OLASILIK VE İSTATİSTİK VİZE ÖDEVİ

Kovaryans

Matematik ve istatistikte kovaryans, iki rastgele değişken arasındaki ilişkinin bir ölçüsüdür. Kovaryans, iki farklı değişkenin sapmalarını tek bir değerde birleştirme işlemini yapar. Pozitif sapma birlikte artan bir doğrusal ilişkiyi, negatif sapma ise ters ilişkiyi gösterir. Korelasyon katsayısının aksine kovaryans birimler cinsinden ölçülür.
Birbirleriyle yüksek pozitif kovaryans sergilemeyen varlıkları seçerek, sistematik olmayan risk kısmen ortadan kaldırılabilir.
Kovaryans formülü:

$$\mathrm{Cov}\ (X, Y) = \frac{\sum(X_i - \bar{X})(Y_j - \bar{Y})}{n}$$

$$\mathrm{Cov}\ (X, Y) = \frac{\sum(X_i - \bar{X})(Y_j - \bar{Y})}{n\ - 1}$$

Formülde;

- Xi - X değişkeninin değerleri
- Yj - Y değişkeninin değerleri
- $\bar{X}$ - X değişkeninin ortalaması (ortalama)
- $\bar{Y}$ - Y değişkeninin ortalaması (ortalama)
- n - veri noktalarının sayısı

Cov(x,y)=0 ise bu değişkenler arasında ilişki yoktur.
Cov(x,y)>0 ise değişkenler aynı yönde
Cov(x,y)<0 ise farklı yönde değişim gösteriyor demektir.

Korelasyon Katsayısı

İki değişken arasındaki doğrusal ilişkinin gücü korelasyon katsayısı kullanılarak ölçülebilir. Korelasyon katsayısı her zaman -1 ile +1 arasındadır. 1, iki değişken arasında aynı yönde tam pozitif doğrusal ilişki olduğunu, -1 ise iki değişken arasında ters yönde tam negatif doğrusal ilişki olduğunu gösterir.

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[\, n\sum x^2 - (\sum x)^2\,]\,[\, n\sum y^2 - (\sum y)^2\,]}}$$

Kullanılan Örnek Veriler

Veride katılımcıların izledikleri reklam sayıları ile aldıkları ürün sayıları içeren iki değişken verilmektedir. Kovaryans ve korelasyon hesabını yazılan program kullanılarak yapılacaktır.

reklam	ürün
5	8
4	9
4	10
6	13
8	15

Örnek:Aşağıdaki tabloda verilen bir ana kütleden seçilmiş örneklem verileri için korelasyon katsayısını hesaplayınız.

x_i	10	15	16	22	25	28	29	40	45	50
y_i	5	12	15	20	10	26	25	38	42	47

Öncelikle x ve y değerleri için ortalamaları hesaplayalım.

$$\bar{X} = \frac{10+15+16+22+25+28+29+40+45+50}{10} = 28$$

$$\bar{Y} = \frac{5+12+15+20+10+26+25+38+42+47}{10} = 24$$

x_i	y_i	$x_i - \bar{X}$	$y_i - \bar{Y}$	$(x_i - \bar{X})^2$	$(y_i - \bar{Y})^2$	$(x_i - \bar{X})(y_i - \bar{Y})$
10	5	-18	-19	324	361	342
15	12	-13	-12	169	144	156
16	15	-12	-9	144	81	108
22	20	-6	-4	36	16	24
25	10	-3	-14	9	196	42
28	26	0	2	0	4	0
29	25	1	1	1	1	1
40	38	12	14	144	196	168
45	42	17	18	289	324	306
50	47	22	23	484	529	506
				Toplam 1600	Toplam 1852	Toplam 1653

Tablodaki değerlere göre formülü uygularsak

$$r = \frac{\sum(x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\sum(x_i - \bar{X})^2 \sum(y_i - \bar{Y})^2}} = \frac{1653}{\sqrt{1600.1852}} \cong \frac{1653}{1721,39} \cong 0,96$$

Sonuç +1 değerine yakın olduğu için aralarında aynı yönde güçlü bir ilişki olduğunu söyleyebiliriz.

