

eğim $\Rightarrow b = \frac{\sum y_i x_i}{\sum x_i^2}$ = ortalamadan sapmalar cinsinden

Paralel şekilde (eğim bulunmadan) regresyon doğrusu çizilebilir.

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$= \sum_{i=1}^n (y_i - b x_i)^2$$

$$\frac{d \sum_{i=1}^n e_i^2}{d b} \rightarrow = 2 \sum_{i=1}^n (y_i - b x_i) (-x_i) = 0$$

$$= -2 \sum_{i=1}^n (y_i - b x_i) x_i = 0$$

$$= \sum_{i=1}^n (y_i - b x_i) x_i = 0$$

$$= \sum_{i=1}^n y_i x_i - b \sum_{i=1}^n x_i^2 = 0$$

$$\boxed{\sum_{i=1}^n y_i x_i = b \sum_{i=1}^n x_i^2}$$

ortalamadan sapmalar cinsinden oluşan en küçük kareler yöntemi

$$\Rightarrow \sum_{i=1}^n Y_i = n \cdot a + b \sum_{i=1}^n X_i \quad 1. \text{ denklem}$$

$$\sum_{i=1}^n Y_i X_i = a \sum_{i=1}^n X_i + b \sum_{i=1}^n X_i^2 \quad 2. \text{ denklem}$$

iki denklemle b denklemi bulunur.
Büyük X-Y

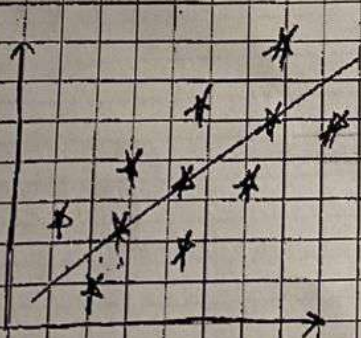
$$\Rightarrow \sum_{i=1}^n y_i x_i = b \sum_{i=1}^n x_i^2 \rightarrow \text{Ortalamadan sapmalar cinsinden OLS}$$

$$b = \frac{\sum_{i=1}^n y_i x_i}{\sum_{i=1}^n x_i^2} = \frac{\text{cov}(Y, X)}{\text{var}(X)} \rightarrow \text{iki denklem kullanmadan b denklemi!}$$

küçük x, y
eğim parametresi

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

Determinasyon katsayısı = Uyumu ölçülüşü = R^2



$$Y_i - \bar{Y} = (\hat{Y}_i - \bar{Y}) + e$$

$$\sum_{i=1}^n y_i^2 = \sum_{i=1}^n \hat{y}_i^2 + \sum_{i=1}^n e_i^2$$

Toplam kareler toplamı Açıklanan kareler toplamı Kalıntı kareler toplamı

23 KASIM 2018

BASIT DOĞRULAR REGRESYON MODELİ İÇİN OLAGAN EN KÜÇÜK KARELER TAHMİNİ

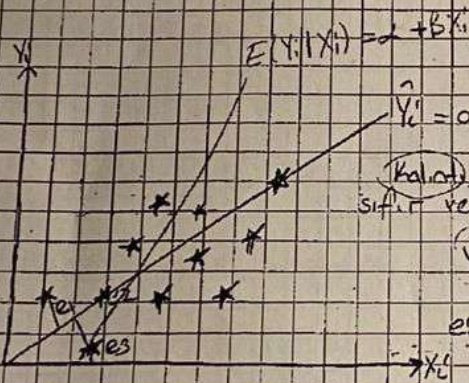
→ $Y_i = \alpha + \beta X_i + u_i$ → Anlatı- regresyon ilişkisi

→ $E(Y_i | X_i) = \alpha + \beta X_i$ → Anlatı- regresyon ilişkisinin beklenen değeri

→ $Y_i = a + bX_i + e_i$ → Örneklem- regresyon ilişkisi
(hata teriminin tahmini) - (orjinal değerle tahmini değerin alınması)

→ $\hat{Y}_i = a + bX_i$ → Tahmin edilen regresyon doğrusu

→ $e_i = Y_i - \hat{Y}_i$ → Kalıntı değerinin elde edilmesi
gerçek değer tahmini değer



(Kalıntı) doğruya olan mesafesi. Kalıntı toplamı sıfır ve ortalaması yine sıfır olur.

$Y_i = \hat{Y}_i + e_i$

$e_i = Y_i - a - bX_i$

→ Kalıntı Kareler Toplamı = Kalıntı kareler toplamını en aza yapan a ve b değerleri "olagan en küçük kareler tahmincisi" olur

$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - a - bX_i)^2$

→ Kısmi Türev : A ve b bilinmediği için iki bilinmeyen var ve 2 denklemler çıkarılır

$\frac{\partial}{\partial a} \sum_{i=1}^n e_i^2 = 2 \cdot \sum_{i=1}^n (Y_i - a - bX_i) \cdot (-1) \rightarrow$ a'ya göre türev

$= -2 \cdot \sum_{i=1}^n (Y_i - a - bX_i) = 0$

$= \sum_{i=1}^n (Y_i - a - bX_i) = 0$

$= \sum_{i=1}^n Y_i - an - b \sum_{i=1}^n X_i = 0$

$\sum_{i=1}^n Y_i = na + b \sum_{i=1}^n X_i \rightarrow$ OLS I. Normal Denklemleri

olagan en küçük kareler tahmincisi

ELASTİKİYET / ESNEKLİK

$$Y_i = \alpha + \beta X_i + u_i$$

Birinden bağımsız analize edilmesi

cto $\eta = \frac{\text{Bağımlı değişkendeki oransal değişme}}{\text{Açıklayıcı değişkendeki oransal değişme}}$

$$\eta = \frac{\frac{\Delta Y}{Y}}{\frac{\Delta X}{X}} = \frac{\Delta Y}{\Delta X} \cdot \frac{X}{Y}$$

$\eta = \frac{\text{Bağımlı değişkendeki yüzde değişme}}{\text{Açıklayıcı değişkendeki yüzde değişme}}$

$$\eta = \frac{\frac{\Delta Y}{Y} \cdot 100}{\frac{\Delta X}{X} \cdot 100} = \frac{\Delta Y}{\Delta X} \cdot \frac{X}{Y}$$

* $E(Y|X_i) = \alpha + \beta X_i \Rightarrow$ Fonksiyon üzerinden

$$\frac{\Delta E(Y|X_i)}{\Delta X_i} = \frac{dE(Y|X_i)}{dX_i} = \beta$$

$$\eta = \frac{\Delta E(Y|X_i) / E(Y|X_i)}{\Delta X_i / X_i} = \frac{\frac{dE(Y|X_i)}{dX_i}}{\frac{X_i}{E(Y|X_i)}} \Rightarrow \eta = \beta \cdot \frac{X_i}{E(Y|X_i)}$$

$\rightarrow$ Tahmin üzerinden

$$Y_i = \alpha + \beta X_i$$

$$\frac{\Delta Y_i}{\Delta X_i} = \frac{dY_i}{dX_i} = \beta \Rightarrow \hat{\eta} = \frac{\Delta \hat{Y}_i / \hat{Y}_i}{\Delta X_i / X_i} = \frac{\Delta \hat{Y}_i}{\Delta X_i} \cdot \frac{X_i}{\hat{Y}_i} \rightarrow \hat{\eta} = b \cdot \frac{X_i}{\hat{Y}_i}$$

$\rightarrow$ Tahmin olduğu için $\hat{Y}$ gelir.

