılır.

$t_{%5} = 2.10$, $t_{%1} = 2.87$

Hesaplanan t değerimiz %5 ve %1 t cetvel değerlerinden küçük olduğundan özel-devlet hastanelerinde sarılık hastalığına yakalanma oranları bakımından fark yoktur denir....

Şayet Hesaplanan t değerimiz %5 ve %1 t cetvel değerlerinden büyük çıksaydı o zaman özel-devlet hastanelerinde sarılık hastalığına yakalanma oranları bakımından farklılık olduğunu söylememiz gerekecekti.....

Bazı t cetvel değerleri

Serbestlik derecesi (SD)	%5	%1
1	12.70	63.65
2	4.30	9.93
3	3.18	5.84
4	2.77	4.60
5	2.57	4.03
6	2.45	3.71
7	2.36	3.50
8	2.31	3.36
9	2.26	3.25
10	2.23	3.17
11	2.20	3.11
12	2.18	3.06
13	2.16	3.01
14	2.15	2.98
15	2.13	2.95
16	2.12	2.92
17	2.11	2.89
18	2.10	2.87
19	2.09	2.86
20	2.086	2.85
21	2.080	2.83
22	2.074	2.82
23	2.069	2.81
24	2.064	2.79
25	2.060	2.78

2. örnek: İki hastaya aiti kalp atışı hızları çıkartılmış ve sonuçları aşağıdaki tabloda verilmiştir.

Tablo 1. İki hastaya ait hesaplamalar.

Kalp atışı hızları		İstatistik hesaplamalar	
1. hasta (X ₁)	2. hasta (X ₂)	1. hasta (X ₁)	2. hasta (X ₂)
20.79	21.22	$\sum X_1=104.53$	$\sum X_2=148.10$
21.47	21.47	$n_1=5$	$n_2=7$
20.18	20.96	$X_{1ort}=20.906$	$X_{2ort}=21.157$
20.96	21.56	$\sum X_1^2=2186.22$	$\sum X_2^2=3134.37$
21.13	21.39	$(\sum X_1)^2/n=2185.30$	$(\sum X_2)^2/n=3133.37$
	20.88	$(\sum X_1 - X_{ort.1})^2=0.9116$	$(\sum X_2 - X_{ort.2})^2=0.7226$
	20.62	$d= X_{1ort} - X_{2ort}=0.251$	

İki hastanın kalp atış hızları arasında fark olup olmadığını t testi ile kontrol ediniz? (örnek sayıları farklı ve varyanslar eşit).

1. adım. Örnek sayıları farklı olduğu için birleştirilmiş varyansın hesaplanması gerekir;

Birleştirilmiş varyans tablosu;

hastalar	n sayısı	Serbestlik derecesi	Ortalama farkın Kareler toplamı
1. hasta (X ₁)	5	4	0.9116
2. hasta (X ₂)	7	6	0.7226
toplam	-	10	1.6342

Birleştirilmiş varyans= $1.6342/10=0.1634$

2. adım. Sd yani ortalama farkın standart hatası hesaplanır.

Örnek sayıları farklı ve varyanslar eşit ($n_1 \neq n_2$ ve $S^2_1=S^2_2$) olarak verildiğinden aşağıdaki formül kullanılır;

$$Sd = \sqrt{(S^2 \times (n_1+n_2) / (n_1 \times n_2))} \quad \text{Serbestlik derecesi: } SD = (n_1-1) + (n_2-1)$$

Buna göre; $Sd = \sqrt{(S^2 \times (n_1+n_2) / (n_1 \times n_2))} = \sqrt{(0.1634 \times (5+7) / (5 \times 7))} = \sqrt{0.056} = Sd = 0.2367$ bulunur. Bulunan değerler formülde yerine konarak t değerimiz hesaplanır.

$$t = \frac{d - P}{Sd} = \frac{0.251 - 0}{0.2367} = 1.06 \text{ bulunur.}$$

6. adım. Hesaplanan t değeri serbestlik derecesi $SD = (n_1-1) + (n_2-1)$ 'den $(5-1) + (7-1)=10$ olan cetvel değeriyle karşılaştırılır.

$$t_{\%5} = 2.23 \quad t_{\%1} = 3.17$$

Hesaplanan t değerimiz %5 ve %1 t cetvel değerlerinden küçük olduğundan İki hastanın kalp atışları arasında fark olmadığı sonucuna varılır.

