KORELASYON VE REGRESYON ANALİZİ

Bağımlı ve Bağımsız Değişkenler

Bağımsız değişken**,** ***x***, **bağımlı değişkendeki,** ***y,*** değişkenliği açıklar.
X değişkenindeki değişkenlik y değişkenindeki değişkenliği açıklar.

Bağımsız değişken(*x*) → Bağımlı değişken(*y*)

TV Ekran Sayısı ---- TV Satış Fiyatı

İşsizlik Oranı ------Suç Oranı

İlaç Dozu-------İyileşme süresi

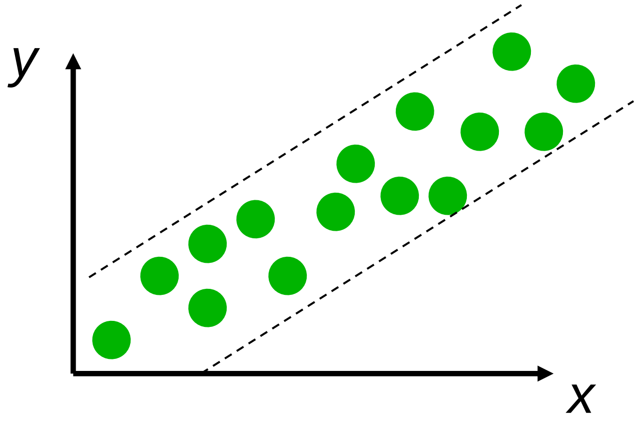
Korelasyon Analizi

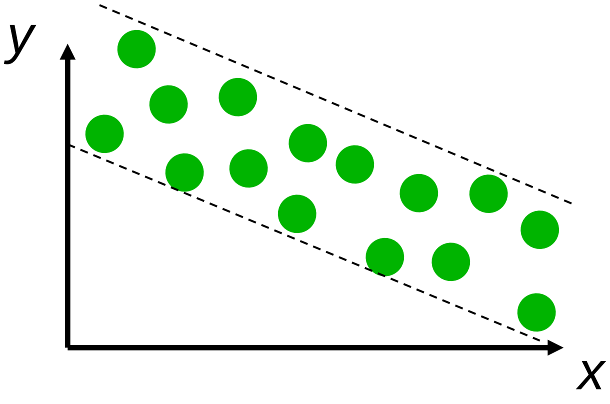
Korelasyon Analizi iki değişken arasındaki ilişkinin yönünü ve büyüklüğünü değerlendirir.

- Serpilme Diyagramı yardımıyla değişkenler arasındaki ilişkinin yapısı görülebilir.

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

$$r_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$



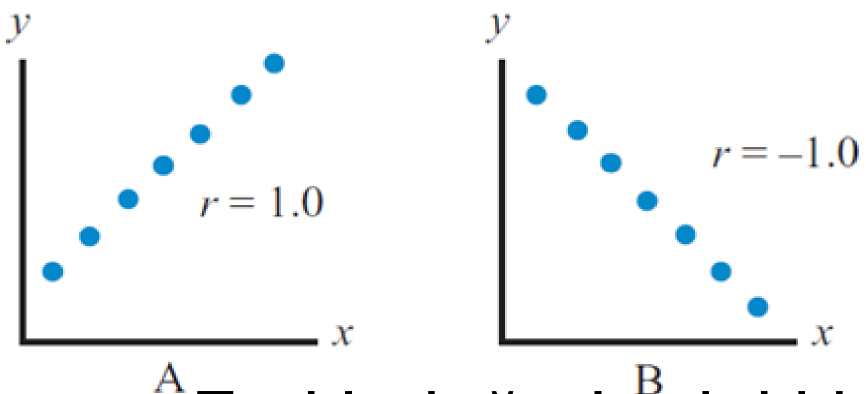


Korelasyon Katsayısı

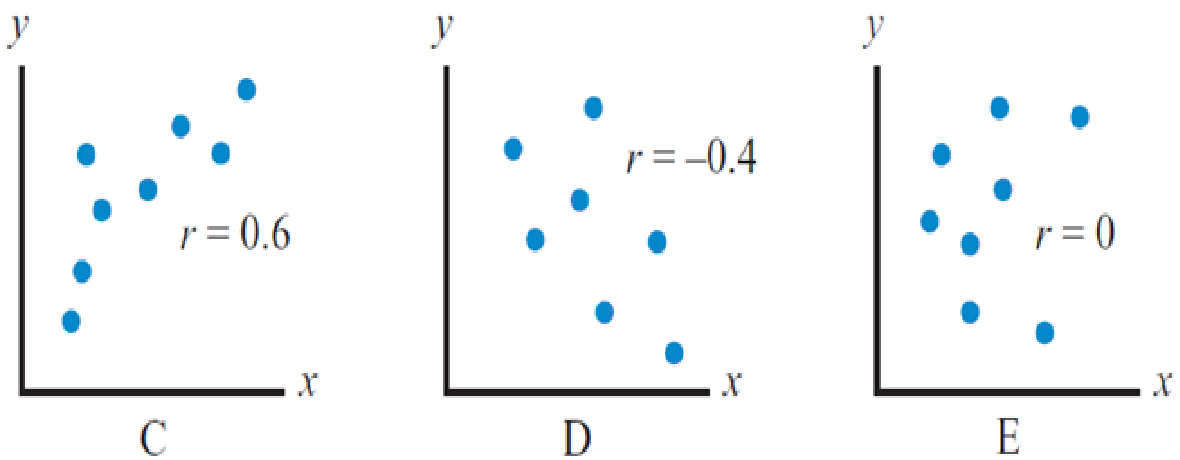
Korelasyon katsayısı,r, iki değişken arasındaki ilişkinin yönünü ve büyüklüğünü gösterir.

