

## EKONOMETRİ

Ekonometri kelime anlamı olarak “ekonomik ölçme” demektir ve ekonominin kantitatif analizi olarak tanımlanabilmektedir (Yerdelen Tatoğlu, 2023, s. 1).

Ekonometri, iktisatçılar tarafından ortaya konulan her tür karmaşık, karşılıklı neden-sonuç içeren, çok bağlantılı ilişkileri modelleyebilmekte, tahmin edebilmekte ve anlaşılır hale getirebilmektedir. Dolayısıyla ekonometri, esneklikler, eğimler, marjinal değerler gibi iktisadi ilişkilerin parametrelerine sayısal değerler bulmak ve iktisat kuramlarını doğrulamak amacıyla, iktisat, matematik ve istatistiğin birleşimi olarak düşünülebilmektedir (Yerdelen Tatoğlu, 2023, s. 1-2).

### Ekonometrik Araştırmanın Aşamaları

#### 1. İktisat Teorisinin ya da Hipotezin Formülasyonu

Ekonometrik araştırmanın ilk aşaması iyi bir teoriye sahip olmak ve bir hipotez ortaya koymaktır. Araştırmacının çalışmak istediği konu genel olarak bir ekonometrik modelin bağımlı değişkenini belirlemektedir. *Bağımlı* değişken değeri model içinde belirlenen, bağımsız değişken ya da değişkenlere bağlı olarak ortaya çıkan değişken olarak tanımlanmaktadır (Yerdelen Tatoğlu, 2023, s. 2).

Yine bu aşamada incelenen literatür neticesinde ele alınan teoriye uygun olarak modelin bağımsız değişkenleri tespit edilmektedir. Bağımsız değişken, bir iktisadi olayın meydana gelmesini sağlayan, bağımlı değişkeni etkileyen ve değeri model dışında tayin edilmiş olan değişkendir (Yerdelen Tatoğlu, 2023, s. 3).

Bu aşamada, incelenen teori literatür taramasının da yardımıyla tüm ayrıntıları ile değişkenler ve değişkenler arası ilişkileri de içerecek şekilde sözel olarak ortaya konulmaktadır. (Yerdelen Tatoğlu, 2023, s. 3)

#### 2. Teori ya da Hipotezin Matematiksel Modelinin Oluşturulması

Bu aşamada, bir önceki aşamada sözel olarak belirlenmiş olan teori bağımlı ve bağımsız değişkenler ve kullanılacak parametreler yardımıyla matematiksel bir modelle ifade edilmektedir. Dolayısıyla stokastik olan iktisat, matematikselleştirilerek deterministik hale getirilmektedir (Yerdelen Tatoğlu, 2023, s. 3-4).

Bir matematiksel modelde iki tür parametre vardır. Bunlardan birincisi, bağımsız değişken sıfır değerini aldığı anda bağımlı değişkenin aldığı değeri veren, doğrunun Y eksenini kestiği noktayı belirleyen sabit parametredir. Sabit parametreye otonom parametre ismi de verilmektedir. Diğeri ise, bağımsız değişkenin bağımlı değişken üzerindeki etkisini gösteren eğim parametresidir. Modeldeki bağımsız değişkenin bir birimlik değişmesi karşısında, bağımlı değişkenin ortalama olarak

ne kadar değiştiğini göstermektedir. Bu parametrenin sayısal değeri, değişkenler arasındaki ilişkinin şiddetini; önündeki işaret ise, değişkenler arasındaki ilişkinin yönünü göstermektedir (Yerdelen Tatoğlu, 2023, s. 4).

### 3. Matematiksel Modelin Ekonometrik Model Haline Getirilmesi

Matematiksel modeller, bağımlı ve bağımsız değişkenler arasında kesin (deterministik) bir ilişki olduğunu varsaymaktadır. Ancak ekonomik ve sosyal olayları etkileyen tüm faktörlerin tam olarak gözlenmesi ve modele dâhil edilmesi mümkün olmadığından, iktisadi değişkenler arasındaki ilişkiler tam anlamıyla deterministik değildir. Bu nedenle, gözlenen değerler teorik ilişkiyi ifade eden doğrudan belirli ölçüde sapabilmektedir (Yerdelen Tatoğlu, 2023, s. 5-6).

Bu sapmalar, ekonometrik modellerde hata terimi (genellikle  $u$  ile gösterilmektedir) aracılığıyla temsil edilmektedir. Böylece yalnızca deterministik ilişkiyi ifade eden matematiksel model, hata teriminin eklenmesiyle stokastik (rastlantısal) yapıya sahip ekonometrik modele dönüşmektedir. Başka bir ifadeyle, matematiksel model  $Y=f(X)$  biçiminde ifade edilirken, ekonometrik model  $Y=f(X, u)$  şeklini almaktadır (Yerdelen Tatoğlu, 2023, s. 6).

### 4. Verilerin Toplanması

Ekonometrik araştırmanın en önemli aşamalarından bir tanesi de değişkenlere ait verilerin toplanmasıdır. Sabit ve eğim parametrelerine ait sayısal değerlerin bulunabilmesi için veriye ihtiyaç vardır. Güvenilir kaynaklardan ve doğru olarak veri toplanması ekonometrik tahminlerin güvenilirliğini büyük ölçüde etkilemektedir. Ekonometrik analizlerde üç çeşit veri türü kullanılmaktadır. Bunlar: zaman serisi verisi, yatay kesit veri ve panel veridir (Yerdelen Tatoğlu, 2023, s. 6-7).

**Zaman Serisi Verisi:** Değişkenlerin değerlerinin gün, ay, mevsim, yıl gibi zaman birimlerine göre değişimini içeren verilere denilmektedir. Yıllık milli gelir, aylık ihracat, günlük hisse senedi getirileri gibi (Yerdelen Tatoğlu, 2023, s. 7).

**Yatay Kesit Veri:** Zamanın belli bir noktasında, farklı birimlerden toplanan verilere denilmektedir. Burada “*birim*”; birey, hane halkı, firma, sektör, ülke gibi ekonometrik birimleri ifade etmek için kullanılmaktadır. 2017 yılı Ocak ayında Türkiye’de illere göre otomobil sayısı, Akdeniz Ülkelerinin her birine 2019 yılı 3. çeyreğinde gelen toplam turist sayısı gibi (Yerdelen Tatoğlu, 2023, s. 7).

**Panel Veri:** Bireyler, ülkeler, firmalar, hane halkları gibi birimlere ait yatay kesit gözlemlerin, belli bir dönemde bir araya getirilmesi olarak tanımlanmaktadır. Panel veri,  $N$  sayıda birim ve her bir birime karşılık gelen  $T$  sayıda gözlemden oluşmaktadır. Türkiye’nin illerinin 1980–2015 yılları arasındaki buğday üretim miktarları, BİST-30’da yer alan hisse senetlerinin 2019 yılı günlük getiri oranları,

1986–2017 döneminde imalat sanayindeki firmalarda çalışan işçi sayıları gibi (Yerdelen Tatoğlu, 2023, s. 7-8).

Uluslararası veri kaynakları: OECD, IMF, Dünya Bankası (WorldBank), Eurostat (Avrupa Birliği İstatistik Birimi), AMECO gibi iken ulusal veri kaynakları: TÜİK (Türkiye İstatistik Kurumu), SBB (Strateji ve Bütçe Başkanlığı), HMB (Hazine ve Maliye Bakanlığı), Ticaret Bakanlığı, TCMB, ITO (İstanbul Ticaret Odası) ve Borsa İstanbul vb. gibidir (Yerdelen Tatoğlu, 2023, s. 8).

## 5. Ekonometrik Modelin Tahmini

Veriler elde edildikten sonra ekonometrik yöntemlerle parametreler tahmin edilebilmektedir (Yerdelen Tatoğlu, 2023, s. 8).

## 6. Hipotez Testleri ve Varsayımların Testleri

Bu aşamada modelin istatistiki ve ekonometrik olarak testleri yapılmaktadır. Ancak tüm temel varsayımları sağlayan modeller güvenilir olarak kabul edilebilir. İstatistiki ve ekonometrik olarak tüm testlerden geçmiş ve varsayımları sağlamış olsa bile bir model iktisadi teori ile uyumlu sonuç vermeyebilir. Bu durumda ya başa dönmekte olası hatalar kontrol edilir ya da bulunan yeni sonuçlar desteklenerek teoriye katkıda bulunulur (Yerdelen Tatoğlu, 2023, s. 8-9).

## 7. Modelin Kullanılması

Hipotez testleri sonucunda seçilen modelin kuramı ya da ön savı doğruladığına karar verilirse, bağımsız değişkenlerin bilinen ya da beklenen değerinden hareketle bağımlı değişkenin gelecekteki değeri kestirilebilmektedir. Yine modelin sonuçları çeşitli para ve maliye politikaları gibi politikaların üretilmesi amacıyla kullanılabilir (Yerdelen Tatoğlu, 2023, s. 9).

## Değişken Çeşitleri

Ekonometrik modellerde yer alan değişkenler modele ve kullanılan veriye göre farklı isimler almaktadır.

- **Bağımlı ve Bağımsız Değişkenler:** Bağımsız değişken, bir iktisadi olayın meydana gelmesini sağlayan değişken iken bağımlı değişken değeri model içinde belirlenen değişkendir (Yerdelen Tatoğlu, 2023, s. 2).
- **İçsel ve Dışsal Değişkenler:** İçsel değişkenler model ya da sistem içerisinde, dışsal değişkenler ise, model ya da sistem dışında belirlenen değişkenlerdir.

- **Nicel ve Nitel Değişkenler:** Sayısal olarak ölçülebilen değişkenler nicel, cinsiyet, medeni durum gibi ölçülemeyen değişkenler ise nitel değişkenlerdir (Yerdelen Tatoğlu, 2023, s. 9).

## Ekonometrik Model Çeşitleri

Hareket edilen iktisat teorisine göre ekonometrik olarak farklı modeller kullanmak gerekebilmektedir. Bunlar aşağıda sıralanmıştır:

- **Değişken Sayısına Göre:** Basit ve çoklu regresyon modelleri. Bir ekonometrik model bir bağımlı ve bir bağımsız değişkenden oluşabilmekte ya da bağımsız değişken sayısı birden fazla olabilmektedir. İlk tanımlanan model basit regresyon modeli, ikincisi ise çoklu regresyon modelidir.
- **Denklem Sayısına Göre:** Tek ve çok denklemlı modeller. Bazen değişkenler arasında karşılıklı ilişkiler söz konusu olabilmekte ve tek denklem iktisadi ilişkileri özetlemekte yetersiz kalmaktadır, bu durumda çok denklemlı modeller kullanılmaktadır.
- **Modelin Fonksiyonel Şekline Göre:** Doğrusal ve doğrusal olmayan modeller. Bağımlı ve bağımsız değişkenler arasındaki ilişkiler doğrusal olabildiği gibi bazen eğrisel de olabilmektedir.
- **Dinamiklik Durumuna Göre:** Statik ve dinamik modeller. Değişkenler arası ilişkiler aynı dönemde gerçekleşebildiği gibi bazen birden fazla döneme de yayılabilmekte ve dinamik modeller söz konusu olabilmektedir.
- **Birim ve Zaman Boyutuna Göre:** Yatay kesit modelleri, belirli bir zamanda farklı birimlere ait verilerle; zaman serisi modelleri, aynı birime ait değişkenlerin zaman içindeki değişimiyle; panel veri modelleri ise hem yatay kesit hem de zaman boyutunu birlikte içeren verilerle kurulmaktadır (Yerdelen Tatoğlu, 2023, s. 10).