Şayet Hesaplanan t değerimiz %5 ve %1 t cetvel değerlerinden büyük çıksaydı o zaman İki hastanın kalp atışları arasında farklılık olduğunu söylememiz gerekecekti.....

t-kontrolü için örnekler;

1. örnek: iki farklı grubun ortalamaları aşağıdaki tabloda verilmiştir.

Tablo 1. iki farklı grup ve hesaplamalar.

Gruplar		İstatistik hesaplamalar	
Gurup I (X ₁)	Gurup II. (X ₂)	Gurup I (X ₁)	Gurup II. (X ₂)
2	4	$\sum X_1=30$	$\sum X_2=60$
4	8	$n_1=5$	$n_2=5$
6	12	$X_{1ort}=6$	$X_{2ort}=12$
8	16	$\sum X_1^2=220$	$\sum X_2^2=880$
10	20	$(\sum X_1)^2/n=180$	$(\sum X_2)^2/n=720$
		$(\sum X_1 - X_{ort.1})^2=40$	$(\sum X_2 - X_{ort.2})^2=160$
		$d= X_{1ort} - X_{2ort}=6$	

1. adım. X₁ ve X₂' nin ayrı ayrı varyansları bulunur.....

$S^2_1 = (\sum X_1 - X_{ort.1})^2/n-1 = 40/4 = 10$ veya $S^2_1 = (\sum X_1^2 - (\sum X_1)^2/n)/n-1 = (220-180)/4 = 10$ bulunur.
(Varyansın 2 tane farklı yöntemle bulunduğunu hatırlayınız!!!)

$S^2_2 = (\sum X_2 - X_{ort.2})^2/n-1 = 160/4 = 40$ veya $S^2_2 = (\sum X_2^2 - (\sum X_2)^2/n)/n-1 = (880-720)/4 = 40$ bulunur...
(Varyansın 2 tane farklı yöntemle bulunduğunu hatırlayınız!!!)

2. adım. d yani iki örnek ortalaması arasındaki fark bulunur.
 $d = X_{1ort} - X_{2ort} = 12 - 6 = 6$ olarak bulunur..

3. adım. P yani İki popülasyon ortalaması arasındaki fark bulunur (bunu soruda verilmedikçe daima 0 (sıfır) olarak alacaksınız.....)

4. adım. Sd yani ortalama farkın standart hatası hesaplanır.
Örnek sayıları ve varyanslar eşit ($n_1=n_2$ ve $S^2_1=S^2_2$) olarak verildiğinden aşağıdaki formül kullanılır.

$$S_d = \sqrt{(S^2_1 + S^2_2)/n} \text{ veya } S_d = \sqrt{(2 \times S^2_1)/n} \quad \text{Serbestlik derecesi: } SD = 2 \times (n-1)$$

Buna göre; $S_d = \sqrt{(10+40)/5} = \sqrt{50/5} = 3.16$ bulunur. Bulunan değerler formülde yerine konarak t değerimiz hesaplanır.

$$t = d - P / S_d = 6 - 0 / 3.16 = 1.89 \text{ bulunur.}$$

5. adım. Hesaplanan t değeri serbestlik derecesi $2 \times (n-1)$ 'den $2 \times (5-1)=8$ olan cetvel değeriyle karşılaştırılır.

$t_{\%5} = 2.31$,

Hesaplanan t değerimiz %5 t cetvel değerinden küçük olduğundan ortalamalar arasında fark yoktur denir....

2. örnek: İki gurubun ortalamaları aşağıdaki tabloda verilmiştir.

Tablo 1. İki gurup ortalamaları ve hesaplamalar.

Kalp atışı hızları		İstatistik hesaplamalar	
Gurup I (X ₁)	Gurup II hasta (X ₂)	Gurup I (X ₁)	Gurup II hasta (X ₂)
3	4	$\sum X_1=30$	$\sum X_2=60$
6	8	$n_1=4$	$n_2=5$
9	12	$X_{1ort}=7.5$	$X_{2ort}=12$
12	16	$\sum X_1^2=270$	$\sum X_2^2=880$
	20	$(\sum X_1)^2/n=225$	$(\sum X_2)^2/n=720$
		$(\sum X_1 - X_{ort.1})^2=45$	$(\sum X_2 - X_{ort.2})^2=160$
		$d= X_{1ort} - X_{2ort}=0.251$	

1. adım. Örnek sayıları farklı olduğu için birleştirilmiş varyansın hesaplanması gerekir;
Birleştirilmiş varyans tablosu;

hastalar	n sayısı	Serbestlik derecesi	Ortalama farkın Kareler toplamı
Gurup I.	4	3	45
Gurup II	5	4	160
toplam	-	7	205

Birleştirilmiş varyans= $205/7= 29.28$

2. adım. Sd yani ortalama farkın standart hatası hesaplanır.