- $-1<r<1$ (korelasyon katsayısı -1 ve +1(dahil) arasında değerler alır
- Eğer *r* = 0 ise x ve y arasında doğrusal bir ilişki yoktur. Eğer ‘r’ +1 değerine yakınsa aynı yönlü güçlü bir ilişki, -1 değerine yakınsa ters yönlü zayıf bir ilişkinin varlığından söz edilir.

Korelasyon Katsayısı



Farklı değerlerdeki korelasyon katsayıları



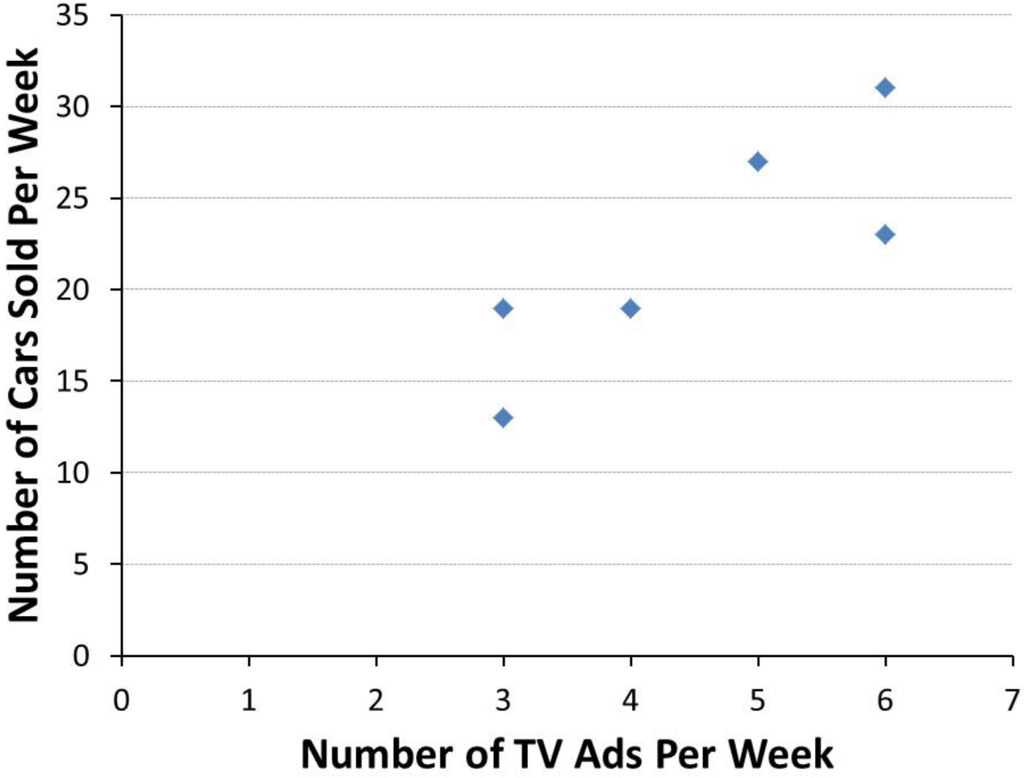
Örnek: Bir satıcı, haftalık olarak televizyonda gösterilen reklam sayısı ile haftalık araba satışı arasındaki ilişkiyi araştırmaktadır. 6 haftaya ait veriler şöyledir:

Hafta	*TV'de reklam sayısı* x	*Satılan araba sayısı* y
1	3	13
2	6	31
3	4	19
4	5	27
5	6	23
6	3	19

Korelasyon Analizi

Hafta	*x*	*y*
1	3	13
2	6	31
3	4	19
4	5	27
5	6	23
6	3	19

Serpilme Diyagramı



Korelasyon Analizi

hafta	x	y	xy	x^2	y^2
1	3	13	39	9	169
2	6	31	186	36	961
3	4	19	76	16	361
4	5	27	135	25	729
5	6	23	138	36	529
6	3	19	57	9	361
	$\Sigma x = 27$	$\Sigma y = 132$	$\Sigma xy = 631$	$\Sigma x^2 = 131$	$\Sigma y^2 = 3110$

Korelasyon Katsayısı

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

Serilerin değeri formüle yerleştirilirse:

$$r = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\left[n\sum y^2 - \left(\sum y\right)^2\right]}} = \frac{(6)(631) - (27)(132)}{\sqrt{\left[(6)(131) - (27)^2\right]\left[(6)(3110) - (132)^2\right]}}$$

$$= \frac{222}{\sqrt{[57]\,[1236]}} = \frac{222}{265.43} = 0.836$$

Araba satışı ve reklam sayısı arasında aynı yönlü ve güçlü bir ilişki vardır r=0.836

Korelasyon katsayısının anlamlılığının testi:

Örnekleme ait korelasyon katsayısı ‘r’ ve anakütleye ait korelasyon katsayısı **‘ρ’ ile gösterilmek üzere,**

Anakütley korelasyon katsayısının,ρ, sıfırdan anlamlı derecede farklı olup olmadığı şu şekilde değerledirilebilir:

- (Tek yönlü hipotez)

H_0: $\rho \leq 0$ (x ve y arasında pozitif bir ilişki yoktur)
H_1: $\rho > 0$ (x ve y arasında anlamlı ve pozitif bir ilişki vardır)

Korelasyon katsayısı için t istatisiği şöyle hesaplanır:

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

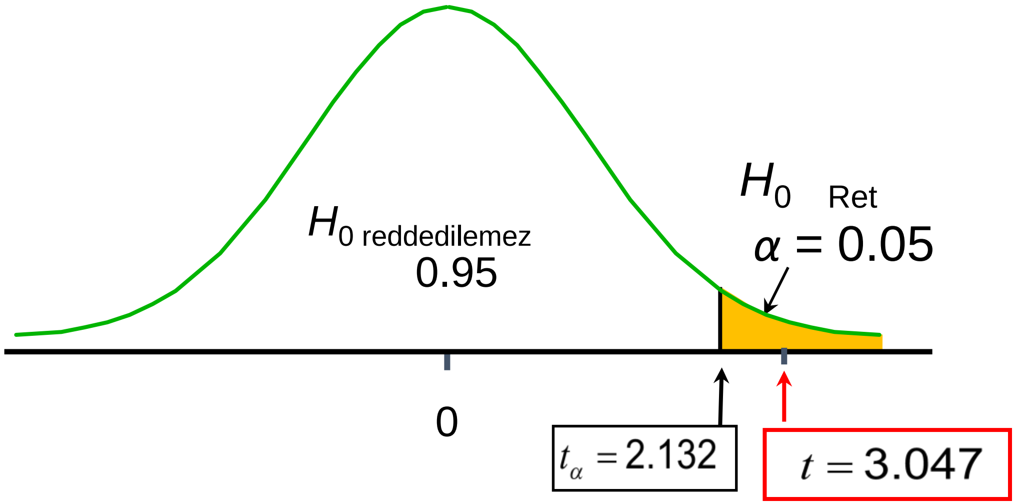
r = örneklem korelasyon katsayısı
n = örneklem büyüklüğü

$r = 0.836$, $n = 6$

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0.836}{\sqrt{\frac{1-(0.836)^2}{6-2}}} = \frac{0.836}{\sqrt{0.0753}} = 3.047$$

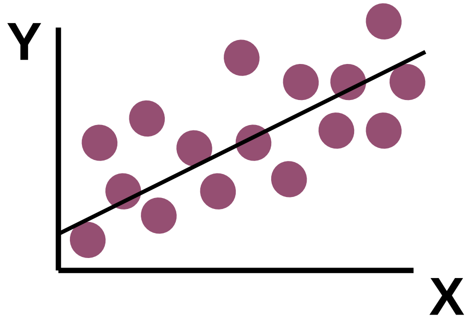
Hesaplanan t istatistiği n-2 serbestlik dereceli t dağılımı ile kıyaslanır:

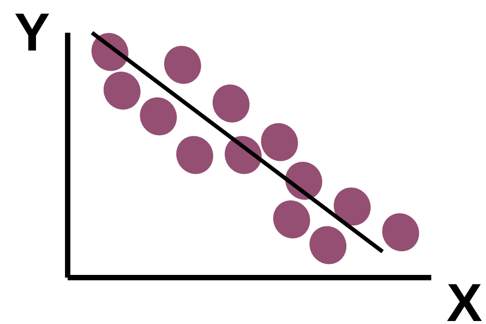
Sıfır hipotezi reddedilir, anakütle korelasyon katsayısı sıfırdan büyüktür.



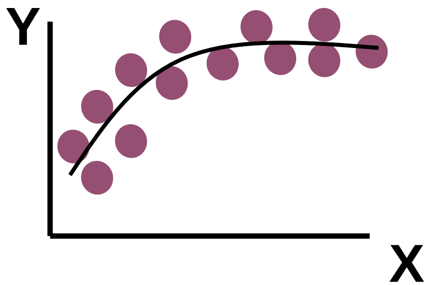
İlişki türlerine göre serpilme diyagramı

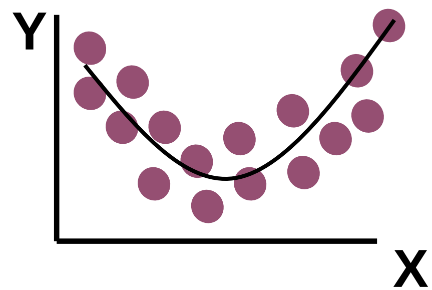
Doğrusal İlişki

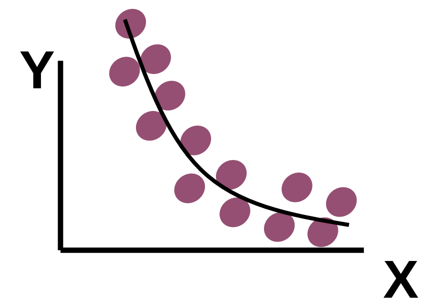




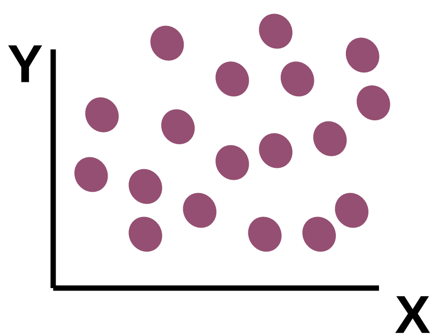
Doğrusal Olmayan İlişki

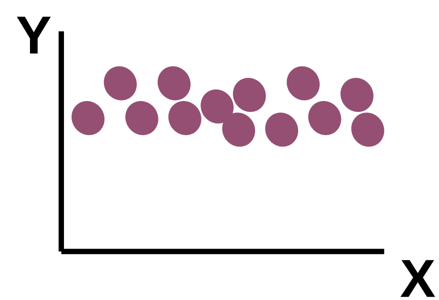






İlişki Yok





Regresyon Analizi Değişkenler Arasındaki İlişkinin Değerlendirilmesi için Kullanılır!