$\rightarrow$ Genelde kullanılan formül:

$$\eta = b \cdot \frac{X}{\bar{Y}} \Rightarrow X \text{ ortalaması}$$

$\rightarrow$ Tahmin edilen regresyonda X gördüğümüzde yerine $\bar{X}$ yazıldığında Y değeri

$\rightarrow$ Ortalamalar noktasındaki elastikiyet

! Elastikiyet Yorumu: Açıklayıcı değişkendeki %1 değişim karşısında bağımlı değişkendeki %... değişme. (Açıklayıcı değişkendeki %1'lik değişim karşısında bağımlı değişken %0.85 artmıştır (pozitifse), azalmıştır (negatifse)).

$\eta > 1 \rightarrow$ Elastik

$\eta < 1 \rightarrow$ İnelastik

$\eta = 1 \rightarrow$ Birim Elastik

OLS TAHMİNCİSİNİN VARYANSLARI

$$\text{var}(a|X) = \sigma_u^2 \cdot \frac{\sum_{i=1}^n X_i^2}{n \sum_{i=1}^n X_i^2}$$

$$\text{var}(b|X) = \sigma_u^2 \cdot \frac{1}{\sum_{i=1}^n X_i^2}$$

$\rightarrow \sigma_u^2 = \text{Hicbir zaman bilinmez}$

$$= \sigma_u^2 \cdot \frac{\sum_{i=1}^n X_i^2}{n \cdot \text{var}(X_i)} = \sigma_u^2$$

$$= \sigma_u^2 \cdot \frac{1}{n \cdot \text{var}(X_i)} = \sigma_u^2$$

SÖZ KESİM

MODEL: TEKLARLI

$$TC = f(q)$$

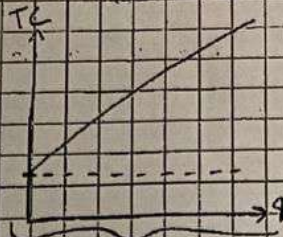
TC → Toplam Maliyet

q → Üretim Miktarı

q ↑ → TC ↑

q = 0 → TC > 0

Sözel-Matematiksel Model



Geometrik Model

$$TC = f(q)$$

$$TC = \alpha + \beta q$$

$$\alpha > 0 \quad \beta > 0$$

$$\frac{dTC}{dq} = \beta > 0$$

İktisadi Model

$$TC = \underbrace{FC}_{\alpha} + \underbrace{VC}_{\beta q}$$

$$TC = \alpha + \beta q_i + u_i$$

$$E(TC_i | q_i) = \alpha + \beta q_i$$

Matematiksel Model

Ekonomik Model

ZARFIYIK

ORJİNDEN GEÇEN REGRESYON DOĞRUSU Bir değişimin orjinden geçmesi demek, fonksiyonun bir kesmesinin olması demektir. Yani açıklayıcı değişken sıfır olduğunda bağımlı değişkenin değeri de sıfır olması. Örnek olarak; girdi faktörlerinden bir tanesi bile olmasa üretim olmaz. Yani kesme terimi sıfırdır.

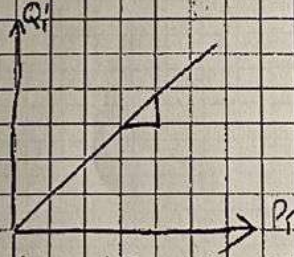
$$Q_i = f(P_i)$$

$$Q_i = \alpha + \beta P_i + u_i$$

$$P_i = 0 \rightarrow Q_i = 0$$

$$\frac{\Delta Q_i}{\Delta P_i} = \frac{dQ_i}{dP_i} = \beta > 0$$

Kesme terimi
gerek
kalmaz
→ $Q_i = \beta P_i + u_i$

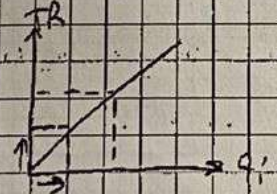


$$TR_i = f(Q_i)$$

$$TR_i = P \cdot q$$

P = sabit sayı

$$TR = \beta q_i + u_i$$



→ Üretim arttıkça toplam
hasılat artar

❗ Kesme terimi olmadığı için orjinden geçen regresyon olur

$$Y_i = \beta X_i + u_i$$

$$Y_i = bX_i + e_i$$

$$e_i = Y_i - \hat{Y}_i$$

$$e_i = Y_i - bX_i$$

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - bX_i)^2$$

$$OLS = \frac{d \sum_{i=1}^n e_i^2}{db} = 0$$

$$= 2 \cdot \sum_{i=1}^n (Y_i - bX_i) \cdot (-X_i)$$

$$= -2 \cdot \sum_{i=1}^n (Y_i - bX_i) \cdot (X_i)$$

$$\sum_{i=1}^n Y_i X_i - b \sum_{i=1}^n X_i^2 = 0$$

$$\sum_{i=1}^n Y_i X_i = b \sum_{i=1}^n X_i^2$$

$$b = \frac{\sum_{i=1}^n Y_i X_i}{\sum_{i=1}^n X_i^2} = \text{Kesme terimi yoksa sadece } b \text{ bulunur.}$$

$$\sum_{i=1}^n y_i^2 = \sum_{i=1}^n \hat{y}_i^2 + \sum_{i=1}^n e_i^2$$

$$1 = R^2 + \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n y_i^2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n y_i^2}$$