Örnek sayıları farklı ve varyanslar eşit ($n_1 \neq n_2$ ve $S^2_1=S^2_2$) olarak verildiğinden aşağıdaki formül kullanılır;

$$Sd= \sqrt{(S^2_x (n_1+n_2/ n_1x n_2))} \quad \text{Serbestlik derecesi: } SD= (n_1-1) + (n_2-1)$$

Buna göre; $Sd= \sqrt{(S^2_x (n_1+n_2/ n_1x n_2))} = \sqrt{(29.28x(4+5)/(4x5))} = \sqrt{13.18}=Sd= 3.63$ bulunur. Bulunan değerler formülde yerine konarak t değerimiz hesaplanır.

$$t= d-P/Sd = 4.5-0/3.63= 1.24 \text{ bulunur.}$$

6. adım. Hesaplanan t değeri serbestlik derecesi $SD= (n-1) + (n_2-1)$ 'den $(4-1) + (5-1)=7$ olan cetvel değeriyle karşılaştırılır.

$$t \%5= 2.36$$

Hesaplanan t değerimiz %5 ve %1 t cetvel değerlerinden küçük olduğundan İki gurup arasında fark olmadığı sonucuna varılır.

7. DERS

F - KONTROL METODU

Daha önce ele aldığımız testlerde sadece 2 konunun karşılaştırılması yapılabiliyordu. F testinde ise 2 den daha fazla konunun karşılaştırılması yapılabilmektedir. Aynı zamanda F- kontrol metoduna “Varyans Analiz Yöntemi de denilmektedir.

F-testinin nasıl uygulandığının bir örnek ile açıklayalım....

1. örnek. 4 farklı aşı dozu 6 tekrarlamalı olarak çocuklar üzerine uygulanmıştır. Aşı dozlarının çocukların kan basıncı seviyesine etki edip etmediğini belirleyiniz. Varyans analiz tablosunu oluşturunuz. Aşı dozları arasında fark var ise ortalamaları LSD testi ile kontrol ediniz?

Aşı doz	TEKRARLAMALAR						TOPLAM	ORTALAMA
	I	II	III	IV	V	VI		
A	17.0	20.8	21.4	23.4	18.9	28.8	130.3	21.7
B	13.4	14.7	16.9	17.9	22.7	28.4	114.0	19.0
C	12.9	13.5	15.7	13.4	12.5	15.1	83.1	13.9
D	24.0	24.5	24.8	20.9	23.2	25.4	142.8	23.8
TOPLAM	67.3	73.5	78.8	75.6	77.3	97.7	470.2	

A) Varyans analizi

1. Düzeltme Faktörü= DF= $(470.2)^2/(24)= 9212$

tüm kan basıncı seviyelerini toplayacağız ve karesini alıp çıkan sonucu toplam rakam sayısına böleceğiz.....

2. Genel kareler toplamı= GKT= $(17.0^2+13.4^2+.....+15.1^2+25.4^2)-DF= 9807-9212=595.8$

Tüm kan basıncı seviyelerinin ayrı ayrı karelerini alıp toplayacağız ve çıkan sonucu Düzeltme Faktöründen çıkaracağız.....

3. Tekrarlamalar arası kareler toplamı=TKT= $(67.3^2+73.5^2+.....+97.7^2)/4-DF=9344.23-DF=132.23$

Her bir tekrarlamasının toplamının ayrı ayrı karelerini alıp toplayacağız ve çıkan sonucu aşı dozu sayısına böleceğiz. En sonunda elde edilen rakamı Düzeltme Faktöründen çıkaracağız.....

4. Aşılar arası kareler toplamı= (İAKT)= $(130.3^2+.....+142.8^2)/6-DF= 9545.76-DF=333.26$

Her bir aşı işlemlerinin toplamının ayrı ayrı karelerini alıp toplayacağız ve çıkan sonucu tekrarlamaya sayısına böleceğiz. En sonunda elde edilen rakamı Düzeltme Faktöründen çıkaracağız.....