Örnek:Aşağıda verilen iki tablonun kovaryansını hesaplayınız.

x_i	y_i
5	4
6	10
7	12
8	16
9	18

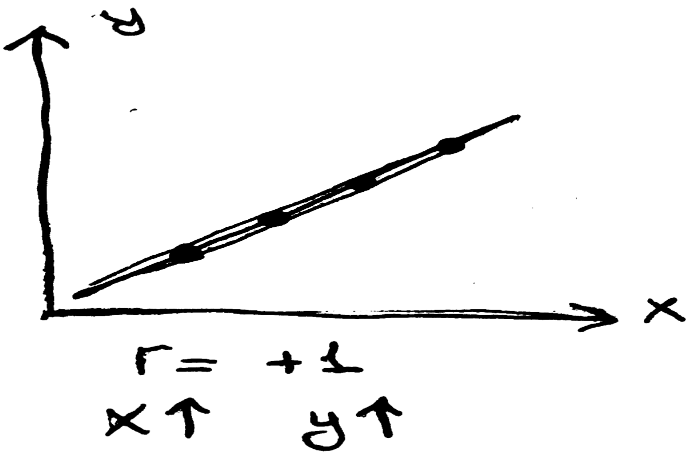
$$\mu_x = \frac{5+6+7+8+9}{5} = 7 \qquad \mu_y = \frac{6+9+11+16+18}{5} = 12$$

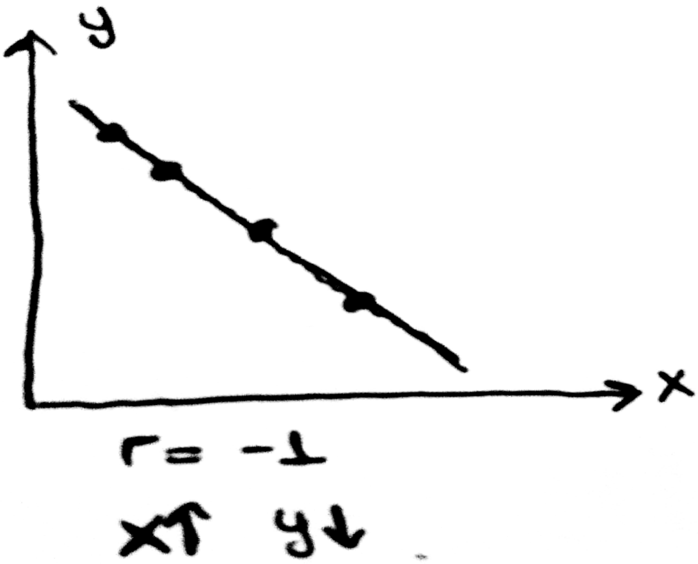
x_i	y_i	$x_i - \mu_x$	$y_i - \mu_y$	$(x_i - \mu_x)(y_i - \mu_y)$
5	6	-2	-6	12
6	9	-1	-3	3
7	11	0	-1	0
8	16	1	4	4
9	18	2	6	12

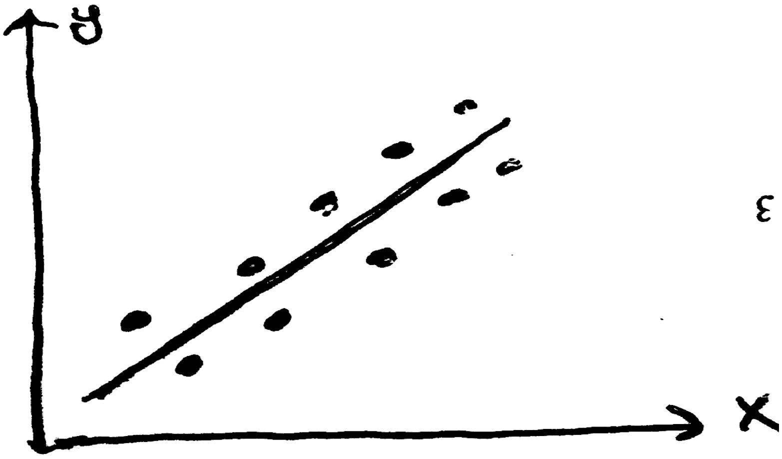
$$\boxed{Cov(x,y) = \frac{\Sigma (x_i - \bar{x}) * (y_i - \bar{y})}{N}} \qquad \frac{12+3+0+4+12}{5} = \frac{31}{5} = 6.2$$

