

# İstatistik 2

## Bölüm 14

### Regresyon Analizi 2

# Farkların Kullanımı ile Model Seçimi

- Birinci farklar büyük ölçüde sabit ise doğrusal trend modeli kullanılacak

$$Y_2 - Y_1 = Y_3 - Y_2 = L = Y_n - Y_{n-1}$$

İkinci farklar büyük ölçüde sabit ise ikinci derece trend modeli kullanılacak

$$\left[ (Y_3 - Y_2) - (Y_2 - Y_1) \right] = L = \left[ (Y_n - Y_{n-1}) - (Y_{n-1} - Y_{n-2}) \right]$$

# Farkların Kullanımı ile Model Seçimi

(Devam)

- Yüzdesele farkların sabit olması durumunda ise üstel trend modeli kullanılacaktır

$$\left(\frac{Y_2 - Y_1}{Y_1}\right) 100\% = \left(\frac{Y_3 - Y_2}{Y_2}\right) 100\% = L = \left(\frac{Y_n - Y_{n-1}}{Y_{n-1}}\right) 100\%$$

# Autoregressive Modeling

- Tahmin için kullanılmakta
- Autocorrelationun avantajına sahiptir
  - > 1. sıra – ard arda gelen deęerler arasındaki korelasyon
  - > 2. sıra – 2 döneme ait deęerler arasındaki korelasyon
- $p$ - inci sıra Autoregressive model

$$Y_i = A_0 + A_1 Y_{i-1} + A_2 Y_{i-2} + \dots + A_p Y_{i-p} + \delta_i$$

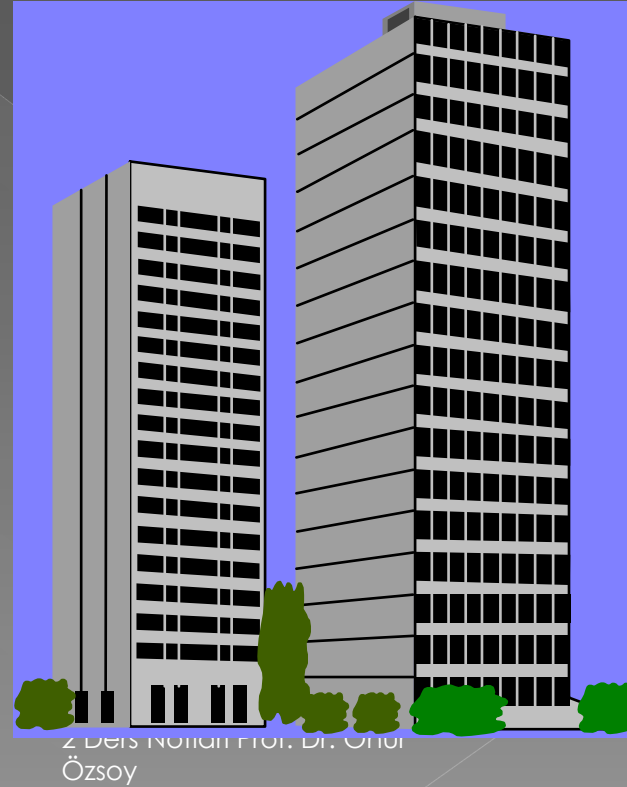
Rassal hata

# Autoregressive Model: Örnek

Aşağıdaki verileri kullanarak Autoregressive Model Modeli oluşturunuz

---

Yıl	Ofis
93	4
94	3
95	2
96	3
97	2
98	2
99	4
00	6



# Autoregressive Model: Örneğin Çözümü

2. Sıra tablosunun  
oluşturulması

Excel kullanılarak regresyon  
denklemini elde edilir

	<b>Coefficients</b>
<b>Intercept</b>	3.5
<b>X Variable 1</b>	0.8125
<b>X Variable 2</b>	-0.9375