5. Hata kareler toplamı= (HKT)= $2-3-4= 595.8-132.23-333.26= 130.31$

Genel kareler toplamından (2 numaralı işlem), Tekrarlamalar arası kareler toplamı (3 numaralı işlem) ve Aşılar arası kareler toplamı (4 numaralı işlem)’ni çıkaracağız.....

B) Varyans Analiz Tablosu

Varyans analiz tablosu

Varyasyon kaynağı	Serbestlik derecesi	Kareler toplamı	Kareler ortalaması	F değerleri
Genel	23	595.8	-	
Tekrarlamalar	5	132.23	26.45	3.044*
Aşı dozları	3	333.26	111.09	12.78**
Hata	15	130.31	8.69	

Tablonun doldurulmuş sıralaması

1. adım: Varyasyon kaynakları sırayla yazılır (A bölümünde 2, 3, 4 ve 5 nolu işlemlerin başlıkları)

2. adım: serbestlik dereceleri; Genel toplam rakam sayımız=24 bunun 1 eksiği yazılır... Tekrarlamamız 6 taneydi, bunun da 1 eksiği yazılır, aşı dozlarımız 4 taneydi bunun da 1 eksiği yazılır. Hata ise Tekrarlamaya ile aşı dozlarının Serbestlik derecelerinin çarpımıyla bulunur($3 \times 5=15$).

3. adım. Kareler toplamı: (A bölümünde bulduğumuz değerler sıraya tabloya yerleştirilir).

4. adım. Her bir Kareler toplamı kendi serbestlik derecelerine bölünerek Kareler ortalaması bulunur.

5. adım. tekrarlamaların ve aşı dozlarının kareler ortalaması ayrı ayrı Hatanın kareler ortalamasına bölünerek hesaplanan F değerleri bulunur.

c-) İşlemler arası farklılık

Aşı dozları arasındaki F hesaplanan değeri 12.78 bulunmuştur. Bu değer Cetveldeki F değeri ile karşılaştırılır. F 15-3 değerine bakılır. (15) hatanın serbestlik derecesi, (3) ise aşı dozlarının serbestlik derecesidir. Buna göre 15-3'ün F(%5) değeri=3.29 bulunur.

Hesaplanan F değeri cetvel F değerinden büyük olduğundan aşı dozları arasında fark var denir. Hangi aşı dozunun kan basıncı seviyesine daha fazla etki ettiğini bulmak için LSD (en küçük önemli fark) testi uygulanır...

$$LSD = t_{H.S.D} \times Sd \quad Sd = \sqrt{2 \times s^2 / n} \quad \text{varyans } (s^2) = \text{Hata kareler ortalaması}$$

$$SD = 15 \text{ için } t(\%5) = 2.13$$

$$Sd = \sqrt{2 \times s^2 / n} = Sd = \sqrt{2 \times 8.69 / 6} = 1.70$$

$$LSD = t_{H.S.D} \times Sd = 2.13 \times 1.70 = 3.62$$

Aşı dozu ortalamaları büyükten küçüğe doğru sıralanır ve her bir ortalamadan hesapladığımız LSD değerleri çıkarılır.

İşlemler	Ortalamalar	harfler	
D	23.8-3.62=20.18	a	bu duruma göre; D işlemi 20.18'e kadar aynı harftir
A	21.7-3.62=18.08	ab	(yani a) A işlemi 18.08'e kadar aynı harftir (yani ab)
B	19.0-3.62=15.38	b	B işlemi ise 15.38'e kadar aynı harftir (yani b)
C	13.9	c	C işlemi ise (c) harfi ile gösterilir....

Bu duruma göre; istatistiki olarak D ve A aşı dozları uygulanırsa en yüksek kan basıncı seviyesi elde edilir. A dozu daha düşük uygulanıp yüksek değer elde edildiğinden; A dozu D aşı dozuna tercih edilmelidir.....

F cetvel değeri (%5)'e göre hazırlanmıştır.