= Determinasyon Katsayısı
 Açıklayıcı değişkendeki değişimlerin bağımlı değişken değişimlerinin % kaçını açıklıyor
 sonuna cevap verir

$$0 \leq R^2 \leq 1$$

Uygun
 Yüklü
 düşük

Uygun
 Yüklü
 yüksek

* Alternatif Yolu

$$y_i = b \cdot x_i$$

$$R^2 = \frac{\sum_{i=1}^n b^2 x_i^2}{\sum_{i=1}^n y_i^2}$$

$$= b^2 \cdot \frac{\sum_{i=1}^n x_i^2}{\sum_{i=1}^n y_i^2}$$

$$b = \frac{\sum_{i=1}^n y_i x_i}{\sum_{i=1}^n x_i^2}$$

yerine koy (b'yi)

$$= \frac{\sum_{i=1}^n y_i x_i}{\sum_{i=1}^n x_i^2} \cdot \frac{\sum_{i=1}^n x_i^2}{\sum_{i=1}^n y_i^2}$$

$$\Rightarrow b = \frac{\sum_{i=1}^n y_i x_i}{\sum_{i=1}^n x_i^2}$$

yerine koy (b'yi)

$$\Rightarrow \frac{\sum_{i=1}^n y_i x_i}{\sum_{i=1}^n x_i^2} \cdot \frac{\sum_{i=1}^n y_i x_i}{\sum_{i=1}^n x_i^2} = \frac{(\sum_{i=1}^n y_i x_i)^2}{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2}$$

> Basit doğrusal regresyon modelinde R^2 bulma

$\text{corr}(Y_i, X_i) = r_{xy}$ Sadece basit doğrusal regresyon modelinde

$$(r_{xy})^2 = R^2$$

b'ye göre türev $\frac{\partial \sum_{i=1}^n e_i^2}{\partial b} = 2 \sum_{i=1}^n (Y_i - a - bX_i)(-X_i) = 0$

$$= -2 \sum_{i=1}^n (Y_i - a - bX_i)(X_i) = 0$$

$$= \sum_{i=1}^n (Y_i - a - bX_i)(-X_i) = 0$$

$$= \sum_{i=1}^n Y_i X_i - a \sum_{i=1}^n X_i - b \sum_{i=1}^n X_i^2 = 0$$

$$\sum_{i=1}^n Y_i X_i = a \sum_{i=1}^n X_i + b \sum_{i=1}^n X_i^2 \rightarrow \text{OLS T. Normal Denklemleri}$$

→ İkinci Türev: Kalan karelerin ikinci türevi alındığında minimum olmasını bakılmalıdır.

$$\frac{\partial^2 \sum_{i=1}^n e_i^2}{\partial a^2} = 2n > 0$$

→ a'nın ikinci türevi

$$\frac{\partial^2 \sum_{i=1}^n e_i^2}{\partial b^2} = 2 \sum_{i=1}^n X_i^2 > 0$$

→ b'nin ikinci türevi

ORTALAMALARDAN SAPMALAR ÇİNSİNDEN OLS (doğru en küçük kareler)

$$\sum_{i=1}^n Y_i = n \cdot \bar{Y} + b \sum_{i=1}^n X_i \rightarrow \text{denklem (bulduğumuz)}$$

Ortalama bulmak için a'ye b' bulundu $\rightarrow \hat{Y} = a + b\bar{X} \rightarrow a = \bar{Y} - b\bar{X}$ a'yı bulmak için form

$X_i - \bar{X} = x_i$ → X serisinden X ortalamaya aktarırsa küçük ve buluruz

$$X_i - \bar{X} = x_i \quad / \quad Y_i - \bar{Y} = y_i$$

$$e_i = Y_i - \hat{Y} = Y_i - \bar{Y} - b\bar{X}$$

$$e_i = Y_i - \hat{Y}$$

$$Y_i = \alpha + \beta X_i + u_i$$

$$\bar{Y} = \alpha + \beta \bar{X} + \bar{u}$$

$$(Y_i - \bar{Y}) = (\alpha - \bar{\alpha}) + (\beta X_i - \beta \bar{X}) + (u_i - \bar{u})$$

$$y_i = 0 + \beta x_i + u_i$$

Ortalama

$Y_i = \beta x_i + u_i$ → Ortalamadan sapmalar çinsinden analitik-regres ilişkisi

$$\hat{Y}_i = a + bX_i$$

$$\hat{Y}_i = \bar{a} + b\bar{X}$$

$\hat{y}_i = b x_i \rightarrow$ Örneklerden elde edilen regresyon doğrusu.

$$e_i = y_i - \hat{y}_i$$

Ortalama dışından örnekler küçük örneklerle, orijinal gözlemlerle ya da ortalamalar dışından örneklerle ya da ortalamalar dışından örneklerle kesme terimi alınır.

Ortalama dışından örnekler mi yoksa orijinal gözlem değerleri mi dikkat.

Kesmesi: Ortalama dışından örnekler mi yoksa orijinal gözlem değerleri mi dikkat.

$$Y_i = \bar{Y} + e_i$$

$$\sum_{i=1}^n Y_i^2 = \sum_{i=1}^n \bar{Y}^2 + \sum_{i=1}^n e_i^2$$

$$\frac{\sum_{i=1}^n Y_i^2}{n} = \frac{\sum_{i=1}^n \bar{Y}^2}{n} + \frac{\sum_{i=1}^n e_i^2}{n}$$

$$1 = R_{uc}^2 + \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n Y_i^2}$$

Diger R'den farkı, diger ortalama dışından örnekler mi yoksa orijinal gözlem değerleri mi dikkat.

$$R_{uc}^2 = 1 - \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n Y_i^2} \rightarrow 0 \leq R_{uc}^2 \leq 1 \rightarrow R_{uc}^2 \approx 0 \text{ uyumun iyiliği düşüktür}$$

Yorum: Açıklayıcı değişkenlerdeki değişimlerin % kaçını açıklıyor.

BDM (Basit Doğrusal Regresyon Modeli):

$$Y_i = \alpha + \beta X_i + u_i$$

$$E(Y_i | X_i) = \alpha + \beta X_i$$

$$X_i = 0 \rightarrow E(Y_i | X_i) = \alpha$$

$$\frac{\Delta E(Y_i | X_i)}{\Delta X_i} = \frac{\Delta (\alpha + \beta X_i)}{\Delta X_i} = \beta$$

$\beta > 0$
 $\beta < 0$
 $\beta = 0$

Örnek:

$$Q_i^d = f(P_i) \rightarrow \text{Talep}$$

$$Q_i^d = \alpha + \beta P_i + u_i$$

$$E(Q_i^d | P_i) = \alpha + \beta P_i$$

$$P_i = 0 \rightarrow E(Q_i^d | P_i) = \alpha > 0 \rightarrow \text{Potansiyel Talep (Fiyat sıfır olsa bile talep var)}$$

$$\frac{\Delta E(Q_i^d | P_i)}{\Delta P_i} = \frac{\Delta (\alpha + \beta P_i)}{\Delta P_i} = \beta < 0 \rightarrow \text{Normal Mal (Fiyat bir birim artarsa talep "B" kadar azalır)}$$

Az olursa: Azalır, pozitifte, negatifte olabilir.