Bulduğumuz sonuç pozitif olduğu için şunu söyleyebiliyoruz bu değişkenler arasında pozitif bir ilişki var.Yani biri artarken diğeri de artıyor.Kovaryans bu sözel ifadeyi bize sayısal olaraktа kanıtlamış oldu.

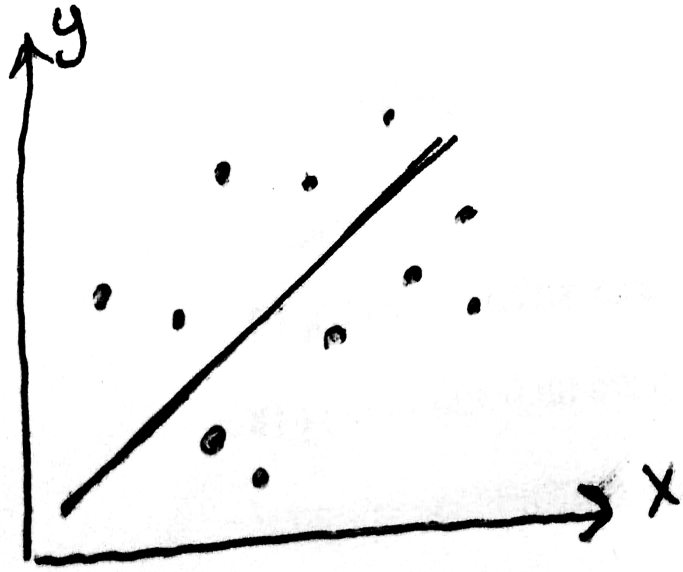
Bazı Serpme Diyagramları

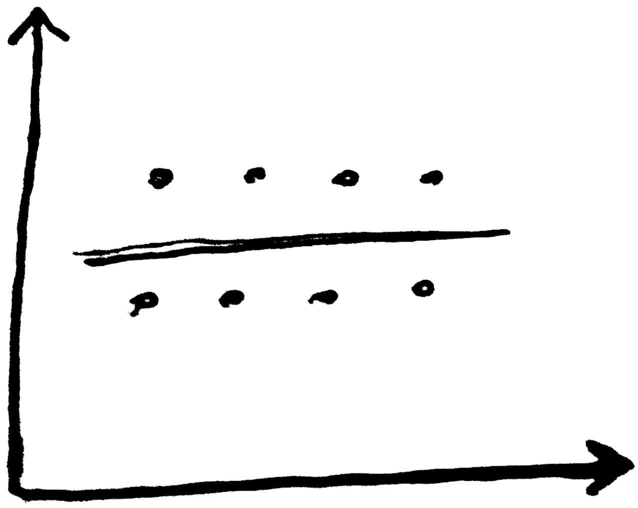




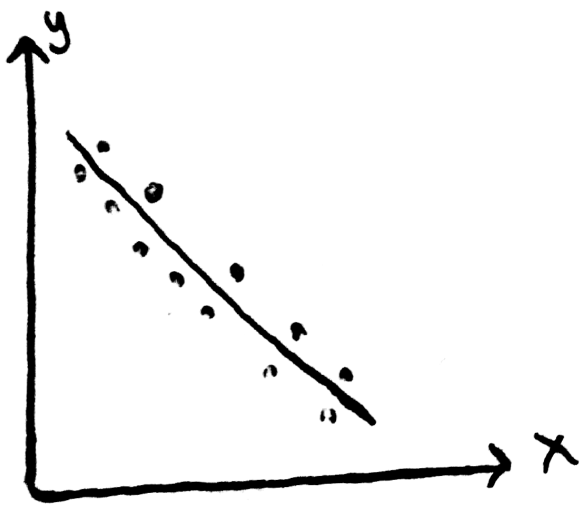


$0 < r < 1$

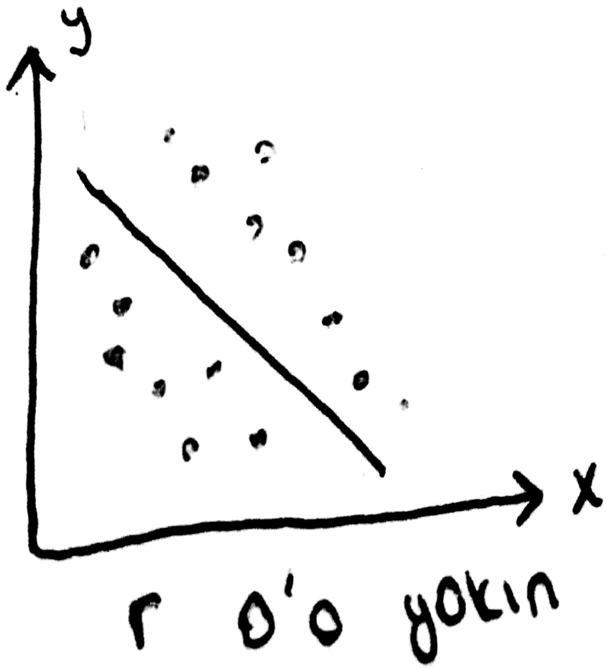




r=0



r, -1 e yakın



Korelasyon katsayısı ve Kovaryans Hesabı yapan Program

```java
import java.io.*;
import java.util.Scanner;

class GFG {
// kovaryans hesaplayan fonksiyon.
static float covariance(float arr1[],
                    float arr2[], int n)
{
    float sonuc1 = 0;

    for(int i = 0; i < n; i++)
        sonuc1 = sonuc1 + (arr1[i] - orta(arr1, n)) *
                        (arr2[i] - orta(arr2, n));
    return sonuc1 / (n - 1);
}
// ortalama bulan fonksiyon.
 static float orta(float arr[], int c)
{
    float sonuc = 0;

    for(int i = 0; i < c; i++)
        sonuc = sonuc + arr[i];