HKO	İşlem serbestlik dereceleri									
	1	2	3	4	5	6	7	8	9	10
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.97
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.86
12	4.75	3.88	3.49	3.26	3.11	3.00	2.92	2.85	2.80	2.76
13	4.67	3.80	3.41	3.18	3.02	2.92	2.84	2.77	2.72	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.77	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.70	2.64	2.59	2.55
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.96	2.81	2.70	2.62	2.55	2.50	2.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.55	2.48	2.43	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.52	2.45	2.40	2.35
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.47	2.40	2.35	2.30
23	4.28	3.42	3.03	2.80	2.63	2.52	2.45	2.38	2.32	2.28
24	4.26	3.40	3.01	2.78	2.62	2.51	2.43	2.36	2.30	2.26
25	4.24	3.38	2.99	2.76	2.60	2.49	2.41	2.34	2.28	2.24
26	4.22	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.30	2.25	2.20
28	4.20	3.34	2.95	2.71	2.56	2.44	2.36	2.29	2.24	2.19
29	4.18	3.33	2.93	2.70	2.54	2.43	2.35	2.28	2.22	2.18
30	4.17	3.32	2.92	2.69	2.53	2.42	2.34	2.27	2.21	2.16
32	4.15	3.30	2.90	2.67	2.51	2.40	2.32	2.25	2.19	2.14
34	4.13	3.28	2.88	2.65	2.49	2.38	2.30	2.23	2.17	2.12
36	4.11	3.26	2.86	2.63	2.48	2.36	2.28	2.21	2.15	2.10
38	4.10	3.25	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.07

2. örnek.

8 farklı muhasebecinin yıl sonu bilançoları 3 tekrarlamalı olarak hesaplanmış ve sonuçlar aşağıdaki tabloda verilmiştir. Muhasebecilerin bilançoları arasında fark olup olmadığını belirleyiniz. Varyans analiz tablosunu oluşturunuz. Muhasebecilerin bilançoları arasında fark var ise ortalamaları LSD testi ile kontrol ediniz?

MUHASEBECİLER	TEKRARLAMALAR			TOPLAM	ORTALAMA
	I	II	III		
A	5.7	13.3	16.3	35.3	11.8
B	5.0	11.6	14.3	30.9	10.3
C	5.3	6.7	10.3	22.3	7.4
D	9.7	9.9	15.0	34.6	11.5
E	6.7	12.7	15.7	35.1	11.7
F	6.0	5.0	9.3	20.3	6.8
G	2.7	5.0	4.7	12.4	4.1
H	8.3	9.7	6.7	24.7	8.2
TOPLAM	49.4	73.9	92.3	215.6	

A) Varyans analizi

1. Düzeltme Faktörü= DF= $(215.6)^2/(24)= 1936.8$

tüm bilançoları toplayacağız ve karesini alıp çıkan sonucu toplam rakam sayısına böleceğiz.....

2. Genel kareler toplamı= GKT= $(5.7^2+5.0^2+.....+4.7^2+6.7^2)-DF= 360.9$

Tüm bilançoların ayrı ayrı karelerini alıp toplayacağız ve çıkan sonucu Düzeltme Faktöründen çıkaracağız.....

3. Tekrarlamalar arası kareler toplamı=TKT= $(49.4^2+73.9^2+92.3^2)/8-DF=115.8$

Her bir tekrarlamasının toplamının ayrı ayrı karelerini alıp toplayacağız ve çıkan sonucu muhasebeci sayısına böleceğiz. En sonunda elde edilen rakamı Düzeltme Faktöründen çıkaracağız.....

4. İşlemler arası kareler toplamı= (İAKT)= $(35.3^2+.....+24.7^2)/3-DF= 164.3$

Her bir muhasebecinin bilançolarının toplamının ayrı ayrı karelerini alıp toplayacağız ve çıkan sonucu tekrarlamaya sayısına böleceğiz. En sonunda elde edilen rakamı Düzeltme Faktöründen çıkaracağız.....

5. Hata kareler toplamı= (HKT)= $2-3-4= 360.9-115.8-164.3= 80.8$

Genel kareler toplamından (2 numaralı işlem), Tekrarlamalar arası kareler toplamı (3 numaralı işlem) ve İşlemler arası kareler toplamı (4 numaralı işlem)'ni çıkaracağız.....