$\rightarrow A_{m2}$

$$Q_i^s = f(P_i) \rightarrow A_{m2}$$

$$Q_i^s = \alpha + \beta P_i + u_i$$

$$E(Q_i^s | P_i) = \alpha + \beta P_i$$

$$P_i = 0 \rightarrow E(Q_i^s | P_i) = \alpha > 0$$

$$\frac{\Delta E(Q_i^s | P_i)}{\Delta P_i} = \frac{\Delta (\alpha + \beta P_i)}{\Delta P_i} = \beta > 0$$

Yorum: Fiyattaki bir birimlik artış karşılığında "B" artışı gösterir $\uparrow - \uparrow$

$\rightarrow T_{k2}$

$$C_i = \alpha + \beta Y_i + u_i$$

$$E(C_i | Y_i) = \alpha + \beta Y_i$$

$$Y_i = 0 \rightarrow E(C_i | Y_i) = \alpha > 0$$

$$\frac{\Delta E(C_i | Y_i)}{\Delta Y_i} = \frac{\Delta (\alpha + \beta Y_i)}{\Delta Y_i} = \beta \rightarrow 0 \leq \beta \leq 1$$

Yorum: Gelirde bir birimlik artış karşılığında "B" artışı gösterir $\uparrow - \uparrow$

Gelir olmasa bile tüketim yapılır.

$\alpha = 250 \rightarrow$ Hiç gelir olmasa bile

250 TL'lik tüketim yapılır.

Sıfıra yaklaşırsa, gelirin az bir bölümü tüketiliyor. Birde yaklaşırsa, gelirin çok bölümü tüketiliyor.

$\beta = 0,85 \rightarrow$ Gelirin %85'i tüketiliyor.

Orjinal tüketim miktarı

$$S_u^2 = \frac{\sum_{i=1}^n e_i^2}{n-k}$$

$$S_p^2 = S_u^2 \cdot \frac{\sum_{i=1}^n x_i^2}{\sum_{i=1}^n y_i^2} = S_u^2 \cdot \frac{\sum_{i=1}^n x_i^2}{n \cdot \text{var}(x_i)}$$

$$S_p^2 = S_u^2 \cdot \frac{1}{\sum_{i=1}^n y_i^2} = S_u^2 \cdot \frac{1}{n \cdot \text{var}(x_i)}$$

$$S_p^2 = S_u^2 \cdot \frac{1}{\sum_{i=1}^n y_i^2} = S_u^2 \cdot \frac{1}{n \cdot \text{var}(x_i)}$$

$\sqrt{S_u^2} = S_u \rightarrow$ A'nın standart hatası

$\sqrt{S_b^2} = S_b \rightarrow$ B'nin standart hatası

OLS Tahmincisinin Sıfır (Küçük) Örnekli Özellikleri (Varsayımları)

1.) Yansızlık: Aşağıdaki model esas alındığında yansızlık için gereken varsayımlardan ilk 3'ü de varsayımlar geçerli kabul edilir. $V_1 \rightarrow V_3$

$$Y_i = \alpha + \beta x_i + u_i$$

$$Y_i = \alpha + \beta x_i + u_i$$

$E(b|X) = \beta$
 $E(a|X) = \alpha$

Tahminlerin koşullu varyansları örneklerine eşitse yansızlık geçerlidir

2.) BLUE = En iyi doğrusal yansız koşulları: 1-Doğrusal olacak, 2-Yansız olacak, 3-En iyi olacak
 V_1 'den $\rightarrow V_3$ 'e kadar geçerli olması gerekir. Tüm doğrusal yansız tahminciler arasında minimum (en küçük) varyansa sahip tahminci.

3.) Etkinlik = V_1 'den $\rightarrow V_3$ 'e kadar geçerli olup birde V_4 'de geçerli ise OLS tahmincisi etkindir. Hipotez testleri ve aralık tahminleri yapmak için kullanılır.
 $V_1 \rightarrow V_3 + V_4 \rightarrow$ Tüm yansız tahminler arasında en küçük varyansa sahip tahminci etkin tahmincidir. (Doğrusal ya da doğrusal dışı)

TİP I HATA VE TİP II HATA (iddia olmadan hipotez olmaz)

	H_0 Reddedilmesi	H_0 Red	
H_0 Doğru	Doğru Karar 1- $\alpha \rightarrow$ Güven Düzeyi	(Yanlış Karar) Tip II Hata $1-\beta \rightarrow$ Anlamlılık Düzeyi	! Diğer α ile β ile ilişkisi yoktur 100 kereden 1 kere tip I hata yapmama için varsa OLS anlamlılık düzeyi, 5 hata veriyorsa OLS anlamlılık düzeyi
H_0 Yanlış	(Yanlış Karar) Tip I Hata α Hatası	Doğru Karar 1- $\beta \rightarrow$ Testin Gücü	100 kereden 95 kereden doğru yapmak için 5 kere yanlış yapmak (güven düzeyi)

! Doğru hipotezi reddedersek Tip I Hataya düşmüş oluruz.

! Önemli sonuçlar:

- 1.) Tip II hata olasılığı testin anlamlılık düzeyi (α) olan tip I hata ile ters orantılıdır.
 - 2.) Baş hipotez parametresi; $\beta=0$ iken β nin (parametre) bilinmeyen değeri 0'a ne kadar yakın ise tip II hata olasılığı o kadar büyütür.
 - 3.) Tip I hata düzeyi veri ile örneklem büyüklüğü arttıkça Tip II hata olasılığı azalmaktadır.
- ? İlk tip hatadan birinden kaçınmak istense hangisinden kaçınmak istenir? I Tip hata daha tehlikeli.