DCOVA

Regreson Analizi ile

- En az bir bağımsız değişkenin aldığı değerden hareketle bağımlı değişkenin alacağı değer tahmin edilir.
- Bağımsız değişkendeki değişimin bağımlı değişkende yaratacağı etki değerlendirilebilir.

Bağımlı Değişken: Tahmin etmek ya da açıklamak istediğimiz değişken

Bağımsız Değişken: Bağımlı değişkeni açıklamak için kullandığımız değişken(ler)

Basit Regresyon Modeli

Modelde bağımsız değişken, X, bir tanedir.

X ve Y arasındaki ilişki doğrusal varsayılır.

Basit Regresyon Modeli

DCOVA

Anakütle sabit terim

Anakütle eğim katsayısı

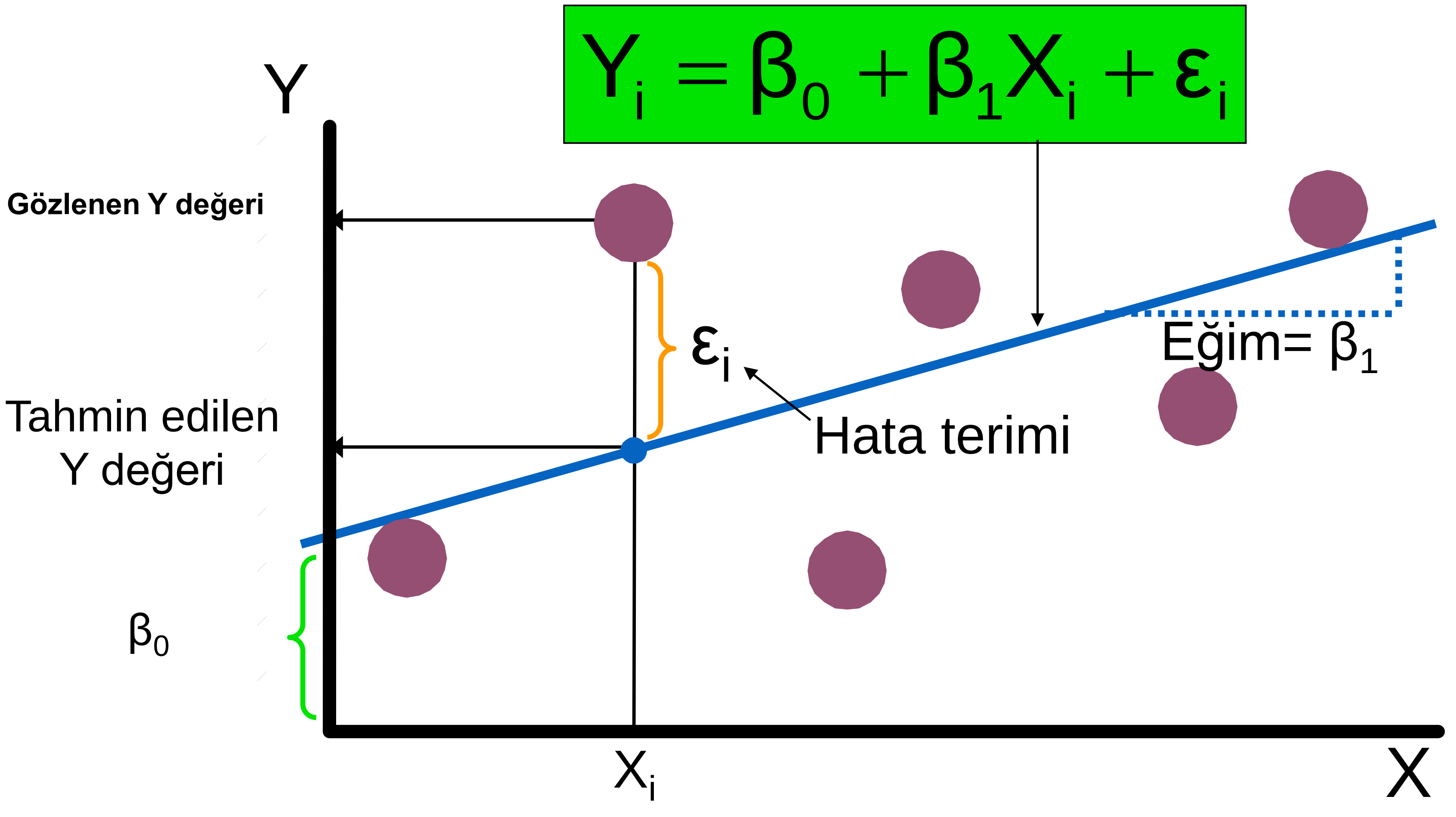
Bağımsız değişken

Hata terimi

Bağımlı Değişken

$$Y_i = \underbrace{\beta_0 + \beta_1 X_i}_{\text{Doğrusal bileşen}} + \underbrace{\varepsilon_i}_{\text{Rastlantısal hata bileşe}}$$