B) Varyans Analiz Tablosu

Varyans analiz tablosu

Varyasyon kaynağı	Serbestlik derecesi	Kareler toplamı	Kareler ortalaması	F değerleri
Genel	23	360.9	-	
Tekrarlamalar	2	115.8	57.9	10.03**
Muhasebeciler	7	164.3	23.47	4.07*
Hata	14	80.8	5.77	

Tablonun doldurulmuş sıralaması

1. adım: Varyasyon kaynakları sırayla yazılır (A bölümünde 2, 3, 4 ve 5 nolu işlemlerin başlıkları)

2. adım: serbestlik dereceleri; Genel toplam rakam sayımız=24 bunun 1 eksiği yazılır... Tekrarlamamız 3 taneydi, bunun da 1 eksiği yazılır, muhasebecimiz 8 taneydi bunun da 1 eksiği yazılır. Hata ise Tekrarlamaya ile muhasebecilerin Serbestlik derecelerinin çarpımıyla bulunur($2 \times 7=16$).

3. adım. Kareler toplamı: (A bölümünde bulduğumuz değerler sıraya tabloya yerleştirilir).

4. adım. Her bir Kareler toplamı kendi serbestlik derecelerine bölünerek Kareler ortalaması bulunur.
5. adım. tekrarlamaların ve muhasebecilerin kareler ortalaması ayrı ayrı Hatanın kareler ortalamasına bölünerek hesaplanan F değerleri bulunur.

c-) İşlemler arası farklılık

Muhasebecilerin bilançoları arasındaki F hesaplanan değeri 4.07 bulunmuştur. Bu değer Cetveldeki F değeri ile karşılaştırılır. F 14-7 değerine bakılır. (14) hatanın serbestlik derecesi, (7) ise muhasebecilerin serbestlik derecesidir. Buna göre 14-7'nin F(%5) değeri=2.77 bulunur.

Hesaplanan F değeri cetvel F değerinden büyük olduğundan muhasebecilerin yaptıkları bilançolar arasında fark var denir. Hangi muhasebecinin bilançosunun daha yüksek olduğunu bulmak için LSD (en küçük önemli fark) testi uygulanır...

$$LSD = t_{H.S.D} \times Sd \quad Sd = \sqrt{2 \times s^2 / n} \quad \text{varyans } (s^2) = \text{Hata kareler ortalaması}$$

$$SD=14 \text{ için } t(\%5)=2.15$$

$$Sd = \sqrt{2 \times s^2 / n} = Sd = \sqrt{2 \times 5.77 / 3} = 1.96$$

$$LSD = t_{H.S.D} \times Sd = 2.15 \times 1.96 = 4.2$$

İşlem ortalamaları büyükten küçüğe doğru sıralanır ve her bir ortalama için hesapladığımız LSD değerleri çıkarılır.

muhasebeciler	ortalamalar	LSD	Fark	Gruplar
A	11.8	4.2	7.6	a
E	11.7	4.2	7.5	a
D	11.5	4.2	7.3	ab
B	10.3	4.2	6.1	abc
H	8.2	4.2	4.0	abcd
C	7.4	4.2		bcd
F	6.8	4.2		cd
G	4.1			d

Bu duruma göre; istatistiki olarak A, H, D, B ve H numaralı muhasebecilerin bilançoları aynı gruptadır ve birbirlerinden farkı yoktur. Ancak, C, F ve G numaralı muhasebecilerin bilançoları farklıdır ve daha düşüktür denir.....

8. DERS

REGRESYON VE KORELASYON

Biri bağımlı diğeri de bağımsız olmak üzere incelenene 2 özellik arasında ilişkinin olup olmadığı regresyon analizi ile belirlenir. Her hangi bir X değişkeninin artması veya azalması durumunda her hangi bir Y değişkeni de azalır veya artar. Y değişkeni X'e göre azalmakta veya artmaktadır. Bu durumda regresyon analizi yapılır.

1. örnek. İnsan yaşı ile aylık dolar bazında yaptığı harcamalar aşağıdaki tabloda verilmiştir. Yaş ile aylık harcama arasındaki ilişkinin regresyon modelini oluşturunuz.

	Yaş (Xi)	Aylık harcama (Yi)	dx (Xi- \bar{X})	dy (Yi- \bar{Y})	dx ² (Xi- \bar{X}) ²	dy ² (Yi- \bar{Y}) ²	dx x dy
1	35	114	-20	-27	400	729	540
2	45	124	-10	-17	100	289	170
3	55	143	0	2	0	4	0
4	65	158	10	17	100	289	170
5	75	166	20	25	400	625	500
Toplam	275	705	00	00	1000	1936	1380
Ort.	$\bar{X}=55$	$\bar{Y}=141$					

1. Farkların çarpımlarının toplamı= $\sum dx \times dy = 1380$
2. (X) değişkeninin ortalamadan farklarının karelerinin toplamı= $\sum dx^2 = 1000$
3. (Y) değişkeninin ortalamadan farklarının karelerinin toplamı= $\sum dy^2 = 1936$
4. regresyon katsayısı= $b_{yx} = \frac{\sum dx \times dy}{\sum dx^2} = \frac{1380}{1000} = 1.38$ bulunur.