Parametre Tahmini: Çift taraflı alternatifte göre çalışır. Bu yüzden anlamlılık düzeyi $\alpha/2$ olur.

$$\mu = a - \alpha \rightarrow \alpha \text{ yalınca} \rightarrow P(a - z^* \sigma_a \leq \mu \leq a + z^* \sigma_a) = 1 - \alpha$$

σ_a bilinmiyor σ üzerinden tahmini kullandık

$$P(a - t_{n-k}^* \sigma_a \leq \mu \leq a + t_{n-k}^* \sigma_a) = 1 - \alpha$$

$\rightarrow z$ = tablo değeri
1.96 - 2.58
0.05 - 0.01

$$\beta = b - \beta \rightarrow \beta \text{ yalınca} \rightarrow P(b - z^* \sigma_b \leq \beta \leq b + z^* \sigma_b) = 1 - \alpha$$

σ_b bilinmiyor σ üzerinden tahmini kullandık

$$P(b - t_{n-k}^* \sigma_b \leq \beta \leq b + t_{n-k}^* \sigma_b) = 1 - \alpha$$

$\rightarrow z$ 'yi kullanmıyoruz çünkü sigma'lar bilinmiyor.

PRAKİK T-TESTİ: Çift taraflı test mantığına göre çalışır. Her parametrenin t -istatistiğinin mutlak değerce 2'den büyük olduğu durumlarda ilgili parametrenin %5'de, mutlak değerce 3'ten büyük olduğu durumlarda ise ilgili parametrenin %1'de istatistiksel olarak anlamlı olduğunu ima eder.

"P" DEĞERİ: Bu değer mantıklı anlamlılık düzeyi olarak bilinmektedir. İma ettiği şey "baş hipotezin reddedilebileceği minimum anlamlılık düzeyi"dir. Eğer bu değer %5'ten küçükse %5'te, %1'den küçükse %1'de istatistiksel anlamlılığı ifade eder. Çift taraflı test mantığında çalışır.

21 - ARALIK

EXCEL ÖRNEĞİ: CORONER

2010 yılı için yapılan çalışma: İki ilen sigaraların hastalıkla ilgisi var mıdır? (Binim=1000)

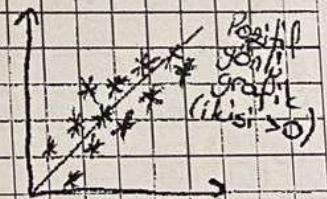
$Y_i \rightarrow i$ ilkesinde koroner kalp yetmezliğinden ölen kişi sayısı (1000 kişi)

$X_i \rightarrow i$ ilkesinde tüketilen sigara sayısı

Model: $Y_i = \alpha + \beta X_i + u_i$ Bağıntı: $Y_i = f(X_i)$

Ölçüm: $E(Y_i | X_i) = \alpha + \beta X_i$ $X_i = 0 \rightarrow E(Y_i | X_i) = \alpha \geq 0$

Tahmin: $\hat{Y}_i = \alpha + \beta X_i$ $\frac{\Delta E(Y_i | X_i)}{\Delta X_i} = \frac{dE(Y_i | X_i)}{dX_i} \rightarrow \beta > 0$



1) Önce en küçük kareler yöntemiyle parametre tahminlerini yapıp yorumlayıp, bekleyişle uyumunu kontrol edelim.

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

$$b = \frac{\sum_{i=1}^n y_i x_i}{\sum_{i=1}^n x_i^2} \rightarrow b = \frac{\sum_{i=1}^n (y_i - \bar{Y}) \cdot (x_i - \bar{X})}{\sum_{i=1}^n (x_i - \bar{X})^2} \Rightarrow b = \frac{727036,6667}{13056523,8095} = 0,0557$$

$\rightarrow$ Sigara tüketimi bin paket artarsa koroner kalp yetmezliği 56 (0,0557 x 1000) kişi artar.

$$a = 149,0667 - (0,0557 \times 2148,0952) = 29,4526 \rightarrow$$

Hıa sigara içilmese bile bir ülkede ortalama 29 bin kişi koroner kalp yetmezliğinden ölür.

$$\hat{Y}_i = 29,4526 + 0,0557 \cdot X_i$$

*Parametre tahminlerinin bekleyişine uygun. Kesme terimin de pozitif, eğim parametresi de pozitif sonuçlandı.

14 - ARAKIS

NORMAL DAĞILIM: İtata teriminin x koşulu altında dağılımı (sıfır ortalama ve sigma kare varyans) normal sekildedir. Yani;

$$X \sim N(0, \sigma^2)$$

(i) $Y \sim N(0, \sigma^2) \rightarrow$ (i) non-kosullu dağılımı normalse kosullu dağılımı da normaldir.

(ii) Normal dağılım birinci momentle (ortalama) ve ikinci momentle (varyansla) karakterize edilebilir.

(iii) Normal dağılıma dogrusali kombinasyon giren tüm rassal degiskenlerin dağılımı da normaldir.

$$Y_i = \alpha + \beta X_i + u_i \quad \text{İtata terimi normal dağılıyorsa (u)}$$

$$Y_i | X \sim N(\alpha + \beta X_i, \sigma^2) \rightarrow Y \text{ de normal demek zorunda}$$

dogrusal kombinasyonu

(iv) Normal dağılım genelliyse bagimsizlik ve korelasyonluk geçerlidir

$$a | X \sim N(\alpha, \sigma_a^2)$$

$$b | X \sim N(\beta, \sigma_b^2)$$

Tahminlerinde normal dağılmaktadır.

* Herhangi bir tahmini ya da rassal degiskeni ortalamasından cikartıp standart sapmasına bölündüğünde dağılım standartlaşır

$$Z_a = \frac{a - \alpha}{\sigma_a} \sim N(0, 1)$$

$$Z_b = \frac{b - \beta}{\sigma_b} \sim N(0, 1)$$

Standart normal dağılım.

Hipotez testlerinde ve aralık tahminlerinde kullanılır.