Yıl	$Y_i$	$Y_{i-1}$	$Y_{i-2}$
93	4	---	---
94	3	4	---
95	2	3	4
96	3	2	3
97	2	3	2
98	2	2	3
99	4	2	2
00	6	4	2

$$\hat{Y}_i = 3.5 + .8125Y_{i-1} - .9375Y_{i-2}$$

# Autoregressive Model Örnek: Tahmin

İkinci sıra modelini kullanarak 2001 için ofis sayısına ilişkin tahmin yapın

$$Y_i = 3.5 + .8125Y_{i-1} - .9375Y_{i-2}$$

$$Y_{2001} = 3.5 + .8125Y_{2000} - .9375Y_{1999}$$

$$= 3.5 + .8125 \times 6 - .9375 \times 4$$

$$= 4.625$$

# Autoregressive Modelde İzlenecek Aşamalar

1.  $P$ 'yi belirle :  $SD = n - 2p - 1$
2. “lag (gecikme) tahmin edici” serisinden oluşan değişkenleri belirle

$$Y_{i-1}, Y_{i-2}, \dots, Y_{i-p}$$

3. Bütün  $p$  vdeğişkenlerini kullanarak Excel'de regresyon modelini tahmin et

4.  $A_p$  ları önem testinden geçir

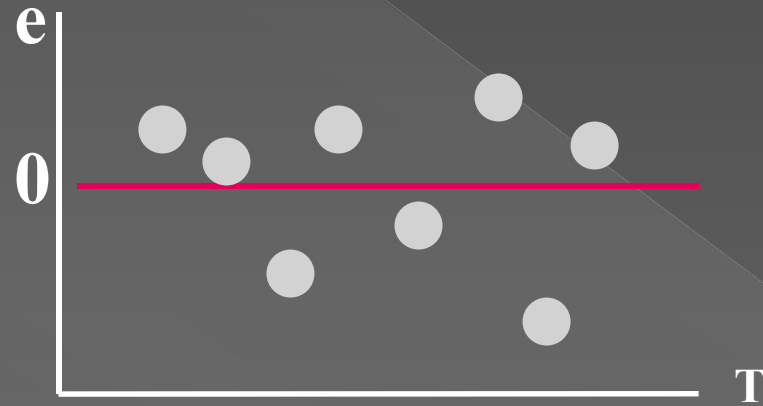
- > Boş hipotez reddedirse, model seçilmeli
- > Boş hipotez kabul edilirse,  $p$ 'yi 1 azaltılarak model yeniden tahmin edilmeli



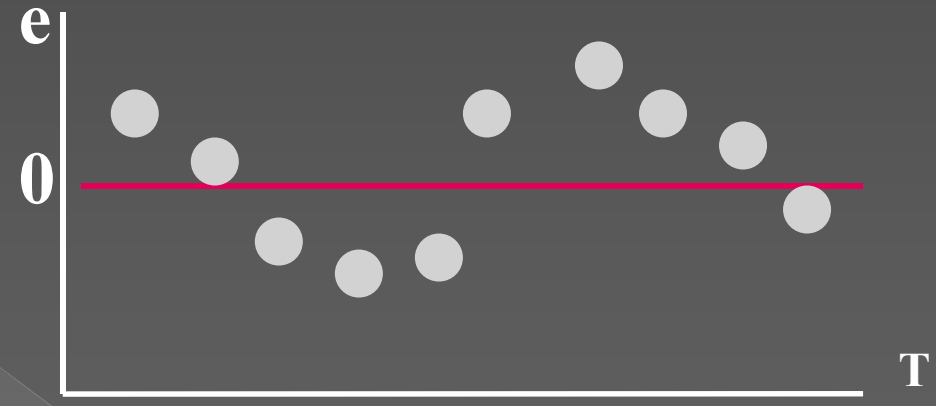
# Tahmin Modelinin Seçimi

- Residual analizi yapılmalı
- pattern yada yönler bakılmalı
- sum of square error - SSE (residual errors) ölçümlenmeli
- Mean Absolute Deviation MAD kullanılarak residual error ölçümlenmeli
- Basit model kullanılmalı