## Ana Kütle, Alt Ana Kütle, Örnek Kavramları

İncelenmesi amaçlanan tüm birimlerin oluşturduğu topluluğa “ana kütle”, ana kütlede seçilen az sayıdaki birimin oluşturduğu topluluğa “örnek”, örnek seçme işlemine ise “örnekleme” denilmektedir (Yerdelen Tatoğlu, 2023, s. 10-11).

Ana kütlede doğrudan gözlenemediği durumlarda ana kütle regresyon fonksiyonu yerine örnek regresyon fonksiyonundan hareket edilmektedir. Ekonometrik model tahmininin amacı, örnek regresyon fonksiyonundan yararlanarak ana kütle regresyon fonksiyonunu tahmin etmektir (Yerdelen Tatoğlu, 2023, s. 11).

## Doğrusallık ve Doğrusal Olmayan Modeller

Bağımlı ve bağımsız değişkenler arasındaki ilişki grafik üzerinde bir doğru ile gösteriliyorsa doğrusal, bir eğri ile gösteriliyorsa doğrusal olmayan (eğrisel) model söz konusudur. Bu nedenle bir denklem, değişkenlerine göre doğrusal veya doğrusal olmayan yapıda olabilmektedir. Değişkenler birbiriyle çarpılmıyor, bölünmüyor veya üs olarak yer almıyorsa denklem değişkenlerine göre doğrusaldır. Bu durumda bağımlı değişkenin her bir bağımsız değişkene göre birinci türevi sabit olmaktadır (Yerdelen Tatoğlu, 2023, s. 12-13).

Bunun yanı sıra denklemler, parametrelerin doğrusallığı bakımından da değerlendirilmektedir. Parametreler birbiriyle çarpılmıyor, bölünmüyor veya üs olarak yer almıyorsa model parametrelerine göre doğrusaldır (Yerdelen Tatoğlu, 2023, s. 13).

Ekonometride model sınıflandırmasında esas alınan ölçüt değişkenlerin değil, parametrelerin doğrusallığıdır. Bu nedenle parametreleri bakımından doğrusal olan modeller, değişkenler doğrusal olmasa bile doğrusal model olarak kabul edilmektedir (Yerdelen Tatoğlu, 2023, s. 13-14).

Değişkenlerine göre doğrusal olmayan modeller arasında en yaygın kullanılanlar logaritmik, ters, parabolik ve kübik modellerdir.

### Logaritmik Modeller

Logaritmik modeller kendi arasında ikiye ayrılmaktadır:

- Yarı logaritmik modeller
- Tam logaritmik modeller (Yerdelen Tatoğlu, 2023, s. 14)

### Yarı Logaritmik Modeller

Doğrusal logaritmik modeller ve logaritmik doğrusal modeller (Yerdelen Tatoğlu, 2023, s. 13).

### Doğrusal Logaritmik Modeller

Doğrusal logaritmik modellerde bağımlı değişken aritmetik, bağımsız değişken(ler) ise geometrik seri özelliği göstermektedir. Bu nedenle bağımlı değişken mutlak, bağımsız değişken(ler) ise nispi olarak değerlendirilmektedir. Tek bağımsız değişkenli doğrusal logaritmik model aşağıdaki gibi gösterilmektedir (Yerdelen Tatoğlu, 2023, s. 14):

$$Y = \alpha + \beta \log X + u$$

$$Y = \alpha + \beta \ln X + u$$

## Logaritmik Doğrusal Modeller

Logaritmik doğrusal modellerde bağımlı değişken geometrik, bağımsız değişken(ler) ise aritmetik seri özelliği göstermektedir. Bu nedenle bağımlı değişken nispi, bağımsız değişken(ler) ise mutlak olarak değerlendirilmektedir. Tek bağımsız değişkenli logaritmik doğrusal model aşağıdaki gibi gösterilmektedir (Yerdelen Tatoğlu, 2023, s. 15).

$$\log Y = \alpha + \beta X + u$$

$$\ln Y = \alpha + \beta X + u \quad \text{ya da} \quad Y = e^{\alpha + \beta X + u}$$

## Tam Logaritmik Modeller

Tam logaritmik modellerde hem bağımlı hem de bağımsız değişkenler geometrik seri özelliği göstermektedir. Bu nedenle her iki değişken de nispi olarak değerlendirilmektedir. Tek bağımsız değişkenli tam logaritmik model aşağıdaki gibi gösterilmektedir:

$$\log Y = \log \alpha + \beta \log X + u \quad \text{ya da} \quad Y = \alpha X^\beta + u$$

$$\ln Y = \ln \alpha + \beta \ln X + u$$

Tam logaritmik modellerin en önemli özelliği  $\beta$ 'nın esnekliği vermesidir (Yerdelen Tatoğlu, 2023, s. 15).

## Ters Modeller

Bağımsız değişkenin tersi ile bağımlı değişken arasındaki ilişkiyi açıklayan modellerde *ters modeller* denilmektedir. Ters modeller aşağıdaki gibi gösterilebilmektedir:

$$Y = \alpha + \beta(1/X) + u$$

$\beta$  katsayısının işaretine göre modelin yorumu değişmektedir.  $\beta < 0$  olduğunda X sonsuza giderken Y sabit parametreye yakınsamaktadır.  $\beta > 0$  olduğunda ise X arttıkça Y de artmakta ve sonsuza yaklaşmaktadır (Yerdelen Tatoğlu, 2023, s. 15-16).

## Parabolik ve Kübik Modeller

Bağımlı ve bağımsız değişkenler arasındaki doğrusal olmadığı, fonksiyonun hem artış hem de azalış gösterdiği durumlarda ikinci veya üçüncü dereceden modeller kullanılmaktadır.

Fonksiyonun tek bir minimum ya da maksimum noktası varsa ikinci dereceden (parabolik veya kuadratik) modeller kullanılmaktadır. Bu modeller aşağıdaki gibi gösterilmektedir:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + u$$

Fonksiyonun ikinci türevi negatifse maksimum, pozitifse minimum nokta vardır. Maksimum ya da minimum noktanın değeri, birinci türevin sıfır olduğu noktaya göre ya da  $-\beta_1/2\beta_2$  formülüyle hesaplanmaktadır.

Üçüncü dereceden ya da kübik modellerin bir minimum bir de maksimum noktası bulunmaktadır ve aşağıdaki gibi gösterilebilmektedir:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 + u$$

$\beta$ 'ların işaretine bağlı olarak fonksiyonun şekli değişmektedir (Yerdelen Tatoğlu, 2023, s. 16).

## DOĞRUSAL VE DOĞRUSAL OLMAYAN REGRESYON MODELLERİ VE TAHMİN YÖNTEMLERİ

Regresyon; bir bağımlı değişkenin ortalama değerinin, bir ya da daha fazla bağımsız değişken yardımıyla tahmin edilmesi olarak tanımlanabilmektedir (Yerdelen Tatoğlu, 2023, s. 1-45).

### Doğrusal Regresyon Modelleri ve Tahmin Yöntemleri

Yatay kesit verilerle kurulan yatay kesit regresyon modeli:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i$$

Zaman serisi verileriyle kurulan zaman serisi regresyon modeli:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + u_t$$

Bu eşitliklerde  $Y$ : bağımlı değişken,  $X$ : bağımsız değişkenler,  $\beta_0$  sabit ve  $\beta_1, \beta_2, \dots, \beta_k$  eğim parametreleri,  $n$ : gözlem sayısı,  $k$ : parametre sayısı,  $i$  ve  $t$  alt indisleri ise sırasıyla birim ve zamanı ifade etmektedir (Yerdelen Tatoğlu, 2023, s. 45-46).

Regresyon modellerinin matris notasyonunda gösterimi aşağıdaki gibidir:

$$Y = X\beta + u$$

Burada (Yerdelen Tatoğlu, 2023, s. 46):

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & X_{11} & \dots & X_{k-1,1} \\ 1 & X_{12} & \dots & X_{k-1,2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & \dots & X_{k-1,n} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} \quad u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

$Y$ :  $n \times 1$  boyutunda bağımlı değişkene ait gözlemleri gösteren bir sütun vektör,

$X$ :  $n \times k$  (basit regresyonda  $n \times 2$ ) boyutunda bağımsız değişkenlere ait gözlemleri ve 1'ler vektörünü içeren bir matris (ilk sütundaki 1'ler sabit terimi gösterir),

$\beta$ :  $k \times 1$  boyutunda bilinmeyen parametreleri gösteren bir sütun vektör,

$u$ :  $n \times 1$  boyutunda hata terimini gösteren bir sütun vektördür.

## En Küçük Kareler Yöntemi

### Parametrelerin Tahmini

Ana külte regresyon fonksiyonu doğrudan gözlenemediğinden aşağıdaki örnek regresyon fonksiyonundan hareket edilmektedir:

$$Y = X\beta + \hat{u} \\ = \hat{Y} + \hat{u}$$

Burada  $\hat{Y}$ ,  $Y$ 'nin tahmin edilmiş değeridir ve,

$$\hat{u} = Y - \hat{Y}$$

$\hat{u}$  (kalıntılar), gözlemlenen ve tahmin edilen  $Y$ 'lerin farkıdır. Örnek regresyon fonksiyonunu, ana külte regresyon fonksiyonuna en yakın olacak şekilde tahmin etmek için, gözlemlenen ve tahmin edilen  $Y$ 'lerin ( $\hat{u}$ ) farkının toplamını minimum yapmak gerekmektedir (Yerdelen Tatoğlu, 2023, s. 47).

$$\sum \hat{u} = \sum (Y - \hat{Y})$$

Bu yaklaşım örnek regresyon fonksiyonuna uzak ya da yakın tüm kalıntılara aynı ağırlığı verdiği için tercih edilmemektedir, ayrıca kalıntıların toplamı sıfırdır. Bunun yerine karelerden hareket etmek daha mantıklı olmakta; kalıntı kareleri toplamı minimum edilmeye çalışılmaktadır.

$$\hat{u}'\hat{u} = \sum \hat{u}^2 = \sum (Y - \hat{Y})^2 \rightarrow \min$$

$$\begin{aligned} \hat{u}'\hat{u} &= (Y - \hat{Y})' (Y - \hat{Y}) \\ &= (Y - X\beta)' (Y - X\beta) \\ &= Y'Y - \beta' X'Y - Y'X\beta + \beta' X'X\beta \\ &= Y'Y - 2\beta' X'Y + \beta' X'X\beta \end{aligned}$$

$\hat{u}'\hat{u}$ 'nin  $\beta$ 'ya göre birinci türevi alınıp sıfıra eşitlenir ve  $\beta$  için çözüldüğünde  $\beta$ 'ların tahmini bulunmaktadır (Yerdelen Tatoğlu, 2023, s. 47-48):

$$\hat{\beta} = (X'X)^{-1} (X'Y)$$

## Bağımlı Değişkendeki Değişkenliğin Kaynakları

$Y'Y - n\bar{Y}^2 \left( = \sum (Y_i - \bar{Y})^2 \right) = \text{TSS}$ : Y'deki toplam değişkenlik ya da toplam kareler toplamıdır.

$\beta' X'X \beta - n\bar{Y}^2 \left( = \sum (\hat{Y}_i - \bar{Y})^2 \right) = \text{RSS}$ : Y'deki beklenen değişkenlik, regresyon kareler toplamı ya da regresyonla açıklanabilen kareler toplamıdır.

$u'u \left( = \sum (Y_i - \hat{Y}_i)^2 \right) = \text{ESS}$ : Y'deki artık değişkenlikten kaynaklanan kısım, regresyonla açıklanamayan kareler toplamı ya da kalıntı kareleri toplamını ifade etmektedir (Yerdelen Tatoğlu, 2023, s. 48-49).