* Eğer tahminciyle igrasıyorsak (gözetim sayısı 30'dan küçükse) t dağılımı kullanılır.

$$Z_a = \frac{a - \alpha}{\sigma_a} \sim N(0, 1)$$

(Anlık parametre)

Hiçbir parametre bilinmez

$$\frac{a - \alpha}{S_a} \sim t(n-k)$$

Tahminci

Hipotez Testlerinde

Kullanılarak

$$Z_b = \frac{b - \beta}{\sigma_b}$$

Hiçbir parametre bilinmez

$$\frac{b - \beta}{S_b} \sim t(n-k)$$

Tahminci

HİPOTEZ TESTLERİ: Öncelikli olarak anlamlılık düzeyine ihtiyaç vardır. Anlamlılık

Düzeyi 1. tip hata yapma olasılığıdır. 1. tip hata yapma olasılığı (2) dogru hipotezi reddetmedir. 100 kez denendiğinde en fazla 5 kez reddedebilirsek 0.05 anlamlılık düzeyidir. 100 kez denendiğinde en fazla 1 kez reddedebilirsek 0.01 anlamlılık düzeyidir. Borsedilen anlamlılık düzeyidir ve 1. tip hatadan söz edilir. n =gözetim sayısı, k =parametre

$$Y_i = \alpha + \beta X_i + u_i$$

parametre

$$t_a = \frac{a - \alpha}{S_a} \sim t(n-k)$$

$$t_b = \frac{b - \beta}{S_b} \sim t(n-k)$$

parametre

H_0 : Alternatifin iddia edilmediği durumlarda geçerliliği kabul edilen hipotez.

$$H_0: \alpha = 0$$

H_1 : Alternatif hipotez

$$H_1: \alpha > 0 \quad \text{Bekleyiş varsa}$$

$$\alpha < 0$$

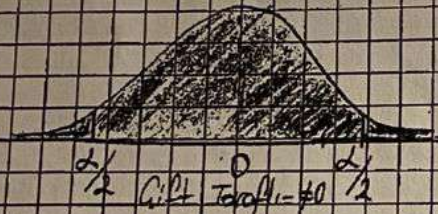
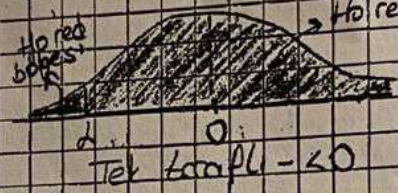
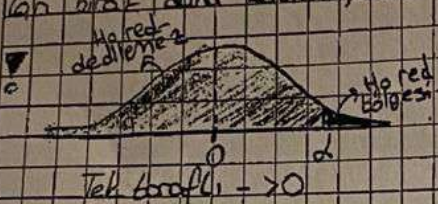
$$\alpha \neq 0 \rightarrow \text{Bekleyiş yoksa}$$

Beta Testi : $H_0: \beta = 0$
 $H_1: \beta \neq 0$
 $\beta > 0$
 $\beta < 0$

Yokluğunun ne esit olduğu durum için "kesme terimi" ve eğim parametresi "istatistiksel olarak anlamlıdır" denir. Eğer reddedersek "eğim parametresi ve kesme terimi" istatistiksel olarak anlamlıdır" denir.

▼ Eğer alternatif " >0 ", " <0 " şeklinde kurulursa hipotez testleri "tek taraflı olur". " $\neq 0$ " dersen hipotez testi çift taraflı olur. Tek taraflıda anlamlılık düzeyi $\alpha/2$ (0,05, 0,01) olur. Çift taraflıda $\alpha/2$ (0,025, 0,005) olur.

▼ χ^2 dağılımı özellikleri: Ortalaması sıfır, ne basıktır ne de dik, sadece kuyrukları biraz daha kalındır, simetrik, normal dağılıma benzer.



- 1) Belirletim neye alternatifte α işaret yapılır " $>$ ", " $<$ ". Eğer belirti yoksa " $\neq$ " yapılır.
- 2) Anlamlılık düzeyine karar verilir. Araştırmacı karar verir ister %0.5, ister %0.1 olur.

▼ Talep Modeli

$Q^d = a + bP_i + u_i$
 $Q^d \rightarrow$ Talep Miktarı
 $P_i \rightarrow$ Malin Fiyatı
 $P_i \uparrow \rightarrow Q^d \downarrow \rightarrow$ Negatif ilişki (fiyat artarsa talepten azalır)
 $dQ^d / dP_i > 0$
 $P_i = 0 \rightarrow Q^d > 0 \rightarrow a > 0$ Bekleyişler
 $Q^d = a + bP_i$ için hipotezleri yaz.

1) A için

$H_0: a = 0$
 $H_1: a > 0$
 $t_a = \frac{a - 0}{\hat{a}} \sim t_{n-2}$
 $t_a > t_{n-2, \alpha} \rightarrow H_0 \text{ red}$
 Kesme terimi anlamlı bir şekilde 0'dan büyüktür (Kesme terimi istatistiksel olarak anlamlıdır).

2) B için

$H_0: \beta = 0$
 $H_1: \beta < 0$
 $t_b = \frac{b - 0}{\hat{b}} \sim t_{n-2}$
 $t_b < -t_{n-2, \alpha} \rightarrow H_0 \text{ red}$
 Eğim parametresi anlamlı bir şekilde 0'dan küçüktür (Kesme terimi istatistiksel olarak anlamlıdır).

▼ Tüketim Modeli

$C_i = a + bY_i + u_i$
 $C_i \rightarrow$ Tüketim
 $Y_i \rightarrow$ Gelir
 $Y_i = 0 \rightarrow C_i > 0 \rightarrow a > 0$
 $Y_i \uparrow \rightarrow C_i \uparrow$
 $dC_i / dY_i = \beta$

a için

$H_0: a = 0$
 $H_1: a > 0$
 $t_a = \frac{a - 0}{\hat{a}} \sim t_{n-2}$
 $t_a > t_{n-2, \alpha} \rightarrow H_0 \text{ red}$
 $t_b = \frac{b - 0}{\hat{b}} \sim t_{n-2} \rightarrow t_b > t_{n-2, \alpha} \rightarrow H_0 \text{ red}$

$0 \leq \beta \leq 1 \rightarrow$ Aralık değeri için $\alpha/2$

2.) Uygunluğun iyiliğini (determinant katsayısı) (R^2) hesaplayıp yorumlayınız.

$$R^2 = 1 - \frac{\sum_{i=1}^n (e_i)^2}{\sum_{i=1}^n y_i^2} = 1 - \frac{41181.4531}{81665.6067} = 0.4957$$

*Yorum: Tüketilen sigara sayısında değişimler koroner kalp yetmezliğinden ölen insan sayısındaki değişimlerin yaklaşık %50'sini açıklar.

*Tüketilen sigara sayısı ile ölen insan sayısı için %50'lik uyum iyiliği oranı çokta düşük sayılmaz.

3.) Ortalamalar noktasındaki elastikiyeti hesaplayıp yorumlayınız.

$$\hat{\eta} = b \cdot \frac{\bar{X}}{\bar{Y}} = 0.0557 \cdot \frac{2148.0852}{149.4015} = 0.802$$

*Yorum: Sigara tüketimindeki %1 artış karşısında koroner kalp yetmezliğinden ölen insan sayısı %0.802 artar.

Regressiyonun ($\hat{Y}$ 'de)

x değişen y'ye $\hat{Y}$ yazılır

Elastikiyet: oransal değişimin oransal değişime oranını hesaplar.