5. regresyon formülü= $\bar{Y} = a + b\bar{X}$ ' dir.

Buna göre; önce a sabit değerini bulmamız gerekir. Formülden a' yı çekersek=

$a = \bar{Y} - b\bar{X} = 141 - 1.38 \times (55) = 141 - 75.9 = 65.1$ buna göre regresyon formülümüz;

$\bar{Y} = a + b\bar{X} = \bar{Y} = 65.1 + 1.38x$ olarak bulunur. Yani; x yerine 1 koyarsak yaş 1 arttıkça aylık harcama da 66.48 dolar da harcama artmaktadır denir.....

Korelasyon

İkiside bağımlı olmak üzere incelenene 2 özellik arasında ilişkinin olup olmadığı korelasyon analizi ile belirlenir. Her hangi bir X1 değişkeni ile X2 değişkeni arasında herhangi bir ilişkinin olup olmadığı korelasyon analizi ile tespit edilir.

1. örnek. Bitkide verim ile boy arasında herhangi bir ilişkinin olup olmadığı incelenmek istenmektedir. Bu amaçla 9 bitki üzerinden yapılan ölçümler aşağıdaki tabloda verilmiştir. Buna göre bitkide verim ile boy arasında herhangi bir ilişkinin olup olmadığını kontrol ediniz?.

Tablo 2. Bitkide verim ile boy değerleri

	Verim (Xi)	Boy (Yi)	Xi x dYi	Xi ²	yi ²
1	1	8	8	1	64
2	4	7	28	16	49
3	6	6	36	36	36
4	8	13	104	64	169
5	12	1	12	144	1
6	14	2	28	196	4
7	16	11	176	256	121
8	22	3	66	484	9
9	26	4	104	676	16
toplam	109	55	562	1873	469
ortalamalar	12.11	6.11			

1) Ortalamadan farkların çarpımlarının toplamı:
 $\sum dx \times dy = \sum dxixdyi - (\sum xi) \times (\sum yi)/n = 562 - (109) \times (55)/9 = 652 - 666.11 = -104.11$

2) X1 değişkeni için ortalamadan farkların karelerinin toplamı:

$$\sum dx^2 = \sum xi^2 - (\sum xi)^2/n = 1873 - (109)^2/9 = 1873 - 132.11 = 552.89$$

3) Y değişkeni için ortalamadan farkların karelerinin toplamı:

$$\sum dy^2 = \sum yi^2 - (\sum yi)^2/n = 469 - (55)^2/9 = 469 - 336.11 = 132.99$$

4) korelasyon katsayısı:

$$r = \sum dx \times dy / \sqrt{\sum dx^2 \times \sum dy^2} = -104.11 / \sqrt{(552.59) \times (132.99)} = -104.11 / 271.06 = r = -0.384$$

Serbestlik derecesi = (n-2)'den bulunur. SD = (n-2) = (9-2) = 7

r %5 = 0.666, r %1 = 0.798 hesaplanan r değerimiz cetvel değerlerinden küçük olduğu için; verim ile boy arasında bir ilişki yoktur denir.....

Korelasyon katsayıları tablosu

Serbestlik derecesi	% 5	% 1
1	0.997	1.000
2	0.950	0.990
3	0.878	0.959
4	0.811	0.917
5	0.754	0.874
6	0.707	0.834
7	0.666	0.798
8	0.632	0.765
9	0.602	0.735
10	0.576	0.708
11	0.553	0.684
12	0.532	0.661
13	0.514	0.641
14	0.497	0.623
15	0.482	0.606
16	0.468	0.590
17	0.456	0.575
18	0.444	0.561
19	0.433	0.549
20	0.423	0.537
21	0.413	0.526
22	0.404	0.515
23	0.396	0.505
24	0.388	0.496
25	0.381	0.487
26	0.374	0.478
27	0.367	0.470
28	0.361	0.463
29	0.355	0.456
30	0.349	0.449
35	0.325	0.418
40	0.304	0.393
45	0.288	0.372