Bu tanımlamalarla:

$$\underbrace{Y'Y - n\bar{Y}^2}_{\text{TSS}} = \underbrace{\beta' X'X \beta - n\bar{Y}^2}_{\text{RSS}} + \underbrace{u'u}_{\text{ESS}}$$

## Belirginlik Katsayısı ve Düzeltilmiş Belirginlik Katsayısı

Örneklem doğrusunun verilere ne kadar yakın olduğu belirginlik katsayısı ( $R^2$ ) hesaplanarak bulunmakta ve kalıntılar ne kadar küçükse,  $R^2$  o kadar büyük olmaktadır (Yerdelen Tatoğlu, 2023, s. 49):

$$R^2 = \frac{\text{RSS}}{\text{TSS}} = \frac{\hat{\beta}' X'X \hat{\beta} - n\bar{Y}^2}{Y'Y - n\bar{Y}^2} = 1 - \frac{\text{ESS}}{\text{TSS}} = 1 - \frac{\hat{u}'\hat{u}}{Y'Y - n\bar{Y}^2}$$

Böylece  $R^2$ , Y'deki toplam değişkenlik içinde regresyon modelince açıklanan oranı, yüzdeyi ölçmekte ve bağımsız değişkenin, bağımlı değişkendeki değişkenliğin yüzde kaçını açıkladığını vermektedir. Kareli bir ifade olduğundan negatif değer almamaktadır, sınırları 0 ile 1 aralığındadır ( $0 \leq R^2 \leq 1$ ).  $R^2=0$  ve  $R^2=1$  deterministik ilişkileri yansıtmaktadır (Yerdelen Tatoğlu, 2023, s. 49-50).

Çoklu regresyon modellerinde özellikle yuvalanmış olan modeller kıyaslanırken  $R^2$ 'nin serbestlik derecesine göre düzeltilmiş hali kullanılmakta ve bu  $R^2$ 'ye *düzeltilmiş  $R^2$*  ( $\bar{R}^2$ ) denilmektedir.  $\bar{R}^2$  aşağıdaki gibi hesaplanmaktadır (Yerdelen Tatoğlu, 2023, s. 50):

$$\bar{R}^2 = 1 - \frac{\hat{u}'\hat{u}/(n-k)}{Y'Y - n\bar{Y}^2/(n-1)} = 1 - \frac{n-1}{n-k} (1 - R^2)$$

## Varyans ve Standart Hatalar

İstatistikte bir değişkenin güvenilirliği standart hatası ile ölçülmektedir. *Standart hata*, tahminlerinin örneklem dağılımının standart sapmasıdır; ortalama etrafındaki dağılım ölçüsüdür. Kalıntı varyansı (Yerdelen Tatoğlu, 2023, s. 51):

$$\hat{\sigma}_u^2 = \frac{\hat{u}'\hat{u}}{n - k}$$

Burada,  $\hat{u}'\hat{u}$  : kalıntıların kareleri toplamı ve  $n-k$ : serbestlik derecesidir. *Serbestlik derecesi*, toplam  $n$  gözlem içinde bağımsız olanların sayısıdır.

Kalıntının varyans kovaryans matrisi hata teriminin varyansının alt ana kütleler itibarıyla birbirine eşit olması (homoskedasite) ve hata terimlerinin gözlemler itibarıyla birbirinden bağımsız olması (otokorelasyon olmaması) varsayımlarıyla aşağıdaki gibidir:

$$E(uu') = \sigma_u^2 I = \sigma_u^2 \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = \begin{bmatrix} \sigma_u^2 & 0 & \dots & 0 \\ 0 & \sigma_u^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_u^2 \end{bmatrix}$$

Burada köşegen elemanlar varyansı göstermekte olup birbirine eşit olması heteroskedasite olmadığını ifade etmektedir. Köşegen dışı elemanlar ise kovaryansı ifade etmektedir ve sıfır olması otokorelasyon olmadığını göstermektedir (Yerdelen Tatoğlu, 2023, s. 51). Heteroskedasite ve otokorelasyon yokken parametre varyans kovaryans matrisi:

$$Var(\hat{\beta}) = E\left[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'\right] = \sigma_u^2 (X'X)^{-1}$$

Bu matriste köşegen elemanlar parametre tahminlerinin varyanslarını, varyansların karekökü standart hataları ve köşegen dışı elemanlar ise kovaryansları vermektedir. Örneğin  $\beta_1$ 'in standart hatası:

$$sh(\hat{\beta}_1) = \sqrt{var(\hat{\beta}_1)}$$

$\hat{\beta}$  'nin  $\beta$  etrafındaki dağılımı ortalama hata karesi (MSE) ile ölçülmektedir (Yerdelen Tatoğlu, 2023, s. 51-53):

$$MSE(\hat{\beta}) = E(\hat{\beta} - \beta)^2$$

## Parametrelerle İlgili Doğrusal Hipotezlerin Testleri (t ve F Testleri)

Doğrusal bir regresyon modelinde  $\beta$ 'lar ile ilgili farklı hipotezler kurulabilmektedir. En yaygın hipotezler aşağıda özetlenmiştir (Yerdelen Tatoğlu, 2023, s. 53-55):

| Hipotez                      | Amaç                                               | Test    |
|------------------------------|----------------------------------------------------|---------|
| $H_0: \beta_i = 0$           | Parametrenin anlamlılığını sınamak                 | t testi |
| $H_0: \beta_i = \beta_{i0}$  | Parametrenin belirli bir değere eşitliğini sınamak | t testi |
| $H_0: \beta_i = \beta_j$     | İki parametrenin eşitliğini sınamak                | t testi |
| $H_0: \beta_i + \beta_j = c$ | Parametrelerin toplamına ilişkin hipotezi sınamak  | t testi |
| Birden fazla kısıt           | Parametrelerin ortak anlamlılığını sınamak         | F testi |

t istatistiği,

$$t = \frac{\hat{\beta}_1 - \beta_1}{sh(\hat{\beta}_1)}$$

şeklinde iken eğer sıfıra eşitlik sınıyanıyorsa aşağıdaki hali almaktadır:

$$t = \frac{\hat{\beta}_1}{sh(\hat{\beta}_1)}$$

Hesaplanan t istatistiği, seçilen anlamlılık düzeyi ( $\alpha$ ) ve serbestlik derecesine göre elde edilen t tablo değeri ile karşılaştırılmaktadır. Genellikle %5 anlamlılık düzeyi ( $\alpha = 0.05$ ) kullanılmaktadır. t istatistiği, t tablo değerinden büyükse  $H_0$  hipotezi reddedilmektedir. Kurulan hipoteze göre bu sonuç; parametrenin sıfırdan, belirli bir değerden veya diğer parametrelerden istatistiksel olarak farklı olduğunu göstermektedir.

Parametrenin yönü hakkında önceden kuramsal bir beklenti bulunmuyorsa çift taraflı hipotez testi, ilişkinin yönü kuramsal olarak biliniyorsa tek taraflı hipotez testi kullanılabilir (Yerdelen Tatoğlu, 2023, s. 56).

Y'deki değişkenliğin X'lerdeki değişkenlikle doğrusal bir ilişki gösterip göstermediği iki değişkenlik tipinin (RSS ve ESS) birbirine oranlanmasıyla elde edilen istatistik yardımıyla test edilmektedir. Bu istatistik F istatistiğidir ve söz konusu test F-testi olarak anılmaktadır.

F testinin hipotezleri sabit parametre ( $\beta_0$ ) dışındaki tüm parametreleri birlikte test etmek için kullanılmaktadır:

$$H_0 : \beta_2 = \beta_3 = \dots = \beta_k = 0$$

$$H_1 : \text{En az biri sıfırdan farklı}$$

F test istatistiği:

$$F = \frac{RSS/(k-1)}{ESS/(n-k)} = \frac{R^2/(k-1)}{(1-R^2)/(n-k)}$$

Basit regresyonda F test istatistiği t test istatistiğinin karesidir:

$$F_{(\alpha;1,n-2)} = (t_{n-2})^2$$

Hesaplanan F değeri F tablo değeri ile karşılaştırılmaktadır. Hesaplanan değer tablo değerinden büyükse  $H_0$  hipotezi reddedilmekte ve modelin anlamlı olduğuna karar verilmektedir (Yerdelen Tatoğlu, 2023, s. 58).

### Parametreler İçin Güven Aralıkları

Nokta tahminleri, örnekten örneğe değişebileceğinden tek başına yeterli bilgi sağlamamaktadır. Bu nedenle, parametrelerin ya da üç standart hata uzaklığını içeren bir aralık oluşturabilmekte ve bu aralık %95 olasılıkla ana kütlelenin gerçek katsayısını içermektedir.  $\beta_1$  parametresi için güven aralığı:

$$P\left[\hat{\beta}_1 - t_{\alpha/2}sh(\hat{\beta}_1) \leq \beta_1 \leq \hat{\beta}_1 + t_{\alpha/2}sh(\hat{\beta}_1)\right] = 1 - \alpha$$

Güven aralığının genişliği, tahminin standart hatasına bağlıdır. Standart hata ne kadar büyükse, güven aralığı o kadar geniş ve ana kütle katsayısının tahmini de o kadar belirsiz olmaktadır. (Yerdelen Tatoğlu, 2023, s. 59).

### Kalıntı Varyansı İçin Güven Aralığı

Kalıntı varyansı ( $\hat{\sigma}_u^2$ ) için güven aralığı aşağıdaki gibi hesaplanmaktadır:

$$P\left[(n-2)\frac{\hat{\sigma}_u^2}{\chi_{\alpha/2}^2} \leq \sigma_u^2 \leq (n-2)\frac{\hat{\sigma}_u^2}{\chi_{1-\alpha/2}^2}\right] = 1 - \alpha$$

$$P\left[\frac{u'u}{\chi_{\alpha/2}^2} \leq \sigma_u^2 \leq \frac{u'u}{\chi_{1-\alpha/2}^2}\right] = 1 - \alpha$$

(n-2) serbestlik dereceli  $\chi^2$  dağılımına uygunluk göstermektedir (Yerdelen Tatoğlu, 2023, s. 59).

## Varyans Analiz (ANOVA) Tablosu

Varyans analiz tablosu bağımlı değişkendeki değişkenliğin kaynağını toplu olarak görmek için yapılan bir tablodur.

| Değişimin Kaynağı                       | Kareler Toplamı                           | Serbestlik Derecesi | Ortalama Kareler Toplamı                          |
|-----------------------------------------|-------------------------------------------|---------------------|---------------------------------------------------|
| Regresyon ile Açıklanabilen Değişkenlik | $\hat{\beta}'X'X\hat{\beta} - n\bar{Y}^2$ | k-1                 | $\hat{\beta}'X'X\hat{\beta} - n\bar{Y}^2 / (k-1)$ |
| Regresyon ile Açıklanamayan Değişkenlik | $\hat{u}'\hat{u}$                         | n-k                 | $\hat{u}'\hat{u} / (n-k) = \hat{\sigma}_u^2$      |
| Toplam Değişkenlik                      | $Y'Y - n\bar{Y}^2$                        | n-1                 |                                                   |

ANOVA tablosu, toplam değişkenliği regresyon modeli tarafından açıklanan ve açıklanamayan bileşenlere ayırmaktadır. Bu tablodan F testi ve belirlilik katsayısı ( $R^2$ ) hesaplanarak modelin genel anlamlılığı ve açıklama gücü değerlendirilmektedir (Yerdelen Tatoğlu, 2023, s. 59-60). (Yerdelen Tatoğlu, 2023, s. 59-60).