Yorum yaparken ölen insan sayısı aynı % koy.

* (%1'den küçük olduğu için): Sigara tüketimi karşısında koroner kalp yetmezliğinden ölen inelastiktir.

$\hat{\eta} \geq 1 \Rightarrow$ Elastik $= 1 \Rightarrow$ Birim Elastik

4.) Tahmincilerin varyanslarını ve standart hatalarını hesaplayınız.

$$S_u^2 = \frac{\sum_{i=1}^n e_i^2}{n-k} = \frac{41181.4531}{21-2} = 2167.4449$$

21-2
v.b. parametre

$$S_a^2 = \frac{S_u^2 \cdot \sum_{i=1}^n x_i^2}{n \cdot \sum_{i=1}^n x_i^2} = \frac{(2167.4449) \cdot (109957100.0)}{21 \cdot (13056523.8095)} = 869.2095 \Rightarrow \sqrt{S_a^2} \Rightarrow S_a = 29.4824$$

$\Rightarrow S_a^2$ bilinmiyor

Amin
standart hatası

$$S_b^2 = \frac{S_u^2}{\sum_{i=1}^n x_i^2} = \frac{2167.4449}{13056523.8095} = 0.0002 \Rightarrow \sqrt{S_b^2} \Rightarrow S_b = 0.0129$$

$\Rightarrow S_b^2$ bilinmiyor

b'nin
standart hatası

5.) Sırası ile %5 ve %1'de kesme ve eğim parametre tahminlerinin bireysel anlamlılıklarını test ediniz (sırayla)

$$H_0: \alpha = 0$$

$$H_1: \alpha > 0$$

baş hipotez
yoktur

$$H_0: \beta = 0$$

$$H_1: \beta > 0$$

$$t_a = \frac{a - \alpha}{S_a} = \frac{29.4526 - 0}{29.4824} = 0.9990$$

$$t_b = \frac{b - \beta}{S_b} = \frac{0.0557}{0.0129} = 4.3218$$

$$t_{a(0.05)} = 1.729$$

$$t_{b(0.01)} = 2.539$$

$$t_a < t_{a(0.05)} \rightarrow H_0 \text{ reddedilmez}$$

$$t_b > t_{b(0.01)} \rightarrow H_0 \text{ red}$$

Yorum: %5 anlamlılık düzeyinde kesme terimi istatistiksel olarak anlamlı değildir.

Yorum: 1x %1 anlamlılık düzeyinde eğim parametresi istatistiksel olarak anlamlıdır.

2x Eğim parametresi %1 anlamlılık düzeyinde anlamlı bir şekilde sıfırdan büyüktür.

* Kesme terimi %5 anlamlılık düzeyinde 0'dan farklı değildir.

1) Olgun en küçük kareler yöntemiyle parametre tahminlerini yapıp yorumlayıp bakiyisine uyumunu tartışınız.

$$a = \bar{Q} - b\bar{P} \quad b = \frac{-481,70}{148,90} = -3,24$$

$$b = \frac{\sum Q_i P_i - \bar{Q} \cdot \bar{P}}{\sum P_i^2 - \bar{P}^2} \quad a = 397,3 - (-3,24 \cdot 23,9) = 674,62$$

$$\hat{Q}_i = a + bP_i \quad \hat{Q}_i = 674,62$$

a) Yorum: Peynir fiyatı sıfır iken peynir talebi 674,62 kilogramdır.
b) Yorum: Peynir fiyatı 1 TL artarsa peynir talebi 3,24 kilogram azalır.
c) Yorum: Parametre tahminlerin bakiyesine uygun. Kesme terimi pozitif, eğim parametresi negatif sonuçlandı.

2) Uyum: iyiliğini (determinant katsayısını - R^2 'yi) hesaplayıp yorumlayınız.

$$R^2 = 1 - \frac{\sum e_i^2}{\sum Q_i^2} = 1 - \frac{1879,77}{3638,10} = 0,45$$

Yorum: Peynir talebindeki değişmelerin 0,45'i peynir fiyatındaki değişmeler tarafından açıklanmaktadır. Uyumun iyiliğine yüksek derebılır.

* Başka bir yol:

$$R^2 = b^2 \cdot \frac{\sum P_i^2}{\sum Q_i^2} = (-3,24)^2 \cdot \frac{148,90}{3638,10} = 0,45$$

3) Ortalamalar noktasındaki elastikiyeti hesaplayıp yorumlayınız.

$$\eta = b \cdot \frac{\bar{P}}{\bar{Q}} = (-3,24) \cdot \frac{23,90}{397,30} = -0,194608$$

$\bar{Q}$ lerde P için $\bar{P}$ yerleştir.

Yorum: Ortalamalar noktasında peynir fiyatındaki %1 artış karşısında peynir talebi %0,1946 azalır. Peynir talebi, peynir fiyatı karşısında inelastiktir.

* Peynir fiyatının 27 TL olduğu noktada elastikiyeti hesaplayıp yorumlayınız.

$$\eta = (-3,24) \cdot \frac{27}{397,30} = -0,22$$

Yorum: Peynir fiyatının 27 TL olduğu noktada fiyat-taki %1 artış karşısında talep edilen peynir miktar %0,22 azalır. Fiyatın 27 TL olduğu noktada peynir fiyatı karşısında peynir talebi inelastiktir.

4) Tahminlerin varyanslarını ve standart hatalarını hesaplayınız.

$$S_b^2 = \frac{\sum e_i^2}{n-k} = \frac{1879,77}{10-2} = 234,97 \quad S_b^2 = \frac{S_u^2}{\sum P_i^2} = \frac{234,97}{148,9} = 1,58$$

$$S_a^2 = \frac{S_u^2 \cdot \sum P_i^2}{n \cdot \sum P_i^2} = \frac{234,97 \cdot 5090}{10 \cdot 148,9} = 792,18 \quad S_b \text{ in standart hatası}$$

$$\sqrt{S_b^2} = S_b = 15,15$$

Ayrıca standart hatası

7) Sırası ile %1 ve %5 anlamlılık düzeyinde parametrelerin bireysel anlamlılıklarını sıralayınız.

$$H_0: \alpha = 0$$

$$H_1: \alpha > 0$$

$$t_a = \frac{a - \alpha}{S_a}$$

$$t_a = \frac{476,62 - 0}{28,15} = 16,86$$

$$t_a^{0,01} = 2,896$$

$$t_a > t_a^{0,01} \rightarrow H_0 \text{ red}$$

$$H_0: \beta = 0$$

$$H_1: \beta < 0$$

$$t_b = \frac{b - \beta}{S_b}$$

$$t_b = \frac{-3,24 - 0}{1,26} = -2,58 \rightarrow \text{Mutlak değerce bak}$$

$$t_b^{0,05} = 1,860$$

$$t_b > t_b^{0,05} \rightarrow H_0 \text{ red}$$

Yorum: %1 anlamlılık düzeyinde kesme terimi istatistiksel olarak anlamlıdır.
* Kesme terimi %1 anlamlılık düzeyinde anlamlı bir şekilde sıfırdan büyüktür.