Basit Regresyon Modeli



Tahmin edilen regresyon doğrusu

DCOVA

Bu basit regresyon doğrusu anakütleye ait regresyon doğrusunun tahminidir

Tahmin edilen değer

Tahmin edilen başlangıç terimi

Tahmin edilen eğim katsayısı

X değeri

$$\hat{Y}_i = b_0 + b_1 X_i$$

En Küçük Kareler Yöntemi

DCOVA

b_0 and b_1 değerlerinin bulunabilmesi için Y ve $\hat{Y}$ değerleri arasındaki farkın karelerinin minimize

$$\min \sum (Y_i - \hat{Y}_i)^2 = \min \sum (Y_i - (b_0 + b_1 X_i))^2$$

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}$$

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

$$b = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

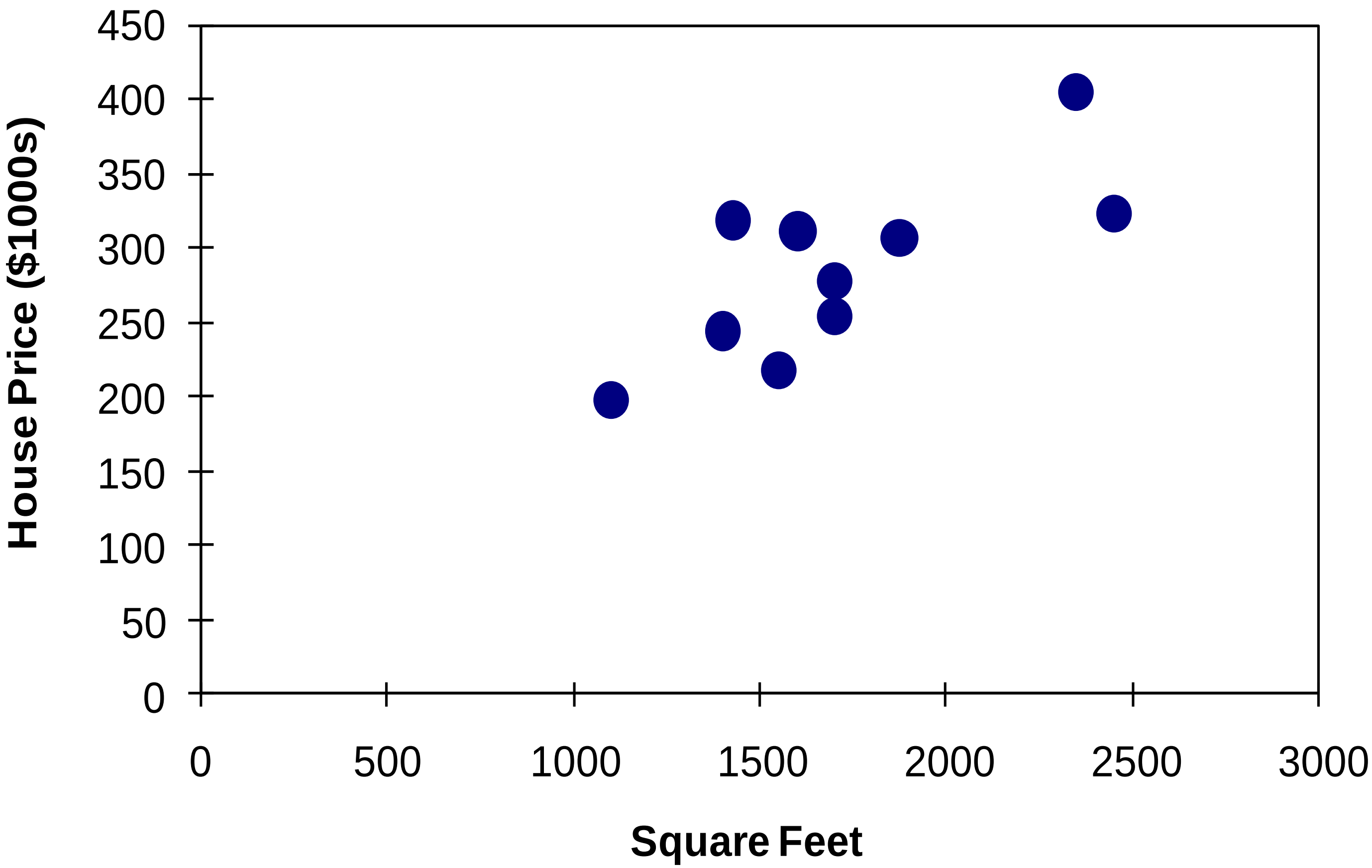
$$a = \bar{Y} - b\bar{x}$$

Örnek:

Bir emlakçı evlerin m^2 ise fiyatları arasındaki ilişkiyi araştırmak istemektedir. Rastgele seçtiği 10 eve ait bilgiler şöyledir:

Ev fiyatı ()	Metrekare ()
245	1,400
312	1,600
279	1,700
308	1,875
199	1,100
219	1,550
405	2,350
324	2,450
319	1,425
255	1,700

Serpilme Diyagramı



Excel çıktısı

DCOVA

Regression Statistics	
Multiple R	0.76211
R Square	0.58082
Adjusted R Square	0.52842
Standard Error	41.33032
Observations	10

Regresyon eşitliği:

$$\widehat{\text{house price}} = 98.24833 + 0.10977\ (\text{square feet})$$

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	18934.9348	18934.9348	11.0848	0.01039
Residual	8	13665.5652	1708.1957		
Total	9	32600.5000			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	98.24833	58.03348	1.69296	0.12892	-35.57720	232.07386
Square Feet	0.10977	0.03297	3.32938	0.01039	0.03374	0.18580

Simple Linear Regression Example: Interpreting b_1

DCOVA

$$\widehat{\text{house price}} = 98.24833 + \boxed{0.10977}\,(\text{square feet})$$

Eğim katsayısı bağımsız değişkenden bir birimlik değişim olduğunda Y'deki değişim miktarı hakkında bilgi verir.

$\boxed{b_1 = 0.10977}$ evin alanının 1 metrekarelik değişiminin fiyatta ortalama nasıl bir değişime neden olacağını gösterir.

Tahmin

DCOVA

2000 metrekarelik bir evin fiyatını tahmin edelim

$$\widehat{\text{house price}} = 98.25 + 0.1098\,(\text{sq.ft.})$$

$$= 98.25 + 0.1098(2{,}000)$$

$$= 317.85$$

Katsayıların bulunması

DCOVA

$$b_1 = \frac{SSXY}{SSX} \qquad b_0 = \bar{Y} - b_1 \bar{X}$$

$$SSXY = \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y}) = \sum_{i=1}^{n} X_i Y_i - \frac{\left(\sum_{i=1}^{n} X_i\right)\left(\sum_{i=1}^{n} Y_i\right)}{n}$$

$$SSX = \sum_{i=1}^{n} (X_i - \bar{X})^2 = \sum_{i=1}^{n} X_i^2 - \frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}$$

$$\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n} \quad \text{and} \quad \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Katsayıların bulunması

DCOVA

House	House Price in $1000s (Y)	Square Feet (X)	X^2	Y^2	XY
1	245	1,400	1,960,000	60,025	343,000
2	312	1,600	2,560,000	97,344	499,200
3	279	1,700	2,890,000	77,841	474,300
4	308	1,875	3,515,625	94,864	577,500
5	199	1,100	1,210,000	39,601	218,900
6	219	1,550	2,402,500	47,961	339,450
7	405	2,350	5,522,500	164,025	951,750
8	324	2,450	6,002,500	104,976	793,800
9	319	1,425	2,030,625	101,761	454,575
10	255	1,700	2,890,000	65,025	433,500
Total	2865	17,150	30,983,750	853,423	5,085,975

$$SSXY = \sum_{i=1}^{n} X_i Y_i - \frac{(\sum_{i=1}^{n} X_i)(\sum_{i=1}^{n} Y_i)}{n} = 5,085,975 - \frac{(17,150)(2,865)}{10} = 172,500$$

$$SSX = \sum_{i=1}^{n} X_i^2 - \frac{(\sum_{i=1}^{n} X_i)^2}{n} = 30,983,750 - \frac{17,150^2}{10} = 1,571,500$$

$$\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n} = \frac{2,865}{10} = 286.5 \quad \text{and} \quad \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{17,150}{10} = 1,715$$

DCOVA

$$SSXY = \sum_{i=1}^{n} X_i Y_i - \frac{(\sum_{i=1}^{n} X_i)(\sum_{i=1}^{n} Y_i)}{n} = 5,085,975 - \frac{(17,150)(2,865)}{10} = 172,500$$

$$SSX = \sum_{i=1}^{n} X_i^2 - \frac{(\sum_{i=1}^{n} X_i)^2}{n} = 30,983,750 - \frac{17,150^2}{10} = 1,571,500$$

$$\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n} = \frac{2,865}{10} = 286.5 \quad \text{and} \quad \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{17,150}{10} = 1,715$$

$$b_1 = \frac{SSXY}{SSX} = \frac{172,500}{1,571,500} = 0.1098$$

$$b_0 = \bar{Y} - b_1 \bar{X} = 286.5 - 0.1098 * 1,715 = 98.193$$

Regresyon analizi değişkenler arasındaki ilişkinin araştırılmasında kullanılan istatistiksel bir araçtır. Bu araçla, bir değişkenin diğer değişken üzerindeki nedensel ilişkisi araştırılır. İncelenen ilişkideki değişkenler aralarındaki ilişki göz önüne alınarak değişkenler bağımlı ve bağımsız olarak isimlendirilir.

İstatistiğin öncelikli ilgi alanını rastlantı değişkeninin davranışını bir modelle tahmin etmek oluşturur. Davranışı tahmin edilecek olan rastlantı değişkeni bir diğer değişken(ler)in fonksiyonu olarak gösterilebilir ve bu değişken bağımlı olarak isimlendirilir ve Y ile gösterilir. Bağımlı değişkeni etkileyen değişken ise X ile gösterilir ve bağımsız değişken olarak isimlendirilir.

Bağımlı değişken, bağımsız değişken(ler) tarafından açıklanmaya çalışılır ve açıklayıcı değişkenlerin modelde bilinen sabitler olduğu varsayılır.

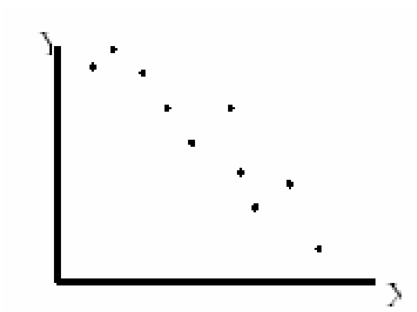
Model şu şekilde gösterilir;

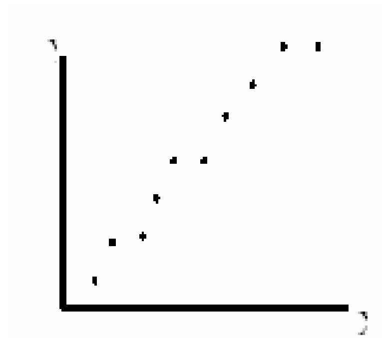
$$Y_i = b_0 + b_1 X_i + \varepsilon_i$$

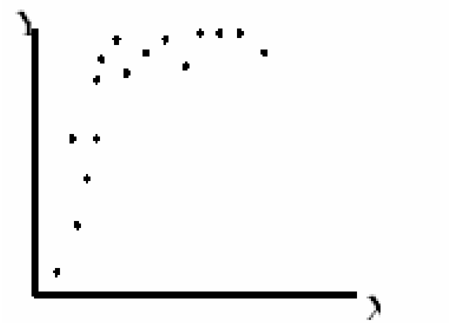
burada b_0 sabit katsayıyı gösterir buna başlangıç parametresi de denir, b_1 ise eğim parametresidir. X'deki 1 birimlik değişmenin Y üzerinde nasıl bir değişim yaptığını gösterir. Denklemdeki ise daha sonra da açıklanacağı gibi hata terimine karşılık gelir.

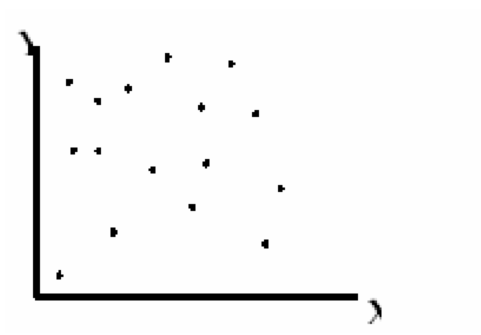
Örneğin, tüketim ve gelir üzerine yapılan bir çalışmada bağımsız değişken gelir, bağımlı değişkense tüketimdir ya da bir hastaya uygulanan ilacın dozu ve hastanın iyileşme süreci çalışmasında bağımsız değişken ilacın dozu ve bağımlı değişkense hastanın iyileşme süreci olur.

Serpilme diyagramı, i. gözlemin bağımlı değeri y_i ve bağımsız değeri x_i olmak üzere tüm gözlem çiftleri üzerinden, her ikili yani Y ve X değişkenlerinin aldığı tüm değerler birer nokta ile temsil edilecek şekilde çizilir. Diyagramdaki dağılıma bakılarak uygun model belirlenir.









Yukarıdaki serpilme diyagramlarında noktaların ortasından geçecek olan eğri dikkate alınır ve bu eğri incelenen ilişki biçimi hakkında bilgi verir. Buna göre, ilk çizimde noktaların bir doğru etrafında toplandığı söylenebilir ve değişkenler arasında aynı yönlü doğrusal bir ilişkinin varlığından söz edilebilir.

İkinci çizimse ters yönlü doğrusal bir ilişkinin varlığını gösterir. Üçüncü çizimde doğrusal olmayan bir ilişkinin varlığı söz konusudur. Son çizim dikkate alındığında ise bir eğri oluşturmak mümkün görünmemektedir, değişkenler arasında bir ilişkinin olmadığı sonucuna varılır.

Serpilme diyagramı çizimi sonrasında uygun modele karar verilir ve modeldeki parametreler En küçük Kareler yöntemi ile tahmin edilir.

En Küçük Kareler Yöntemi ile bulunacak eğrinin her (xi,yi) gözlem çiftine karşılık gelen nokta ile bu noktanın EKK ile elde edilecek eğri üzerindeki dik izdüşümü arasındaki farklar toplamı sıfır olmalıdır. Bu farklar, yani Y_i değerlerinin regresyon doğrusuna olan uzaklığı, daha sonrada bahsedileceği gibi 'hata' olarak isimlendirilir. Y_i değerlerinin regresyon doğrusu üzerindeki görüntüsü $\hat{Y}_i$ (tahmini Y_i) ile arasındaki fark hataya karşılık gelir. İdeal regresyon doğrusu, bu farkların karelerinin toplamını $\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$, minimum verenle elde edilir.

Basit Regresyon denkleminin EKK ile tahmin sürecine notlarda detaylı bir şekilde değinilmiştir. Sabit ve Eğim parametrelerinin tahmininde kullanılacak olan eşitlikler şöyledir:

$$b_1 = \frac{n\sum XY - \sum X \sum Y}{n\sum X^2 - (\sum X)^2} \qquad b_0 = \frac{\sum Y \sum X^2 - \sum X \sum XY}{n\sum X^2 - (\sum X)^2}$$

Eğim ve sabit parametre tahminlerini ortalamadan sapmalar üzerinden giderekte belirlemek mümkündür:

$$x = X - \bar{X} \qquad y = Y - \bar{Y}$$

$$\bar{Y} = b_0 + b_1\bar{X}$$

$$b_0 = \bar{Y} - b_1\bar{X}$$

$$b_1 = \frac{\sum xy}{\sum x^2} \qquad \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

ÖRNEK: Aşağıda bir sınıftaki öğrencilerin muhasebe ve matematik derslerine ait veri bulunmaktadır. Muhasebe dersinden başarının matematik dersinden başarıya bağımlı olup olmadığını sınamak için regresyon denklemini oluşturunuz.

Muhasebe	Matematik
1	2
2	3
3	5
5	6
6	7
7	10
8	7
8	8

Yort=5

Xort=6

Y	X	Y^2	X^2	YX	X-Xort=x	Y-Yort=y	xy	x^2
1	2	1	4	2	-4	-4	16	16
2	3	4	9	6	-3	-3	9	9
3	5	9	25	15	-1	-2	2	1
5	6	25	36	30	0	0	0	0
6	7	36	49	42	1	1	1	1
7	10	49	100	70	4	2	8	16
8	7	64	49	56	1	3	3	1
8	8	64	64	64	2	3	6	4
40	**48**	**252**	**336**	**285**	**0**	**0**	**45**	**48**

Şu sonuçlara ulaşılır:

$$b_0 = \frac{\sum Y \sum X^2 - \sum X \sum XY}{n\sum X^2 - (\sum X)^2}$$

$$b_0 = \frac{40(336) - 48(285)}{8(336) - 48^2} = -0.625$$

$$b_1 = \frac{n\sum XY - \sum X \sum Y}{n\sum X^2 - (\sum X)^2}$$

$$b_1 = \frac{8(285) - 48(40)}{8(336) - 48^2} = 0.9375$$

$$b_0 = \bar{Y} - b_1\bar{X} = 5 - 0.9375(6) = -0.625$$

$$b_1 = \frac{\sum xy}{\sum x^2} = \frac{45}{48} = 0.9375$$

$$\hat{y} = -0.625 + 0.9375\text{x}$$

ÖRNEK:

Aşağıda bir eyaletteki suç ve işsizlik oranlarına ilişkin veri mevcuttur.

işşizlik oranı	**suç oranı**
0,8	3
1,4	6
2,3	7
3,5	15
4,5	19

İşsizlik ve suç işleme oranları arasındaki ilişkiyi gösteren regresyon denklemini oluşturunuz.

Sorunun çözümünün ilk aşamasında bağımlı ve bağımsız değişkenleri belirleyelim. İşssizlik oranı bağımsız değişkendir (x), suç oranını etkiler ki bu da bağımlı değişken (Y) olarak isimlendirilir. İlgili kolonların toplam ve çarpımlarının toplamına ait bilgiler aşağıdaki gibi özetlenmiştir.

$$\sum Y = 50 \quad \sum X = 12.5 \quad \sum XY = 164.9$$

$$\sum X^2 = 40.39 \quad \sum Y^2 = 680$$

$$\sum (X - \bar{X})(Y - \bar{Y}) = 39.9$$

$$\sum (X - \bar{X})^2 = 9.14 \quad \sum (Y - \bar{Y})^2 = 180$$

$$b1 \equiv \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} \qquad b_0 = \bar{Y} - b_1 \bar{X}$$

$$b_1 = \frac{39.9}{9.14} = 4.365$$

$$b_0 = \frac{50}{5} - 4.365(\frac{12.5}{5}) = -0.9$$

$$Y = -0.9 + 4.365X$$

İşsizlik oranı 1 birim arttığında suç oranı 4.365 birim artar.

Ortalamadan sapmalar serisi yerine orijinal seriden hareket edilirse eğim katsayısı şöyle bulunacaktır:

$$b_1 = \frac{n\sum XY - \sum X \sum Y}{n\sum X^2 - (\sum X)^2} = \frac{5(164.9) - (12.5)(50)}{5(40.39) - 12.5^2} = 4.365$$

ÖRNEK:

Bir firmanın reklam harcamaları ve satış rakamlarına ilişkin veri mevcuttur. Regresyon denklemini oluşturunuz

Reklam Harca.	**Satışlar**
1.6	6
2.8	12
4.6	14
7	30
9	38

Bağımlı değişken satışlardır., Reklam harcamaları ise bağımsız değişkendir.

$$\sum Y = 100 \quad \sum X = 25 \quad \sum XY = 659.6$$
$$\sum X^2 = 161.56 \quad \sum Y^2 = 2720$$
$$\sum (X - \bar{X})(Y - \bar{Y}) = 159.6$$
$$\sum (X - \bar{X})^2 = 36.56 \quad \sum (Y - \bar{Y})^2 = 720$$

$$b_1 = \frac{159.9}{36.56} = 4.37$$
$$b_0 = \frac{100}{5} - 4.37(\frac{25}{5}) = -1.8$$
$$Y = -1.8 + 4.37X$$

Reklam harcamaları 1 birim artarsa satış 4.37 birim artar.

$$b_1 = \frac{n\sum XY - \sum X \sum Y}{n\sum X^2 - (\sum X)^2} = \frac{5*659.6 - 25*100}{5*161.56 - 25^2} = 4.37$$

$$b_0 = \frac{\sum Y \sum X^2 - \sum X \sum XY}{n\sum X^2 - (\sum X)^2} = -1.8$$

Örnek: Aşağıdaki x,y serilerinden hareketle regresyon denklemini bularak yorumlayınız.

x	y
20	12
19	10
17	9
16	8
13	6

$$n = 5, \quad \sum X = 85, \quad \sum Y = 45$$

$$\sum XY = 789, \quad \sum X^2 = 1475, \quad \sum Y^2 = 425$$

$$b_1 = \frac{789 - 5(85/5)(45/5)}{1475 - 5(85/5)^2} = 0.8$$

$$b_0 = 45/5 - (0.8)(85/5) = -4.6$$

Örnek:Aşağıda verilen x,y ikilisinden hareketle regresyon denklemini bulunuz.

x	y
10.2	7
8.4	5
6.2	4
4.2	1
11	8

$$n = 5, \quad \sum X = 40, \quad \sum Y = 25$$

$$\sum XY = 230.4, \quad \sum X^2 = 351.68, \quad \sum Y^2 = 155$$

$$b_1 = \frac{230.4 - 5(40/5)(25/5)}{351.68 - 5(40/5)^2} = 0.96$$

$$b_0 = 25/5 - (0.96)(40/5) = -2.677$$

$$y = -2.677 + 0.96x$$

Örnek: Aşağıdaki veriden hareketle regresyon denklemini bulunuz

x	y
20	6
19	8
17	9
16	10
13	12

$$n = 5, \quad \sum X = 85, \quad \sum Y = 45$$

$$\sum XY = 741, \quad \sum X^2 = 1475, \quad \sum Y^2 = 425$$

$$b_1 = \frac{741 - 5(85/5)(45/5)}{1475 - 5(85/5)^2} = -0.8$$

$$b_0 = 45/5 - (-0.8)(85/5) = 22.6$$