## Korelasyon Katsayısı

Değişkenler arasındaki ikili ilişkinin yönü ve gücü korelasyon katsayısı yardımıyla ölçülmektedir. Korelasyon katsayısı -1 ile +1 arasında değer almakta olup, katsayının işareti ilişkinin yönünü, mutlak değeri ise ilişkinin gücünü göstermektedir. Korelasyon katsayısının +1 olması tam pozitif doğrusal ilişkiyi, -1 olması tam negatif doğrusal ilişkiyi, 0 olması ise doğrusal bir ilişkinin bulunmadığını ifade etmektedir. Ancak korelasyon katsayısı yalnızca değişkenler arasındaki doğrusal ilişkinin derecesini göstermekte olup, nedensellik ilişkisini ifade etmemektedir (Yerdelen Tatoğlu, 2023, s. 60).

Y ile X arasındaki korelasyon katsayısı ( $r_{YX}$ ):

$$r_{YX} = \frac{\sum yx}{\sqrt{\sum y^2 \sum x^2}}$$

Burada  $y = Y_i - \bar{Y}$  ve  $x = X_i - \bar{X}$  'dır.

Basit regresyon modellerinde  $R^2$ 'nin karekökü korelasyon katsayısıdır:

$$r = \pm\sqrt{R^2}$$

Korelasyon katsayısının anlamlılığını sınamak için aşağıdaki hipotezler kurulmaktadır:

$$H_0: r = 0$$

$$H_1: r \neq 0$$

Test istatistiği:

$$t_r = \frac{r}{sh(r)} = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

(n-2) serbestlik dereceli t-tablo değerinden büyükse  $H_0$  hipotezi reddedilmekte ve korelasyon katsayısının anlamlı olduğuna karar verilmektedir (Yerdelen Tatoğlu, 2023, s. 61-62).

Kısmi korelasyon katsayısı üçüncü değişkenin etkisi sabitken iki değişken arasındaki korelasyon katsayısı olarak tanımlanmaktadır. Örneğin  $X_3$ 'ün etkisi sabitken  $X_1$  ve  $X_2$  arasındaki korelasyon katsayısı (Yerdelen Tatoğlu, 2023, s. 62):

$$r_{X_1X_2.X_3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1-r_{13}^2}\sqrt{1-r_{23}^2}}$$

Sadece değişkenlerden birisi için etki sabit tutularak *yarı-kısmi korelasyon katsayısı* elde edilmektedir. Örneğin  $r_{X_1X_2.X_3}$  için (Yerdelen Tatoğlu, 2023, s. 63):

$$r_{X_1(X_2.X_3)} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1-r_{23}^2}}$$

### En Küçük Karelerin Ardındaki Varsayımlar

**V1.** Regresyon modeli parametrelerine göre doğrusaldır.

**V2.** Örnekler ana kütleden tesadüfi örnekleme yoluyla çekilmektedir.

**V3.** X matrisi stokastik (rassal) olmayan bir matristir.

**V4.** Hata teriminin (u) koşulsuz ve koşullu beklenen değeri sıfırdır.

**V5.**  $\text{rank}(E(X'X))=k$ 'dir. Bu varsayım  $X'X$  matrisinin rankının tam olduğunu yani modelde tam çoklu doğrusal bağlantı olmadığını ifade etmektedir.

**V6.**  $E(uu') = \sigma_u^2 I$  'dir.

- X değerlerine karşılık gelen Y ana kütlelerinin varyansları aynıdır; *heteroskedasite* yoktur.
- Köşegen dışı elemanlar sıfıra eşittir, yani hata terimleri arasında ardışık bağımlılık yoktur; *otokorelasyon* yoktur

V7. Gözlem sayısı  $n$ , tahmin edilecek katsayı sayısından fazla olmalıdır ( $n > k$ ).

V8. Regresyon modeli doğru kurulmuş olmalı; model kurma hatası bulunmamalıdır.

V9. Hata terimleri normal dağılmaktadır.

Bu varsayımlar, en küçük kareler tahmincilerinin sapmasızlık, en küçük varyans gibi özelliklerini gerçekleştirmeye yardımcı olmaktadır (Yerdelen Tatoğlu, 2023, s. 62-65).

### En Küçük Kareler Tahmincilerinin Özellikleri

#### Küçük Örnek Özellikleri

EKK tahmincileri doğrusal, en iyi, sapmasız tahmincilerdir (DESTE ya da BLUE).

- 1) **Doğrusallık:** Tahminciler rassal bağımlı değişkenin doğrusal bir fonksiyonudur.
- 2) **Sapmasızlık:** En küçük kareler tahmincileri sapmasızdır, bir başka ifade ile tahmincilerin ortalaması yani beklenen değeri, gerçek değere eşittir.
- 3) **Etkinlik:** Etkin tahminci bütün doğrusal ve sapmasız tahminciler içinde en düşük varyanslı olanıdır.

V1-V6 varsayımları altında  $\hat{\beta}$  örnek birimlerinin doğrusal bir fonksiyonudur,  $\beta$ 'nin sapmasız bir tahmincisidir ve minimum varyansa sahiptir. Bir başka ifade ile  $\hat{\beta}$  'lar ana kütle parametresi  $\beta$ 'nin doğrusal en iyi sapmasız tahmincisidir (Yerdelen Tatoğlu, 2023, s. 65-66).

#### Büyük Örnek Özellikleri

- 1) **Tutarlılık:** Örnek birim/zaman sayısı arttıkça  $\hat{\beta}$  ,  $\beta$ 'ya yaklaşmaktadır.
- 2) **Asimptotik Normallik:** Hata terimleri normal dağılmasa bile, hangi dağılıma sahip olursa olsun  $n$  (ya da  $t$ ) sonsuza giderken dağılımı normal dağılıma yaklaşmaktadır.  $u$  normal dağılırsa  $\beta$ 'lar da normal dağılacaktır.
- 3) **Asimptotik Etkinlik:**  $\beta$ 'nin sonlu ortalaması ve sonlu varyansı ile asimptotik bir dağılımı varsa, tutarlıysa ve tutarlı tahmincileri arasında asimptotik varyansı en küçük olan tahmincisi ise  $\hat{\beta}$  ,  $\beta$ 'nin asimptotik etkin bir tahmincisidir (Yerdelen Tatoğlu, 2023, s. 67).

## En Çok Olabilirlik Tahmincisi

Aşağıdaki model ele alındığında:

$$Y = X\beta + u \quad u \sim N(0, \sigma^2)$$

$u$ 'nun normal dağıldığı ( $u \sim N(0, \sigma^2)$ ) varsayımı altında log-olabilirlik fonksiyonu:

$$\begin{aligned} l = \ln f(y|X) &= \ln f(u) = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} u'u \\ &= -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} (Y - X\beta)'(Y - X\beta) \end{aligned}$$

log-olabilirlik fonksiyonunun bilinmeyen parametrelere göre birinci türevleri:

$$\begin{aligned} \frac{\partial l}{\partial \beta} &= -\frac{1}{\sigma^2} (-X'Y + X'X\beta) \\ \frac{\partial l}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} (Y - X\beta)'(Y - X\beta) \end{aligned}$$

Bu türevler sıfıra eşitlenmekte ve çözülmektedir. Bu aşamada çeşitli iteratif süreçler uygulanmaktadır. Buradan aşağıdaki tahminciler elde edilmektedir (Yerdelen Tatoğlu, 2023, s. 86-87):

$$\hat{\beta} = (X'X)^{-1} (X'Y)$$

$$\hat{\sigma}_u^2 = \frac{(Y - X\hat{\beta})'(Y - X\hat{\beta})}{n} = \frac{\hat{u}'\hat{u}}{n}$$

En küçük kareler yönteminde parametre tahmini için normal dağılım varsayımı gerekli değildir. Ancak hipotez testleri ve güven aralıklarının oluşturulabilmesi için hata terimlerinin normal dağıldığı varsayılır. En çok olabilirlik yönteminde ise dağılım varsayımı modelin temelini oluşturur ve log-olabilirlik fonksiyonu seçilen dağılıma göre tanımlanır. Bu nedenle normal dağılım varsayımı altında EKK ile benzer sonuçlar verirken, normal dağılımın sağlanmadığı durumlarda uygun dağılım seçilerek de kullanılabilir (Yerdelen Tatoğlu, 2023, s. 87). (Yerdelen Tatoğlu, 2023, s. 87).

## Genelleştirilmiş En Küçük Kareler Tahmincisi

Genelleştirilmiş en küçük kareler tahmincisi:

$$\hat{\beta} = (X'\Omega^{-1}X)^{-1} (X'\Omega^{-1}Y)$$

Otokorelasyon ve heteroskedasite yoksa Genelleştirilmiş En Küçük Kareler (GEKK) tahmincisi En Küçük Kareler (EKK) tahmincisine eşittir (Yerdelen Tatoğlu, 2023, s. 90-91):

$$\Omega = \sigma_u^2 I \rightarrow \hat{\beta}_{GEKK} = \hat{\beta}_{EKK}$$

EKK'in yanında,  $\Omega$  matrisinin farklı yapıları altında Ağırlıklı (Tartılı) En Küçük Kareler, Genelleştirilmiş Farklar ve benzeri tahmin yöntemleri, Genelleştirilmiş En Küçük Kareler (GEKK) tahmincisinin özel halleri olarak ortaya çıkmaktadır (Yerdelen Tatoğlu, 2023, s. 91).

### Araç Değişkenler Tahmincisi

Dışsalılık varsayımı bozulduğunda tutarlı tahminler elde etmenin yolu içsel olan değişken yerine araç değişken kullanmaktır. Kullanılacak araç değişken Z iken araç değişkenin iki koşulu sağlaması beklenmektedir:

- Z yerine geçeceği değişkenlerle korelasyonlu olmalıdır.
- Z hata terimi ile korelasyonsuz olmalıdır.

Bu koşullar sağlandığında parametreler genelleştirilmiş en küçük kareler yöntemi ile aşağıdaki gibi tahmin edilmektedir:

$$\hat{\beta}_{GEKK} = \hat{\beta}_{IV} = \left( X'Z(Z'Z)^{-1}Z'X \right)^{-1} X'Z(Z'Z)^{-1}Z'Y$$

$\beta$ 'nın varyans-kovaryans matrisi ise aşağıdaki gibi olmaktadır (Yerdelen Tatoğlu, 2023, s. 91-92):

$$\text{var}(\hat{\beta}_{IV}) = \sigma^2 \left( X'Z(Z'Z)^{-1}Z'X \right)^{-1}$$

### Doğrusal Olmayan Regresyon Modelleri

Parametrelerine göre doğrusal olmayan modeller, doğrusal olmayan en küçük kareler, Gauss-Newton ve Steepest-Descent gibi yöntemlerle tahmin edilmektedir (Yerdelen Tatoğlu, 2023, s. 92-93).

## TEMEL VARSAYIMLAR, TESTLERİ VE SAPMALAR HALİNDE DÜZELTME YÖNTEMLERİ

### Çoklu Doğrusal Bağlantı

Çoklu regresyon modellerinde bağımsız değişkenlerin arasında ilişki olması çoklu doğrusal bağlantı olarak adlandırılmaktadır (Yerdelen Tatoğlu, 2023, s. 111).

## **Çoklu Doğrusal Bağlantının Nedenleri**

- Bağımsız değişkenler aynı trende sahip olabilmektedir.
- Veri seti yetersiz olabilmektedir.
- Modelde bağımsız değişkenlerin önceki dönem değerleri de bağımsız değişkenler olarak bulunuyorsa yani dağıtılmış gecikmeli modellerle çalışılıyorsa çoklu doğrusal bağlantı görülebilmektedir (Yerdelen Tatoğlu, 2023, s. 111).

