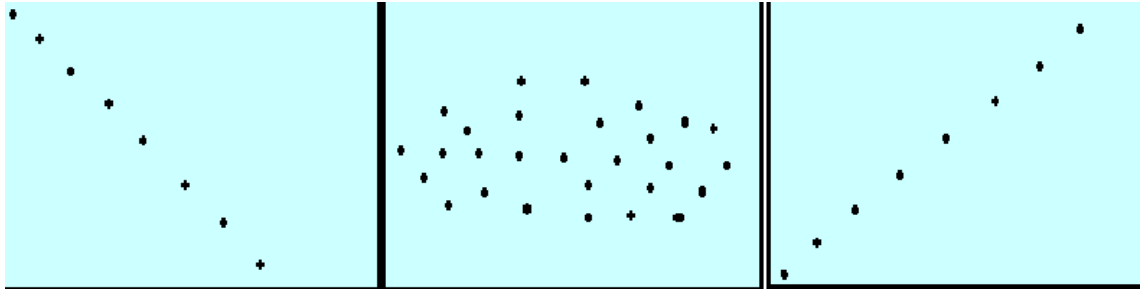


## KORELASYON, REGRESYON (Zaman Serisi)

### I. Korelasyon

Korelasyon katsayısı ile iki değişken arasındaki ilişkinin derecesi, yönü ve istatistiksel açıdan önemli olup olmadığı saptanır. Bu katsayı +1 ile -1 arasında değerler alır. Bulunan katsayı 0 ise iki değişken arasında hiçbir ilişki yok demektir. Katsayı -1 ya da +1'e yakınsa iki değişken arasında kuvvetli bir ilişki var demektir. Katsayı (-) ise iki değişkenden birinin değeri artarken diğerinin değeri azalıyor, katsayı (+) ise değişkenlerden birinin değeri artarken diğerinin de değeri artıyor ya da tersi demektir.

X ve Y değerlerine ilişkin n adet çiftten oluşan bir örnek için korelasyon katsayısı çeşitli formüllerle hesaplanır. Bunlardan pratikte en çok kullanılanı aşağıda verilmiştir.



a) Mükemmel negatif ilişki

b) İlişki yok

c) Mükemmel pozitif ilişki

Yukarıdaki saçılım grafikleri ;

- a) Değişkenlerden birisinin artışına bağlı olarak diğerinde azalma olan doğrusal ilişki olduğu
- b) İki değişken arasında ilişki olmadığı
- c) Değişkenlerden birisindeki artışa bağlı olarak diğerinde de artış olan doğrusal ilişki olduğu şeklinde açıklanır.

Korelasyonun katsayısının (r) gücü ile ilgili olarak aşağıdaki tanımlamalar yapılmıştır:

r=0.00 - 0.25 Çok zayıf ilişki

r =0.26 - 0.49 Zayıf ilişki

r =0.50 - 0.69 Orta ilişki

r =0.70 - 0.89 Yüksek ilişki

r =0.90 - 1.0 Çok yüksek ilişki

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

**Örnek:**

Aylar	Reklam Giderleri (X)	Ürün Satışları (y)	XY	X <sup>2</sup>	Y <sup>2</sup>
Ocak	1,00	101	101,00	1,00	10201
Şubat	1,25	116	145,00	11,56	13456
Mart	1,50	165	247,50	2,25	27225
Nisan	2,00	209	418,00	4,00	43681
Mayıs	2,50	264	660,00	6,25	69696
<b>Toplam</b>	<b>8,25</b>	<b>855</b>	<b>1571,50</b>	<b>15,06</b>	<b>164259</b>

$r=0,99$  Çok yüksek ilişki

**II. Regresyon (Basit Doğrusal)**

Korelasyon katsayısı ile X değişkenindeki değişikliğin Y değişkenini nasıl etkileyeceği belirlenemez. Sadece iki değişken arasında ilişki olup olmadığı eğer ilişki varsa bunun yönü ve derecesi belirlenir. Eğer iki değişken arasında ilişki varsa regresyon analizi ile X değişkeninin alacağı değere bağlı olarak y değişkeni tahmin edilebilir. Bu tahmin en küçük kareler yöntemine göre yapılır. Eldeki verileri en iyi şekilde temsil eden basit doğrusal regresyon denklemi hesaplanır. Genelde bir doğrunun denklemi  $y=a+bx$  şeklindedir. Bu denklem sonsuz sayıda doğruyu temsil eder. En küçük kareler yönteminde eldeki veri serisini en iyi temsil eden doğrunun denklemini bulmak gerekir. En küçük kareler yöntemine göre eldeki verileri en iyi şekilde temsil eden doğru, üzerindeki değerlerle bunlara karşı gelen gerçek değerler arasındaki farkların (sapmaların) kareleri toplamını en küçük (minimum) yapan doğrudur.

Bu doğruyu (trend doğrusu) bulmak için iki tane denklem vardır.

- $\sum y = Na + b \sum x$  n serideki veri sayısını ifade eder
- $\sum xy = a \sum x + b \sum x^2$

**Örnek: En küçük kareler yöntemi ile trend hesabı**

Yıllar	X	Süt Üretimi (ton)	xy	x <sup>2</sup>
2002	-7	9890	-69230	49
2003	-5	9633	-48165	25
2004	-3	9618	-28854	9
2005	-1	10240	-10240	1
2006	1	10279	10279	1
2007	3	10406	31218	9
2008	5	10562	52810	25
2009	7	10602	74214	49
<b>Toplam</b>	<b>0</b>	<b>81230</b>	<b>12032</b>	<b>168</b>

**Çözüm:**

$$81230=8a+0b$$

$$12032=0a+168b$$

$$a= 10153,75$$

$$b= 71,62$$

Trend denklemi  $Y=10153,75+71,62x$  olarak bulunur. Eğer 2010 yılı üretimini tahmin etmek istenirse x yerine 9 konularak Y hesaplanır ve 10798,33 ton bulunur