    return sonuc / c;
}
```

```java
static float correlationCoefficient(int X[],
                                    int Y[], int n)
{

    int sonuc_X = 0, sonuc_Y = 0, sonuc_XY = 0;
    int squareSum_X = 0, squareSum_Y = 0;

    for (int i = 0; i < n; i++)
    {
        // X dizisinin elemanları toplamı
        sonuc_X = sonuc_X + X[i];

        // Y dizisinin elemanları toplamı
        sonuc_Y = sonuc_Y + Y[i];

        // X[i] * Y[i] toplamı.
        sonuc_XY = sonuc_XY + X[i] * Y[i];

        // array elementlerinin kareleri toplamı.
        squareSum_X = squareSum_X + X[i] * X[i];
        squareSum_Y = squareSum_Y + Y[i] * Y[i];
    }

    // korelasyon katsayısı bulma formülü.
    float corr = (float)(n * sonuc_XY - sonuc_X * sonuc_Y)/
                 (float)(Math.sqrt((n * squareSum_X -
                 sonuc_X * sonuc_X) * (n * squareSum_Y -
                 sonuc_Y * sonuc_Y)));

    return corr;
}
```

Grup Üyeleri

Saltuk Yaşar
Özge Akdaş
Melike Yoğurtcu
Mehmet Salih Önder