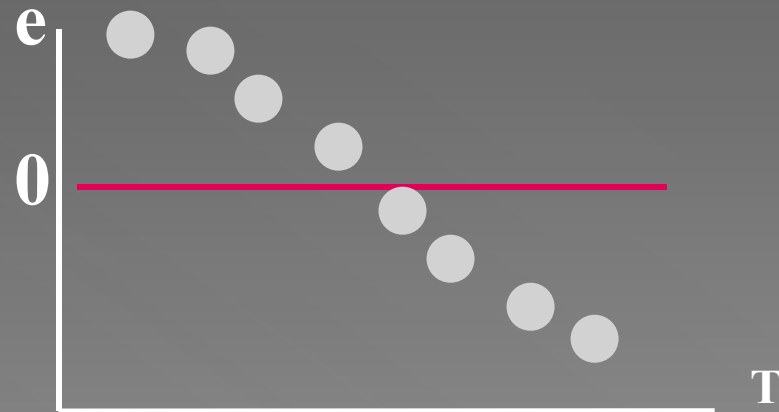
# Residual Analysis



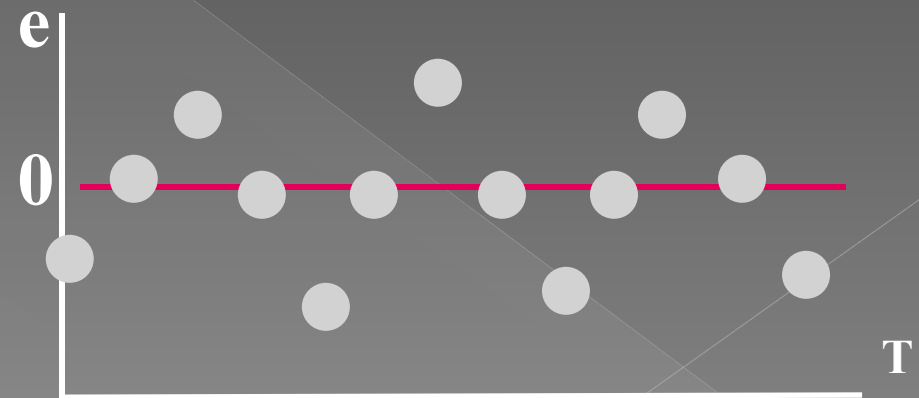
Rastsal Hata



Hesaplanmayan döngüsel etkiler



Hesaplanmayan Trentler



Hesaplanamayan Mevsimsel Etki

# Ölçümleme Hatası

- En küçük ölçümleme hatası veren modeli seç
- Sum of square error (SSE)

> 
$$SSE = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

- > Aşırı uç değerlere oldukça duyarlıdır

# Ölçümlenme Hatası

(Devam)

## ● Mean Absolute Deviation (MAD)

$$MAD = \frac{\sum_{i=1}^n |Y_i - \hat{Y}_i|}{n}$$

> Aşırı uç değerlere duyarlı değildir

# Basit Model Seçimi

- İki veya daha fazla sayıda modelin veri setini iyi açıkladığını varsayalım
- Basit modeli seç
  - > Basit model türleri:
    - En küçük kareler doğrusal modeli
    - İkinci derece doğrusal modeli
    - 1. sıra autoregressive modeli
  - > Kompleks model türleri:
    - 2. Sıra ve 3. Sıra autoregressive modeli
    - Üstel en küçük kareler yöntemi

# Mevsimsel Verilerle Tahmin

- En küçük kareler trend modeli ile beraber kategorik tahmin edici deęişkenler kullan
- Çeyrek dönemlik verilerle üstel trend modeli kullan:

$$\hat{Y} = b_0 b_1^{X_i} b_2^{Q_1} b_3^{Q_2} b_4^{Q_3}$$

- >  $b_j$  ler çeyreğe kıyasla  $i$  inci çeyrek için çarpanırlar
- >  $Q_i = i$  inci çeyrek ise 1 deęilse 0
- >  $X_j =$  zaman dilimini gösteren bağımsız deęişken

# Mevsimsel Verilerle Tahmin : Örnek

Standards and Poor's Composite Stock Price  
Index:

1	445.77	500.71	645.5	757.12
2	444.27	544.75	670.63	885.14
3	462.69	584.41	687.31	947.28
4	459.27	615.93	740.74	970.43

Excel Çıktısı

<i>Regression Statistics</i>	
Multiple R	0.989936819
R Square	0.979974906
Adjusted R Square	0.972693054
Standard Error	0.043867069
Observations	16

$r^2$  is .98

Modelin iyi olduğunu  
göstermektedir

# Mevsimsel Verilerle Tahmin : Örnek

(Devam)

Excel Çıktısı

	<i>Coefficients</i>	<i>Standard Error</i>
Intercept	6.011187697	0.031115484
Coded X	0.055372493	0.002452244
Q1	0.010421639	0.031879168
Q2	0.023885562	0.031404042
Q3	0.019342411	0.031115484

## 1. Çeyrek için regresyon denklemi:

$$\begin{aligned}\ln \hat{Y}_i &= \ln b_0 + X_i \ln b_1 + Q_1 \ln b_2 \\ &= 6.011 + .0553 X_i + 0.0104 Q_1\end{aligned}$$