## **Çoklu Doğrusal Bağlantının Sonuçları**

- Bağımsız değişkenler arasında tam korelasyon varsa tam çoklu doğrusal bağlantı vardır. Bu durumda parametreler tahmin edilemez.
- Bağımsız değişkenler arasında hiç korelasyon yoksa, çoklu doğrusal bağlantı yoktur. Bu durumda tahminler hiç etkilenmez.
- Bağımsız değişkenler arasında düşük düzeyde çoklu doğrusal bağlantı varsa, bu durumun modele olumsuz etkisi olmayabilir.
- Orta ya da yüksek düzeyde çoklu doğrusal bağlantı varsa ilk bakışta bile çoklu doğrusal bağlantının önsel göstergeleri görülebilir.

Modelde çoklu doğrusal bağlantı varsa, parametreler sapmalı olmaktadır. Kalıntı varyansı büyümekte ve bu nedenle tahmincilerin varyansları gerçek değerlerinden büyük tahmin edilmektedir. Böylece, aralık tahminleri geniş ve t test istatistiği gerçek değerlerinden küçük tahmin edilmektedir.  $R^2$  ve F istatistiği de büyümektedir (Yerdelen Tatoğlu, 2023, s. 112).

## **Çoklu Doğrusal Bağlantının Belirlenmesi**

### **Önsel Göstergeler**

$R^2$ 'nin büyük, F testinin anlamlı fakat t testlerinin anlamsız olması çoklu doğrusal bağlantının önsel bir göstergesidir (Yerdelen Tatoğlu, 2023, s. 112).

### **Bağımsız Değişkenler Arasındaki Korelasyonlar**

Bağımsız değişkenler arasında yüksek korelasyon olması çoklu doğrusal bağlantı hakkında fikir vermektedir. Genelde pozitif ya da negatif yönde %80'in üzerindeki korelasyonlar yüksek, %50'nin altındaki korelasyonlar düşük ve %50-%80 aralığındaki korelasyonlar orta olarak değerlendirilmektedir (Yerdelen Tatoğlu, 2023, s. 113).

## Varyans Büyütme Faktörü (VIF)

Bu faktör ile parametre tahminlerinin varyanslarının çoklu doğrusal bağlantı nedeniyle gerçek değerinden ne kadar uzaklaştığı belirlenmektedir ve aşağıdaki gibi hesaplanmaktadır:

$$VIF_i = \frac{1}{1 - R_i^2} \quad i = 1, 2, \dots, k$$

Burada  $R_i^2$ , incelenen modelde bağımlı değişkenlerin tek tek bağımlı değişken olduğu, diğer bağımsız değişkenlerin bağımsız değişken olduğu modellere ait  $R^2$  değeridir. VIF 5'den küçükse çoklu doğrusal bağlantı yoktur. 5-10 arasındaysa orta şiddette, 10'dan büyükse güçlü çoklu doğrusal bağlantı vardır.

1/VIF oranı tolerans sayısı olarak adlandırılmakta ve değeri sıfıra yaklaştıkça çoklu doğrusal bağlantının şiddeti artmaktadır (Yerdelen Tatoğlu, 2023, s. 115).

## Diğer Kriterler ve Testler

VIF kriteri çoklu doğrusal bağlantının belirlenmesinde en çok kullanılan kriterdir. Bunun yanında F testi, Klein kriteri, Theil-m kriteri gibi çeşitli kriter ve testler bulunmaktadır (Yerdelen Tatoğlu, 2023, s. 116-117).

## Çoklu Doğrusal Bağlantının Düzeltilmesi

Çoklu doğrusal bağlantıyı düzeltmek için aşağıdaki yöntemler kullanılabilir (Yerdelen Tatoğlu, 2023, s. 117-118):

- Örnek birim sayısının eğer mümkünse artırılması ilk başvuru yöntemidir.
- Panel veri kullanımı hem gözlem sayısını arttırdığı için hem de birim ve zaman boyutundan aynı anda bilgi alındığı için çoklu doğrusal bağlantıyı azaltabilmektedir.
- Çoklu doğrusal bağlantıya neden olan değişkenin modelden çıkarılması.
- Birinci farklar ya da logaritmik dönüşüm gibi değişken dönüşümleri yapılması.
- Çoklu doğrusal bağlantının sebebi gecikmeli bağımsız değişkenler ise bu modellerin tahmini için önerilen tahmin yöntemlerine başvurulmalıdır.
- Değişkenler arasındaki ilişkiler dikkate alınarak modele yeni denklemler eklenmesi.
- Ridge regresyon, asal bileşenler analizi, faktör analizi, QR ayrıştırması gibi çeşitli istatistikî yöntemlerin kullanılması

## Heteroskedasite ve Otokorelasyon

### Heteroskedasite (Değişen Varyans)

Homoskedasite, alt ana kütle hata terimlerinin varyanslarının bağımsız değişken ya da değişkenlerin alacağı değerlere göre sabit olması olarak tanımlanırken heteroskedasite varyansların sabit olmaması, bağımsız değişkenin alacağı değerlere göre oluşturulan alt ana kütleler itibariyle değişmesi olarak tanımlanmaktadır (Yerdelen Tatoğlu, 2023, s. 119).

### Heteroskedasitenin Nedenleri

Heteroskedasite örneklerle değil ana kütle ile ilgili bir problem olup zaman serilerinden çok yatay kesit verilerde görülmektedir. Heteroskedasitenin temel nedenleri aşağıda sıralandığı gibidir (Yerdelen Tatoğlu, 2023, s. 119-120):

- İhmal edilen değişkenler
- Veri toplama hataları
- Bağımlı değişkende ölçme hatası olması

### Heteroskedasitenin Sonuçları

Heteroskedasite durumunda parametre tahmincileri sapmasız, asimptotik sapmasız ve doğrusaldır, fakat etkin değildir. Bu durumda yapılacak aralık tahminleri, t ve F testleri ve  $R^2$  değeri de etkilenecektir (Yerdelen Tatoğlu, 2023, s. 120).

### Heteroskedasitenin Belirlenmesi

#### Breusch-Pagan / Cook-Weisberg Testi

Heteroskedasitenin Breusch-Pagan (1979) / Cook-Weisberg (1983) testi ile sınılabilmesi için, öncelikle modelin en küçük kareler yöntemi ile tahmininden kalıntılar elde edilmektedir. İki bağımsız değişkenli bir yatay kesit regresyon modelinden hareket edildiğinde:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i$$

Bu model için yardımcı regresyon modelleri:

$$u_i^2 = \delta_0 + \delta_1 \hat{Y}_i + \varepsilon_i$$

$$u_i^2 = \delta_0 + \delta_1 X_{1i} + \delta_2 X_{2i} + \varepsilon_i$$

Temel hipotez iki yardımcı regresyon modeli için sırasıyla aşağıdaki gibidir:

$$H_0 : \delta_1 = 0$$
$$H_0 : \delta_1 = \delta_2 = 0 \quad (\text{heteroskedasite yoktur})$$

Bu hipotezler,  $u_i^2$ 'nin  $X_i$ 'nin fonksiyonları ile korelasyonuz olduğunu söylemektedir. Temel hipotez F testi ile sınanmaktadır. Alternatif olarak, yardımcı regresyon modellerinden hesaplanan ve  $\chi^2$  dağılım  $N \cdot R^2$  istatistiği kullanılmaktadır. Her iki test için de hesaplanan değer tablo değerinden büyükse  $H_0$  hipotezi reddedilmektedir, heteroskedasite vardır (Yerdelen Tatoğlu, 2023, s. 120-121).

### White Testi

Normallik varsayımına dayanmayan White (1980) testinde öncelikle heteroskedasitenin varlığının sınanacağı model tahmin edilmekte ve kalıntılar elde edilmektedir. Daha sonra yardımcı regresyon modeli kurulmaktadır. İki bağımsız değişkenli durumda yardımcı regresyon modeli aşağıdaki gibidir:

$$u_i^2 = \alpha_0 + \alpha_1 X_{1i} + \alpha_2 X_{2i} + \gamma_1 X_{1i}^2 + \gamma_2 X_{2i}^2 + \delta_1 X_{1i} X_{2i} + \varepsilon_i$$

F testi ile katsayılar toplu olarak aşağıdaki gibi sınanmaktadır:

$$H_0: \alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = \delta_1 = 0 \quad (\text{homoskedasite})$$

$$H_1: \text{en az biri sıfırdan farklı} \quad (\text{heteroskedasite})$$

Alternatif olarak  $N \cdot R^2$  değeri hesaplanmakta ve  $\chi^2$  tablo değeri ile karşılaştırılmaktadır.

Her iki durumda da hesaplanan değer tablo değerinden büyükse  $H_0$  hipotezi reddedilmekte ve heteroskedasite olduğuna karar verilmektedir (Yerdelen Tatoğlu, 2023, s. 123-124).

### Diğer Testler

Heteroskedasitenin varlığını sınamak için Spearman'ın sıra korelasyonu testi, benzerlik oranı testi, Ramsey Reset, Goldfeld Quandt, Park ve Glesjer testleri gibi testler de kullanılabilir (Yerdelen Tatoğlu, 2023, s. 125).

### Heteroskedasitenin Düzeltilmesi

#### Logaritmik Dönüşüm

Ele alınan model doğrusal bir modelse logaritmik dönüşüm ölçek farkını azalttığından bir başka ifade ile varyansı küçülttüğünden heteroskedasiteyi düzeltebilmektedir (Yerdelen Tatoğlu, 2023, s. 125).

## Tartılı (Ağırlıklı) En Küçük Kareler Yöntemi

İncelenen modelde heteroskedasite varsa modeli uygun şekilde tartılandırarak heteroskedasiteyi düzeltebilmektedir. Aşağıdaki modelde heteroskedasite varsa,

$$Y_t = \beta_0 + \beta_1 X_t + u_t$$

Aşağıdaki varyanslarla tartılandırılmış model kullanılabilir:

$$\frac{Y_i}{\hat{\sigma}^2} = \beta_0 \frac{1}{\hat{\sigma}^2} + \beta_1 \frac{X_i}{\hat{\sigma}^2} + e_i \quad (3.6)$$

Bu modelinin en küçük kareler ile tahmini tartılı en küçük kareler yöntemi olarak bilinmektedir.  $\hat{\sigma}^2$  bilinmediği için yerine tahmini değeri ya da heteroskedasiteye sebep olan değişken için çeşitli kalıplar kullanılabilir (Yerdelen Tatoğlu, 2023, s. 126).

## White Dirençli Standart Hatalar

Dirençli standart hatalar için ilk çalışmalar Huber (1967), Eicker (1967) ve White (1980) tarafından yapılmıştır. Kalıntıların bağımsız dağılımlı olması durumunda,  $\Omega$  matrisinin bilindiği ve diyagonal olduğu fakat diyagonal elemanların birbirlerine eşit olmadığı varsayımı altında, bir başka ifade ile sadece heteroskedasite olduğu durumda varyansların tahmini için aşağıdaki tahminciyi önermişlerdir (Yerdelen Tatoğlu, 2023, s. 128-129):

$$\text{Var}(\hat{\beta}) = (X'X)^{-1} X'VX (X'X)^{-1}$$

$$\text{Var}(\hat{\beta}) = (X'X)^{-1} X' \text{diag}(\hat{u}_i^2) X (X'X)^{-1}$$

Bu, “*Heteroskedastik Dirençli Varyans Tahmincisi*”dir, “*Huber Tahmincisi*”, “*White Tahmincisi*” ya da “*Eicker Tahmincisi*” olarak da bilinmektedir. MacKinnon ve White (1985) ve Hinkley (1977), bu tahminciyi serbestlik derecesi düzeltilmesi yaparak küçük örneklerde de kullanılabilecek hale getirmişlerdir:

$$\begin{aligned} \text{Var}(\hat{\beta}) &= \frac{N}{N-k} (X'X)^{-1} \left( \sum_{i=1}^N \hat{u}_i^2 x_i x_i' \right) (X'X)^{-1} \\ &= \frac{N}{N-k} (X'X)^{-1} X' \text{diag}(u_i^2) X (X'X)^{-1} \end{aligned}$$

Yine aynı araştırmacılar tarafından önerilen başka bir yaklaşımda, küçük örnek düzeltilmesi ( $N/N-k$ ) yerine ( $1/1-h_i$ ) ( $h_i$ : tahmin edilen matrisinin diyagonal elemanlarıdır) ile yapılmaktadır. Long ve Ervin (2000), küçük örnek özelliklerini iyileştirmek ve sapan gözlemlere daha az ağırlık vermek için  $1/(1-h_i)^2$

düzeltilmesini, Cribari-Neto ve Zarkos (2004) ise, yine özellikle sapan değerlerin varlığında küçük örnek özelliklerini iyileştirmek için,  $1/(1-h_i)^{\delta_i}$  düzeltilmesini önermiştir (burada  $\delta_i = \min\{4, h_i/\bar{h}\}$  'dir) (Yerdelen Tatoğlu, 2023, s. 129).

## Otokorelasyon

Hata terimlerinin diğer dönemlere ait hata terimleri ile ilişkili olması *otokorelasyon* olarak adlandırılmaktadır (Yerdelen Tatoğlu, 2023, s. 130).

Hata teriminin t dönemindeki değeri ile t-1 dönemi arasındaki korelasyon yani birinci mertebeden otokorelasyon katsayısı aşağıdaki gibi hesaplanmaktadır:

$$\rho = \frac{\text{cov}(u_t, u_{t-1})}{\sqrt{\text{var}(u_t)\text{var}(u_{t-1})}} = \frac{\sum_{t=2}^T u_t u_{t-1}}{\sum_{t=2}^T u_{t-1}^2}$$

$\rho$ 'ya *anakütle otokorelasyon katsayısı* ismi verilmektedir ve  $\rho$  korelasyon katsayısı gibi  $\pm 1$  aralığında yer almaktadır, 0 ise otokorelasyon yoktur.

Birinci mertebeden otokorelasyon olması durumuna sık rastlanmakla birlikte bazı modellerde şokların etkilerinin birden fazla döneme yayılması nedeniyle üst mertebelerden otokorelasyon da görülebilmektedir. Ayrıca aylık verilerle çalışıldığında 12., çeyreklik verilerle çalışıldığında 4. mertebeden otokorelasyon söz konusu olmaktadır (Yerdelen Tatoğlu, 2023, s. 130-131).

## Otokorelasyonun Nedenleri

Otokorelasyon, heteroskedasite gibi ana kütle ilgili bir sorundur ve sadece zaman serisi ve panel verilerle çalışırken karşılaşılmaktadır. Otokorelasyonun nedenleri aşağıda sıralanmıştır (Yerdelen Tatoğlu, 2023, s. 131-132):

- İhmal edilen değişkenler,
- Modelin fonksiyonel şeklinin yanlış belirlenmesi,
- Verilerde ölçme hatası olması,
- Şokların etkileri,
- Gecikmeli değişkenler,
- Durağan olmama.

## Otokorelasyonun Sonuçları

Otokorelasyonun sonuçları heteroskedasitenin sonuçları ile benzerdir; en küçük kareler tahmincileri sapmasız ve tutarlıdır, fakat etkin değildir. Bu nedenle hesaplanan t istatistiği ve güven aralıkları da sapmalı, belirsizlik katsayısı olduğundan büyük tahmin edilmekte ve F istatistiği de etkilenmektedir (Yerdelen Tatoğlu, 2023, s. 132).

## Otokorelasyonun Belirlenmesi

### Grafik İncelemesi

Hata terimlerinin bir önceki dönem hata terimleri ile grafikleri incelenerek birinci mertebeden (AR(1) korelasyon) otokorelasyonun varlığına karar verilmektedir. Eğer grafik pozitif eğimli ise pozitif otokorelasyon, negatif eğimli ise negatif otokorelasyon, sıfır etrafında dağınık bir örüntü gösteriyorsa otokorelasyon olmadığı ifade etmektedir (Yerdelen Tatoğlu, 2023, s. 132).

### Durbin Watson Testi

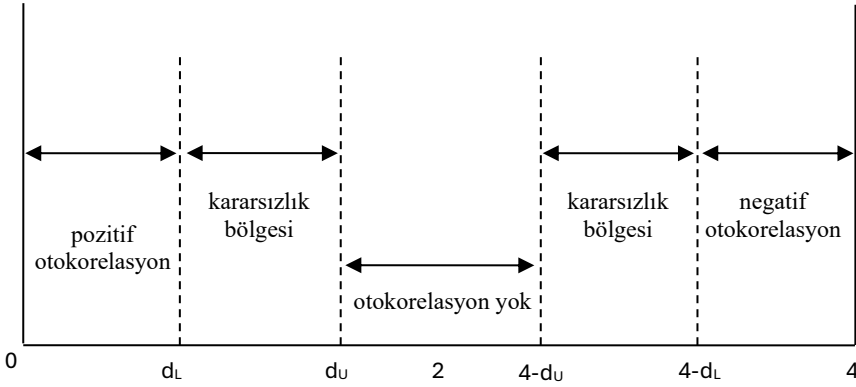
Otokorelasyonun belirlenmesinde en sık kullanılan test Durbin Watson testidir. Uygulanabilmesi bazı koşullara bağlıdır (Yerdelen Tatoğlu, 2023, s. 133):

- Modelde sabit terim yer almalıdır.
- Otoregresif modellerde geçerli değildir.
- Sadece birinci mertebeden otokorelasyonu belirlemek için kullanılmaktadır.
- Bazen bu test otokorelasyonun varlığına kesin olarak karar verememektedir.

Durbin Watson istatistiği:

$$d = \frac{\sum_{t=2}^T (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^T (\hat{u}_t)^2} = 2 \left( 1 - \frac{\sum_{t=2}^T \hat{u}_t \hat{u}_{t-1}}{\sum_{t=1}^T \hat{u}_t^2} \right) = 2(1 - \hat{\rho})$$

$\rho$ , -1 ile +1 arasında değer alacağından,  $\rho = 1$  ise,  $d = 0$ ,  $\rho = 0$  ise,  $d = 2$  ve  $\rho = -1$  ise  $d = 4$ 'tür. Dolayısıyla  $d$ , 0 ile 4 arasında değer almaktadır ( $0 \leq d \leq 4$ ).  $k$  bağımsız değişken sayısı ve  $n$  gözlem sayısı ile Durbin Watson tablosundan  $d_L$  ve  $d_U$  değerleri bulunmaktadır ve daha sonra değerler aşağıdaki tabloya yerleştirilmektedir.  $d$  istatistiği hangi bölgede kalıyorsa ona göre yorum yapılmaktadır (Yerdelen Tatoğlu, 2023, s. 134).



### Otokorelasyon Karar Bölgeleri

Aylık ya da üçer aylık verilerle çalışıldığında 12 ve 4. mertebeden otokorelasyonu sınamak için sırasıyla King ve Wallis testleri tercih edilmektedir. Test istatistiği Durbin Watson test istatistiğine benzemekle birlikte tek fark  $\hat{u}_{t-1}$  yerine  $\hat{u}_{t-12}$  ve  $\hat{u}_{t-4}$  kullanılmasıdır (Yerdelen Tatoğlu, 2023, s. 135).

Otoregresif modellerle çalışıldığında Durbin-h testi kullanılabilmektedir.

$$d_h = \left(1 - \frac{1}{d}\right) \sqrt{\frac{T}{1 - TVar(\alpha)}}$$

Burada  $\alpha$  gecikmeli değişkenin parametresidir. Bu testin bir kısıtı  $TVar(\alpha) < 1$  olması gerekliliğidir (Yerdelen Tatoğlu, 2023, s. 136).

### Breusch Godfrey Lagrange Çarpamı (LM) Testi

Breusch Godfrey LM testi yüksek mertebeden otokorelasyonun varlığını sına-maya da imkân vermekte, otoregresif ve orijinden geçen regresyon modellerinde de kullanılabilmektedir (Yerdelen Tatoğlu, 2023, s. 137).

Öncelikle model tahmin edilmekte, kalıntılar elde edilmekte ve p. mertebeden otokorelasyonun varlığını sınamak için aşağıdaki model tahmin edilmektedir.

$$u_t = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \rho_1 u_{t-1} + \rho_2 u_{t-2} + \dots + \rho_p u_{t-p} + v_t$$

Daha sonra aşağıdaki hipotezlerin sınanması için test yapılmaktadır:

$$H_0: \rho_1 = \rho_2 = \dots = \rho_p = 0 \text{ (p. mertebeden otokorelasyon yoktur)}$$

$$H_1: \rho_1 \neq \rho_2 \neq \dots \neq \rho_p \neq 0 \text{ (p. mertebeden otokorelasyon vardır)}$$

Modelde yer alan gecikmeli hata terimlerinin katsayılarının birlikte anlamlılığı F ya da büyük örnekler için LM testi ile test edilmektedir. LM test istatistiği:

$$LM = T \cdot R^2$$

Burada  $R^2$  yardımcı regresyon modelinin belirginlik katsayısıdır. Test istatistiği p serbestlik dereceli  $\chi^2$  tablo değeri ile karşılaştırılmakta, hesaplanan değer tablo değerinden büyükse  $H_0$  hipotezi reddedilmekte ve p. mertebeden otokorelasyon olduğu anlaşılmaktadır (Yerdelen Tatoğlu, 2023, s. 137-138).

### **Durbin'in Alternatif Testi**

Durbin'in alternatif testi yüksek mertebeden otokorelasyonun varlığını sınılamaya imkân vermektedir. Ayrıca otoregresif ve orijinden geçen regresyon modellerinde de kullanılabilir (Yerdelen Tatoğlu, 2023, s. 139).

p. mertebeden otokorelasyonun varlığını sınamak için aşağıdaki yardımcı regresyon modeli tahmin edilmektedir.

$$u_t = \beta_0 + \rho_1 u_{t-1} + \rho_2 u_{t-2} + \dots + \rho_p u_{t-p} + v_t$$

Aşağıdaki hipotezler sınanmaktadır:

$H_0: \rho_1 = \rho_2 = \dots = \rho_p = 0$  (p. mertebeden otokorelasyon yoktur)

$H_1: \rho_1 \neq \rho_2 \neq \dots \neq \rho_p \neq 0$  (p. mertebeden otokorelasyon vardır)

Gecikmeli hata terimlerinin katsayılarının birlikte anlamlılığı F ya da büyük örnekler için LM testi ile test edilmektedir (Yerdelen Tatoğlu, 2023, s. 139).

### **Diğer Testler**

Bu testlerin dışında sıra testi, ki-kare testi, Farebrother testi, Berenblut-Webb testi, Von-Neuman oranı testi gibi testler de vardır (Yerdelen Tatoğlu, 2023, s. 141).

### **Otokorelasyonun Düzeltilmesi**

#### **Birinci Farklar Yöntemi**

Otokorelasyonun etkisini ortadan kaldırmak için değişkenlerin bir önceki devre değerlerinden farkları alınıp yeni bir model oluşturulmaktadır. Aşağıdaki modelin:

$$Y_t = \beta_0 + \beta_1 X_t + u_t$$

Birinci farkı:

$$Y_t - Y_{t-1} = \beta_1 (X_t - X_{t-1}) + (u_t - u_{t-1})$$
$$\Delta Y_t = \beta_1 \Delta X_t + \Delta u_t$$

Bu model *birinci farklar modeli* olarak bilinmektedir. Sabit terim ve birinci devre düşmüştür. Bu yöntem sadece birinci mertebeden otokorelasyonun düzeltilmesinde kullanılmaktadır (Yerdelen Tatoğlu, 2023, s. 141-142).

### **Cochrane-Orcutt ve Prais-Winsten Dönüşümleri ile Genelleştirilmiş En Küçük Kareler Yöntemi**

Cochrane–Orcutt yöntemi, hata terimindeki otokorelasyon katsayısını ( $\rho$ ) tahmin ederek değişkenleri dönüştürmekte ve daha sonra dönüştürülmüş modeli EKK ile tahmin etmektedir.

$$Y_t - \rho Y_{t-1} = (\beta_0 - \rho \beta_0) + \beta_1 X_t - \rho \beta_1 X_{t-1} + u_t - \rho u_{t-1}$$
$$Y_t - \rho Y_{t-1} = \beta_0 (1 - \rho) + \beta_1 (X_t - \rho X_{t-1}) + (u_t - \rho u_{t-1})$$

Bu dönüşüm, genelleştirilmiş farklar (Cochrane–Orcutt dönüşümü) olarak adlandırılmaktadır. Bu son eşitlik  $\rho=1$  için birinci farklar yöntemine eşittir. Bu yöntemde ilk gözlem dönüşüm sırasında kaybedilmektedir.

Prais–Winsten yöntemi Cochrane–Orcutt yaklaşımına benzemekle birlikte ilk gözlemi korumaktadır. Bu nedenle özellikle küçük örneklemelerde daha fazla tercih edilmektedir.

Prais–Winsten ya da Cochrane–Orcutt dönüşümü uygulanmış model, Genelleştirilmiş En Küçük Kareler (GEKK) yöntemi ile tahmin edilmektedir:

$$\hat{\beta} = (X' \Omega^{-1} X)^{-1} (X' \Omega^{-1} Y)$$

$\Omega$ , hata teriminin varyans-kovaryans matrisini göstermektedir (Yerdelen Tatoğlu, 2023, s. 145-146). Otokorelasyon katsayısı ( $\rho$ ) bilinmediğinden uygulamada çeşitli yöntemlerle tahmin edilmektedir. En yaygın yaklaşımlar kalıntılar regresyonu ve Durbin–Watson istatistiğine dayalı tahmin yöntemleridir (Yerdelen Tatoğlu, 2023, s. 148):

### **Heteroskedasite ve Otokorelasyonun Düzeltilmesi**

### **Cochrane-Orcutt ve Prais-Winsten Dönüşümleri ile Genelleştirilmiş En Küçük Kareler Yöntemi**

Cochrane–Orcutt ve Prais–Winsten dönüşümleri, Genelleştirilmiş En Küçük Kareler (GEKK) yaklaşımıyla birlikte kullanılarak hem heteroskedastisiteyi hem de

birinci mertebeden otokorelasyonu dikkate alacak şekilde genişletilebilmektedir (Yerdelen Tatoğlu, 2023, s. 152-153).

### Newey-West Dirençli Standart Hatalar

Newey-West (1987, 1994) tarafından önerilen yaklaşımda heteroskedasite ve otokorelasyona karşı dirençli standart hatalar elde edilmektedir. Newey-West'in genelleştirilmiş momentler temelli kovaryans tahmincisi White'in tahmincisinin genişletilmiş bir halidir ve otokorelasyon yokken Newey-West tahmincisi White tahmincisine eşittir. Bu tahminci otokorelasyon için kalıntının ağırlıklı çarpaz çarpımlarını da hesaba katmaktadır.

$$Var(\hat{\beta}) = \frac{1}{T} \sum_{t=1}^T \hat{u}_t^2 X_t X_t' + \frac{1}{T} \sum_{\tau=1}^{T-1} w_{m(T)^{(\tau)}} \sum_{t=\tau+1}^T (X_{t-\tau} \hat{u}_{t-\tau} \hat{u}_t X_t' + X_t \hat{u}_t \hat{u}_{t-\tau} X_{t-\tau}') )$$

Burada,  $w_{m(T)}$  Bartlett Kernel fonksiyonunu ifade etmektedir.

Bu yöntemde parametre tahminleri değiştirilmemekte, yalnızca standart hatalar düzeltilmektedir (Yerdelen Tatoğlu, 2023, s. 155).

### Normal Dağılım

En küçük kareler yönteminde hata terimlerinin normal dağılması, parametre tahmini için gerekli değildir. Ancak hipotez testlerinin yapılabilmesi ve güven aralıklarının oluşturulabilmesi için hata terimlerinin normal dağıldığı varsayılmaktadır. Bu nedenle uygulamada hata terimlerinin normal dağılıp dağılmadığı çeşitli normallik testleri ile sınanmaktadır (Yerdelen Tatoğlu, 2023, s. 156-157).

### Jarque Bera Testi

Jarque ve Bera (1987) tarafından önerilen bu testte hata teriminin ortalama etrafındaki dağılımının çarpıklık ve basıklık değerlerinin normal dağılımın çarpıklık ve basıklık değerlerinden istatistikî olarak farklı olup olmadığı sınanmaktadır (Yerdelen Tatoğlu, 2023, s. 157).

Temel hipotez:

$H_0: S=0$  ve  $K=3$  (hata terimleri normal dağılmaktadır)

Burada S (skewness) çarpıklığı ve K (kurtosis) basıklığı ifade etmekte olup normal dağılım için alması gereken değerler sırasıyla 0 ve 3'tür. Dağılımın çarpıklık değeri sıfırdan küçükse (negatifse) sola, büyükse (pozitifse) sağa çarpıktır; basıklık değeri ise üçten büyükse sivri, küçükse basıktır. Jarque Bera test istatistiği (Yerdelen Tatoğlu, 2023, s. 157):

$$JB = n \left[ \frac{S^2}{6} + \frac{(K-3)^2}{24} \right]$$

Burada S üçüncü ve K dördüncü moment yardımıyla elde edilmektedir:

$$S = \frac{\mu_3}{\sigma^3} \text{ ve } K = \frac{\mu_4}{\sigma^4} \quad \left( \mu_3 = \frac{\sum_{i=1}^N \hat{u}_i^3}{n} \text{ ve } \mu_4 = \frac{\sum_{i=1}^N \hat{u}_i^4}{n} \right)$$

Jarque Bera test istatistiği iki serbestlik dereceli  $\chi^2$  dağılımına ( $JB \sim \chi^2_2$ ) uyumaktadır.  $H_0$  hipotezi reddedilemezse, hata teriminin normal dağıldığı sonucuna varılmaktadır (Yerdelen Tatoğlu, 2023, s. 157-158).

### D'Agostino, Belanger ve D'Agostino Testi

D'Agostino, Belanger ve D'Agostino (1990), çarpıklık ve basıklığın ayrı ayrı ve birlikte sınanması için bir test süreci önermiştir (Yerdelen Tatoğlu, 2023, s. 160).

Temel hipotezler sırasıyla,

$$H_0: S=0 \quad (\text{çarpıklık sıfırdır}) \quad (1)$$

$$H_0: K=3 \quad (\text{basıklık üçtür}) \quad (2)$$

$$H_0: S=0, K=3 \quad (\text{hata terimleri normal dağılmaktadır}) \quad (3)$$

Çarpıklığın testi için test istatistiği (Yerdelen Tatoğlu, 2023, s. 160-161):

$$Z_1 = \frac{1}{\sqrt{\ln W}} \ln \left[ \frac{Y}{\alpha} + \left\{ \left( \frac{Y}{\alpha} \right)^2 + 1 \right\}^{1/2} \right]^{-1}$$

Bu test istatistiği temel hipotez altında yaklaşık standart normal dağılıma sahiptir. Burada:

$$Y = S \left\{ \frac{(n+1)(n+3)}{6(n-2)} \right\}^{1/2}, \quad W^2 = -1 + \left[ 2 \{ \beta_2(S) - 1 \} \right]^{1/2}$$

$$\beta_2(S) = \frac{3(n^2 + 27n - 70)(n+1)(n+3)}{(n-2)(n+5)(n+7)(n+9)}, \quad \alpha = \left\{ \frac{2}{W^2 - 1} \right\}^{1/2}$$

Basıklığın testi için test istatistiği (Yerdelen Tatoğlu, 2023, s. 161):

$$Z_2 = \frac{1}{\sqrt{2/(9A)}} \left[ \left( 1 - \frac{2}{9A} \right) - \left\{ \frac{1 - (2/A)}{1 + X \sqrt{2/(A-4)}} \right\}^{1/3} \right]$$

Bu test istatistiği de temel hipotez altında yaklaşık standart normal dağılıma sahiptir. Burada:

$$X = \frac{K - E(K)}{\sqrt{\text{var}(K)}}, \quad E(K) = \frac{3(n-1)}{(n+1)}, \quad \text{var}(K) = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}$$

$$A = 6 + \frac{8}{\sqrt{\beta_1(K)}} \left[ \frac{2}{\sqrt{\beta_1(K)}} + \left\{ 1 + \frac{4}{\sqrt{\beta_1(K)}} \right\}^{1/2} \right]$$

$$\sqrt{\beta_1(K)} = \frac{6(n^2 - 5n + 2)}{(n+7)(n+9)} \left\{ \frac{6(n+3)(n+5)}{n(n-2)(n-3)} \right\}^{1/2}$$

Normal dağılımı sınamak için birleşik test istatistiği ise:

$$Z_{1,2}^2 = Z_1^2 + Z_2^2$$

Bu test istatistiği normal dağılım temel hipotezi altında 2 serbestlik dereceli  $\chi^2$  dağılımına sahiptir (Yerdelen Tatoğlu, 2023, s. 161).

## Diğer Testler

Bu testlerin dışında ki-kare uygunluk testi, Shapiro Wilk ve Shapiro Francia testleri gibi testler de kullanılmaktadır (Yerdelen Tatoğlu, 2023, s. 163-164).

## Düzeltilme Yöntemleri

Normal dağılmayan serilerde örnek hacmi arttıkça Merkezi Limit Teoremi gereği örnekleme dağılımı normal dağılıma yaklaşmaktadır. Bu nedenle mümkün olduğunda örnek birim sayısının artırılması tercih edilmektedir. Bunun mümkün olmadığı durumlarda ise logaritmik veya Box-Cox dönüşümleri uygulanabilmekte ya da büyük örneklem özelliklerinden yararlanan dirençli (robust) tahmin yöntemleri kullanılabilmektedir (Yerdelen Tatoğlu, 2023, s. 164).

## KUKLA DEĞİŞKENLER

Ekonometrik modellerde kullanılan değişkenler ölçülebilir nicel değişkenler olabileceği gibi cinsiyet, eğitim durumu, meslek, bir ülkenin uluslararası bir topluluğa üyeliği veya bir şehrin denize kıyısının bulunması gibi nitel özellikleri de temsil edebilmektedir. Bu tür değişkenler kukla değişken, gölge değişken (dummy değişken) veya kategorik değişken olarak adlandırılmaktadır. Kukla değişkenler genellikle 0-1 ya da birden fazla kategori söz konusu olduğunda 0-1-2-3 gibi kodlarla ifade edilmektedir (Yerdelen Tatoğlu, 2023, s. 179).