Yorum: %5 anlamlılık düzeyinde eğim parametresi istatistiksel olarak anlamlıdır.
* Eğim parametresi %5 anlamlılık düzeyinde anlamlı bir şekilde sıfırdan küçüktür.

❗ β hipotezlerinde: $t_a^{0,01} = 2,896$ olarak alınsaydı, $t_b < t_b^{0,01}$ olucaktı ve H_0 reddedilemezdi;

* 0,01 anlamlılık düzeyinde eğim parametresi anlamlı bir şekilde sıfırdan farklı değildir.
* 0,01 anlamlılık düzeyinde eğim parametresi istatistiksel olarak anlamlı değildir.

6) Sırası ile %1 ve %5 anlamlılık düzeyinde kesme ve eğim parametrelerinin güven aralıklarını hesaplayıp yorumlayınız.

$$* P(a - t_{n-k}^{1/2} * S_a \leq \alpha \leq a + t_{n-k}^{1/2} * S_a) = 1 - \alpha$$

$$P(476,62 - 3,355 * 28,15 \leq \alpha \leq 476,62 + 3,355 * 28,15) = 1 - 0,01$$

$$P(380,19 \leq \alpha \leq 569,05) = 0,99$$

Yorum: 100 kere tahmin edilse bunun 99'unda kesme terimi 380,19 ile 569,05 arasında olacaktır.

$$* P(b - t_{n-k}^{1/2} * S_b \leq \beta \leq b + t_{n-k}^{1/2} * S_b) = 1 - \alpha$$

$$P(-3,24 - 2,306 * 1,26 \leq \beta \leq -3,24 + 2,306 * 1,26) = 1 - 0,05$$

$$P(-6,13 \leq \beta \leq -0,34) = 0,95$$

Yorum: 100 kere tahmin edilse bunun 95'inde eğim parametresi -6,13 ile -0,34 arasında olacaktır.

❗ β eğimi 0,01 kabul edilirse;

$$P(-3,24 - 3,355 * 1,26 \leq \beta \leq -3,24 + 3,355 * 1,26) = 0,99$$

$$P(-7,45 \leq \beta \leq 0,98)$$

↳ Sağ taraf pozitif çıktı, H_0 reddedilemez olur. Çünkü sıfır bu aralığa düşer.
-, - aralığı ya da +, + aralığında sorun yoktur. -, + olursa sıfır olur.

7) Pratik t-testine ve marginal anlamlılıklarına göre bireysel anlamlılıkları sıralayınız.

* Pratik t-test çift taraflı: α 2'den ve 3'ten büyük, 0,01'de anlamlı

β 2'den büyük 3'ten küçük, 0,05'de anlamlı, 0,01'de anlamlı

* P değeri çift taraflı test: α %5'ten de %1'den de küçük %0,1'de anlamlı

β %5'ten küçük, %1'den büyük, %0,5'te anlamlı

6) Sırası ile %5 ve %1 anlamlılık düzeyinde kesme ve eğim parametrelerinin güven aralıklarını oluşturunuz. (güven aralığı dediği için $\frac{\alpha}{2}$)

$$* P(a - t_{(n-k)}^{1/2} * S_a \leq \alpha \leq a + t_{(n-k)}^{1/2} * S_a) = 1 - \alpha$$

$$P(29,4526 - 2,093 * 29,4824 \leq \alpha \leq 29,4526 + 2,093 * 29,4824) = 1 - 0,05 = 0,95$$

$$P(-32,1655 \leq \alpha \leq 91,0707)$$

Yorum: 100 kez tahmin edilse bunun 95'inde kesme terimi -32,1655 ile 91,0707 arasında çıkar...

$$* P(b - t_{(n-k)}^{1/2} * S_b \leq \beta \leq b + t_{(n-k)}^{1/2} * S_b) = 1 - \alpha$$

$$P(0,0557 - 2,861 * 0,0129 \leq \beta \leq 0,0557 + 2,861 * 0,0129) = 1 - 0,01 = 0,99$$

$$P(0,0188 \leq \beta \leq 0,0925)$$

Yorum: 100 kez tahmin edilse bunun 99'unda eğim parametresi 0,0188 ile 0,0925 arasında çıkar

▼ - ile + arasında H_0 zaten reddedilemez. Ama - ile -, + ile + olsaydı bunu diyenezdik. (Çift taraflı test mantığına göre)

▼ Pratik t-testine göre t mutlak değerce 2'den ve 3'den büyük olduğu için anlamsız. β ise; mutlak değerce 2'den büyük ama 3'ten de büyük olduğu için %1'de anlamlı

▼ P değeri testi marginal anlamlılık düzeyinden hesaplanır

▼ Rapor :

$$\hat{Y}_i = 29,4526 + 0,0557 X_i$$

$$(s) = (29,4824), (0,0129)$$

$$(t) = (0,9990), (1,3218)$$

$$R^2 = 0,4957 \quad n = 21$$

28 ARALIK

Örnek: Excel - Peynir örneği

Q_i = Peynir talebi miktarı (kg)

P_i = Peynirin fiyatı (₺)

Model:

$$Q_i^d = \alpha + \beta P_i + \epsilon_i$$

$$E(Q_i^d | P_i) = \alpha + \beta P_i$$

$$\text{örneğin } Q_i = \alpha + \beta P_i + \epsilon_i$$

$$\text{tahmin } \hat{Q}_i = \alpha + \beta P_i$$

Tahmin: $Q_i^d = f(P_i)$

$$P_i = 0 \rightarrow \alpha > 0$$

(Peynir fiyatı sıfır olsa bile)

Peynir talebi vardır Potansiyel)

$$\frac{\Delta E(Q_i^d | P_i)}{\Delta P_i} = \frac{dE(Q_i^d | P_i)}{dP_i} \rightarrow \beta < 0$$

(Peynir fiyatındaki 1₺'lik)

ortalama veya peynir talebi)

β kadar azalır

Sayıları:

