

TEKLİ REGRESYON ANALİZİ (EN KÜÇÜK KARELER YÖNTEMİ) Bölüm 2

Prof. Dr. Fazıl GÖKGÖZ

1

TEK DEĞİŞKENLİ REGRESYON MODELİ

- İki ya da daha fazla değişken arasında bir ilişkinin bulunup bulunmadığını test eden ve bunu doğrusal veya eğrisel olarak ifade eden denklemlere **regresyon modeli** denir.
- Ancak, burada doğrusal olarak ifade edilen regresyon modelleri üzerinden işlem yapacağız.
- İçerisinde yalnızca **bir tane bağımsız değişken** ile buna bağlı tahmin edilen tek **bağımlı değişken** bulunduran denklemlere tek değişkenli veya **tekli regresyon modeli** denilir.

Prof. Dr. Fazıl GÖKGÖZ

2

Doğrusal Regresyon Denklemi Formu

X_i :bağımsız (*açıklayıcı*) değişken

Y_i :bağımlı (*açıklanan*) değişken

e_i :hata payı

a,b: regresyon katsayıları

n:gözlem sayısı

olmak üzere;

$$Y_i = a + bX_i + e_i$$

$(i=1,2,\dots,n)$

Prof. Dr. Fazıl GÖKGÖZ

3

Regresyon Denklemine ait Varsayımlar

- İstatistiksel rasgele örneklem ve popülasyonu en iyi temsil eden örneklem kullanılmaktadır.
- Her x değerine karşılık mutlaka bir y değeri vardır.
- Bağımlı değişken içerisinde ortalaması 0 ve varyansı σ^2 (*otokorelasyonsuz*) olan normal dağılım sergilemektedir.
- Bağımsız değişkenler hatasız olarak veriyi yansıtmaktadır.
- Denkleme hata teriminin eklenmiş olması denklemde dışlanan değişkenlerin (*denkleme bulunmayan bağımsız değişkenler*) var olabileceği olasılığını göstermektedir.

Prof. Dr. Fazıl GÖKGÖZ

4

Tahmin

- Kurulan regresyon modeli ilgilenilen problemle ilgili örnek olarak alınmış gözlem değerleri kullanılarak hesaplanmaya çalışılır. Bu nedenle, kurduğumuz modeldeki değerler tahmini değerler olacaktır.
- Tahmin sonrası açıklanmaya çalışılan değerler ile bunu açıklayan değerler şapka (^) ile yazılırsa **regresyon tahmin modeli** elde edilmiş olur.
- Tahmin modeli kurulmasının ardından tahmin edilen katsayıların güvenilirliği (*ana kütleyi ne kadar temsil ettiği*) test edilmelidir. Çünkü, bu katsayılar bağımsız değişkenin bağımlı değişkeni ne oranda açıkladığını gösterir.

Prof. Dr. Fazıl GÖKGÖZ

5

Regresyon Tahmin Modeli

Kurulan regresyon modeli $Y=a+bX+e$ ise bu modele ait tahmin modeli:

$$\hat{Y}_i = \hat{a} + \hat{b}X_i + e_i$$

Prof. Dr. Fazıl GÖKGÖZ

6

En Küçük Kareler Yaklaşımı

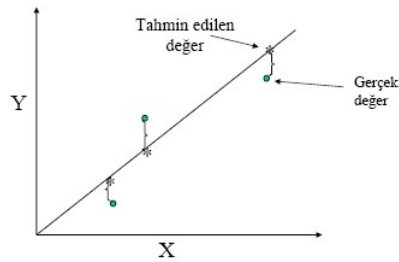
- Tahmin edilen katsayıların ve değişkenlerin hesaplanması için geliştirilen ve gerçek katsayılara en yakın sonuçları veren yöntemlerden bir tanesi **En Küçük Kareler (EKK)** yöntemidir.
- Bir diğer ifade ile regresyon modelindeki hata terimini minimum yapan yöntem EKK yöntemidir.

Prof. Dr. Fazıl GÖKGÖZ

7

.....
EKK yöntemi sayesinde gerçek Y değerleri ile tahmin edilen Y değerleri arasındaki farkın kareleri toplamını minimum yapılır.

Diğer ifadeyle, hataların karelerinin toplamının $[\sum(Y - \hat{Y})^2]$ minimum olması koşulunu sağlayabilecek değişkenlerin katsayıları tahmin edilir.



Prof. Dr. Fazıl GÖKGÖZ

8

Regresyon Katsayılarının Kestirimi

- Regresyon denkleminin belirlenebilmesi için alfa (\hat{a}) ve beta (\hat{b}) parametrelerinin kestiriminin yapılması gerekir. \hat{a} denklem için parametre olarak da bilinmekte olup sabit değer diye nitelendirilmektedir. Bu değer X bağımsız değişkeni sıfır değerini aldığı zaman bağımlı değişken olan Y'nin alacağı değeri göstermektedir. Ayrıca regresyon doğrusu çizilirken doğrunun Y eksenindeki başlangıç noktasını ifade eder.

$$\hat{a} = \bar{Y} - \hat{b}\bar{X} \quad \left\{ \begin{array}{l} \bar{X} = \frac{\sum X}{n} \\ \bar{Y} = \frac{\sum Y}{n} \end{array} \right.$$

Prof. Dr. Fazıl GÖKGÖZ

9

Beta Katsayısı (\hat{b})

Beta katsayısı (\hat{b}) Diğer tüm faktörlerin sabit olması halinde bağımsız değişkendeki değişimin bağımlı değişkeni ne kadar değiştireceğini gösterir. Ayrıca regresyon doğrusunun eğimi olan parametre olarak da bilinmektedir.

1.YOL

$$\hat{b} = \frac{\sum_{i=1}^n (Y_i - \bar{Y}) \cdot (X_i - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

2.YOL

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

Prof. Dr. Fazıl GÖKGÖZ

10

.....

- Katsayıların hesaplanmasından sonra, \hat{Y} olarak gösterilen regresyon tahmini değerleri, regresyon denkleminde tahmin katsayıları yerine koyularak ve hata terimi göz ardı edilerek bulunan değerlerdir.
- Bulunan \hat{Y} tahmin değerleri Y değerlerinden çıkarılırsa o veriye ait hata terimi bulunmuş olur.
- Tahmin modeline ait katsayılar bulunup regresyon doğru denklemini yazıldıktan sonra x-y koordinat eksenlerinde doğru denklemini çizilebilir.

Prof. Dr. Fazıl GÖKGÖZ

11

Örnek:

- Korelasyon konusu içerisindeki örneğimiz için ilgili regresyon modelini bulup doğruyu çizelim: veriler;

Satış elemanı	Telofonla yapılan arama sayısı	Satış yapılan ürün adeti
Ali	14	28
Veli	35	66
Ayşe	22	38
Gül	29	70
Hüsnü	6	22
Necati	15	27
Zehra	17	28
Fatma	20	47
Zeynep	12	14
Ahmet	29	68

Prof. Dr. Fazıl GÖKGÖZ

12

Katsayılar için gerekli değerler;

- Bir tablo ile her bir gözlem için

Satış elemanı	Telofonla yapılan arama sayısı(X)	Satış yapılan ürün adeti (Y)	X ²	XY	Y ²
Ali	14	28	196	392	784
Veli	35	66	1225	2310	4356
Ayşe	22	38	484	836	1444
Gül	29	70	841	2030	4900
Hüsnü	6	22	36	132	484
Necati	15	27	225	405	729
Zehra	17	28	289	476	784
Fatma	20	47	400	940	2209
Zeynep	12	14	144	168	196
Ahmet	29	68	841	1972	4624
TOPLAM	199	408	4681	9661	20510

Prof. Dr. Fazıl GÖKGÖZ

13

Katsayıları hesaplırsak;

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

$$= \frac{10(9661) - (199)(408)}{10(4681) - (199)^2}$$

$$= 2.1387$$

$$\hat{a} = \bar{Y} - \hat{b}X$$

$$= 40,8 - 2,1387(19,9)$$

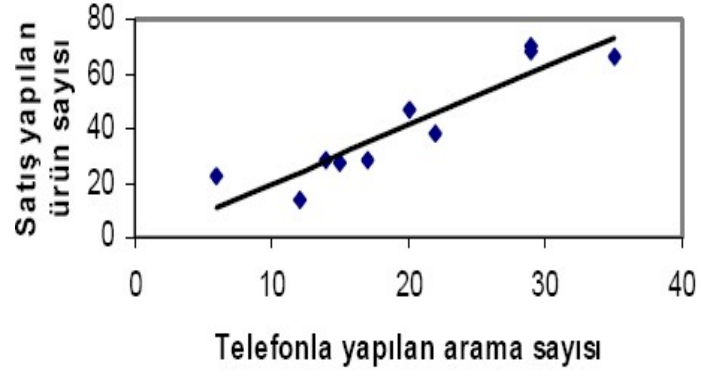
$$= -1,706$$

Prof. Dr. Fazıl GÖKGÖZ

14



$$\hat{Y} = -1,7601 + 2,1387 X$$



Prof. Dr. Fazıl GÖKGÖZ

15