Kodlamada hangi kategorinin referans (temel) kategori olarak seçildiği katsayıların yorumunu etkilemekte, ancak modelin tahmin sonuçlarını değiştirmemektedir (Yerdelen Tatoğlu, 2023, s. 179).

Kukla değişkenler ayrıca krizler, politika değişiklikleri, savaşlar, salgın hastalıklar, yapısal kırılmalar ve mevsim etkileri gibi ekonomik olayların modellenmesinde de yaygın olarak kullanılmaktadır (Yerdelen Tatoğlu, 2023, s. 180).

### İkili Değer Alan Kukla Değişkenli Modeller

İkili değer alan kukla değişkenler, modelde tek başına veya diğer bağımsız değişkenlerle birlikte farklı model spesifikasyonlarında yer alabilmektedir (Yerdelen Tatoğlu, 2023, s. 180).

#### Tek Bağımsız Değişkene Sahip Modeller

Bir modelde sadece tek bağımsız değişken olduğu ve onun da kukla değişken olduğu düşünüldüğünde model aşağıdaki gibi ifade edilebilmektedir:

$$Y_i = \beta_0 + \beta_1 D_i + u_i$$

Burada kukla değişken  $D_i$  sadece 0 ve 1 gibi iki değer aldığı durumda aşağıdaki gibi tanımlanmaktadır (Yerdelen Tatoğlu, 2023, s. 180):

$$D_i = 0 \text{ ya da } E(Y_i | D_i = 0) = \beta_0$$

$$D_i = 1 \text{ ya da } E(Y_i | D_i = 1) = \beta_0 + \beta_1$$

### İki Bağımsız Değişkenden Birisinin Kukla Değişken Olduğu Modeller

Eğer modeldeki bağımsız değişkenlerden birisi kukla değişkense modele dahil ediliş biçimine bağlı olarak sadece sabit parametreyi, sadece eğim parametresini ya da her ikisini birden etkileyebilecektir (Yerdelen Tatoğlu, 2023, s. 182).

### Sadece Sabit Parametrenin Etkilendiği Modeller

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_i + u_i$$

Burada (Yerdelen Tatoğlu, 2023, s. 182-183):

$$\begin{aligned} D_i = 0 \text{ ya da } E(Y_i | D_i = 0) &= \beta_0 + \beta_1 X_i \\ D_i = 1 \text{ ya da } E(Y_i | D_i = 1) &= \beta_0 + \beta_1 X_i + \beta_2 \\ &= (\beta_0 + \beta_2) + \beta_1 X_i \end{aligned}$$

### Sadece Eğim Parametresinin Etkilendiği Modeller

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_i X_i + u_i$$

Burada (Yerdelen Tatoğlu, 2023, s. 184):

$$\begin{aligned} D_i = 0 \text{ ya da } E(Y_i | D_i = 0) &= \beta_0 + \beta_1 X_i \\ D_i = 1 \text{ ya da } E(Y_i | D_i = 1) &= \beta_0 + \beta_1 X_i + \beta_2 X_i \\ &= \beta_0 + (\beta_1 + \beta_2) X_i \end{aligned}$$

### Sabit ve Eğim Parametrelerinin Birlikte Etkilendiği Modeller

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_i + \beta_3 D_i X_i + u_i$$

Burada (Yerdelen Tatoğlu, 2023, s. 185-186):

$$\begin{aligned} D_i = 0 \text{ ya da } E(Y_i | D_i = 0) &= \beta_0 + \beta_1 X_i \\ D_i = 1 \text{ ya da } E(Y_i | D_i = 1) &= \beta_0 + \beta_1 X_i + \beta_2 + \beta_3 X_i \\ &= (\beta_0 + \beta_2) + (\beta_1 + \beta_3) X_i \end{aligned}$$

### İkiden Fazla Değer Alan Kukla Değişkenli Modeller

$$Y_i = \beta_0 + \beta_1 D_i + u_i$$

Burada (Yerdelen Tatoğlu, 2023, s. 187-188):

$$\begin{aligned} D_i = 0 \text{ ya da } E(Y_i | D_i = 0) &= \beta_0 \\ D_i = 1 \text{ ya da } E(Y_i | D_i = 1) &= \beta_0 + \beta_1 \\ D_i = 2 \text{ ya da } E(Y_i | D_i = 2) &= \beta_0 + 2\beta_1 \end{aligned}$$

## Birden Fazla Kukla Değişkene Sahip Modeller

Bu modeller iki şekilde kurulabilmektedir (Yerdelen Tatoğlu, 2023, s. 190):

- Kukla değişkenler birbirinden bağımsız ise;

$$Y_i = \beta_0 + \beta_1 D_{1i} + \beta_2 D_{2i} + u_i$$

Bu durumda modelin tahmini ve katsayıların yorumu tek kukla değişkenli modellere benzer şekilde yapılmaktadır.

- Kukla değişkenler aynı nitel değişkenin kategorilerinden oluşuyorsa:

$$Y_i = \beta_0 + \beta_1 D_i + u_i$$

Bu durumda kategori sayısından bir eksik sayıda kukla değişken oluşturulmaktadır. Böylece her katsayı, ilgili kategorinin referans kategoriye göre etkisini gösterdiğinden daha fazla bilgi vermektedir (Yerdelen Tatoğlu, 2023, s. 190).

## Kukla Değişken Tuzağı

Eğer modele sınıf (kategori) sayısı kadar kukla değişken eklenirse, bu kategorilerin birbirinden bağımsız olmaması sebebiyle meydana gelen tam çoklu doğrusal bağlantı modelin tahmin edilememesine neden olabilmektedir. Bu duruma *kukla değişken tuzağı* denilmektedir.

Kukla değişken tuzağına düşülmemesi için kategori sayısından bir eksik sayıda kukla değişken modele ilave edilmelidir. M sayıda kategori varsa M-1 sayıda kukla değişken kullanılmalıdır (Yerdelen Tatoğlu, 2023, s. 191).

## Mevsim ve Ay Etkilerinin Kukla Değişkenlerle Belirlenmesi

Üçer aylık ya da aylık verilerin yer aldığı zaman serisi verileriyle çalışılan bir ekonometrik modele mevsimlerin ya da ayların etkilerinin ilave edilmesi istenildiği takdirde kukla değişken tuzağından kaçınmak için 4 mevsim için 3, 12 ay için 11 adet kukla değişken kullanılmalıdır (Yerdelen Tatoğlu, 2023, s. 191-192).

## Mevsimsellikten Arındırma

Seride mevsim etkileri varsa değişkenler mevsimsellikten arındırılarak modele alınmalıdır. Mevsimsel düzeltme için çok sayıda yöntem vardır, bunlar hareketli ortalamalar, Census X-11, 12, 13, treamo-seats, Holt-Winters mevsimsel düzeltme yöntemleri olarak sayılabilmektedir (Yerdelen Tatoğlu, 2023, s. 195-196).

## Logaritmik Modellerde Kukla Değişkenler

Bağımlı değişkenin logaritmik olduğu modellerde de bağımsız değişkenler olarak kukla değişkenler yer alabilmektedir:

$$\ln Y_i = \beta_0 + \beta_1 D_i + u_i$$

$\beta_1$  parametresinin yorumlanmasında literatürde birkaç yaklaşım önerilmekle birlikte en fazla Halvorsen ve Palmquist (1980) tarafından önerilen yaklaşım kullanılmaktadır. Bu yaklaşıma göre kukla değişkenin bağımlı değişken üzerindeki etkisi aşağıdaki gibi bulunmaktadır (Yerdelen Tatoğlu, 2023, s. 198):

$$100 * (e^{\beta_1} - 1)$$

## YAPISAL DEĞİŞİKLİK

İktisadi değişkenleri etkileyen faktörler değişkenlerin trendinde ve regresyon doğrusunda bazen geçici bazense kalıcı değişiklikler oluşturmaktadır. Bu değişiklikler politika değişiklikleri (vergi gelirlerini arttırıcı, enflasyonu düşürücü, ihracatı arttırıcı tedbirler gibi) sonucunda, teknolojik bir değişiklik sonucunda ya da savaş, kuraklık, deprem, büyük grevler, krizler, salgın hastalıklar gibi şoklar sonucunda meydana çıkabilmektedir. Geçici değişiklikler regresyon doğrusunun trendinde geçici bir kırılma yaratmaktadır, şok geçtikten sonra etkisi de kaybolmaktadır (Yerdelen Tatoğlu, 2023, s. 207).

### Yapısal Değişiklik Testleri

Regresyon doğrusunda yapısal bir kırılmanın olup olmadığı çeşitli testler yardımıyla sınanmaktadır. Bunlardan bir kısmında kırılma tarihi içsel olarak yani testler sonucunda belirlenmekte, bir kısmında ise dışsal olarak verilmekte ve bu tarihte yapısal bir kırılma olup olmadığına bakılmaktadır (Yerdelen Tatoğlu, 2023, s. 208).

#### 1. Grafiksel İncelemeler ve Testler

##### Serpilme Diyagramı İncelemesi

Bağımlı ve bağımsız değişkenin serpilme diyagramının incelenmesi yapısal kırılma olup olmadığının önsel tespiti hakkında önemli bilgi vermektedir (Yerdelen Tatoğlu, 2023, s. 208).

##### Cusum Testi

Brown, Durbin ve Evans (m kalıntılarından hesaplanan istatistiğin grafiği çizilmekte ve alt ve üst güven sınırları oluşturulmaktadır. Çizilen grafiğin alt veya üst

sınırların dışına çıkması durumunda yapısal değişikliğin olduğu sonucuna varılmaktadır. Test istatistiği (Yerdelen Tatoğlu, 2023, s. 210):

$$W_t = \sum_{r=k+1}^t \frac{w_r}{\sigma} \quad t = k+1, k+2, \dots, T$$

Burada  $w_r$  ardışık kalıntılar ve  $\sigma_w$  ise ardışık kalıntılar  $w_t$ 'nin standart hatasıdır. Sıfırdan anlamlı bir sapma olup olmadığı, aralarındaki mesafe  $t$  ile artan bir çift %5 anlamlılık çizgi aralığı ile değerlendirilmektedir. %5 anlamlılık çizgisi aşağıdaki noktalar birleştirilerek bulunmaktadır:

$$\left[ k, \pm 0.948(T-k)^{(1/2)} \right] \text{ ve } \left[ T, \pm 3 \times 0.948(T-k)^{(1/2)} \right]$$

$W_t$  çizilen sınırlar aralığındaysa yapısal değişiklik olmadığına, bu aralık dışına çıkarsa parametrelerin kararsız olduğuna, bir başka ifade ile yapısal değişiklik olduğuna karar verilmektedir (Yerdelen Tatoğlu, 2023, s. 210).

### Cusum-Square Testi

Cusum-Square test istatistiği (Yerdelen Tatoğlu, 2023, s. 211):

$$S_t = \frac{\sum_{r=1}^t w_r^2}{\sum_{r=k+1}^T w_r^2}$$

$E(S) \pm C_0$  güven sınırları içindeyse yapısal değişiklik yoktur. Burada  $E(S)$ :

$$E(S_t) = \frac{t-k}{T-k}$$

$C_0$ ,  $T$  gözlem ve  $k$  parametre sayısı ile tablo değeridir (Brown, Durbin ve Evans (1975)).  $S_t$ 'nin grafiği çizilen alt veya üst sınırların dışına çıkarsa yapısal değişikliğin olduğu sonucuna varılmaktadır (Yerdelen Tatoğlu, 2023, s. 211).

## 2. Yapısal Kırılma Tarihinin İçsel Olarak Belirlendiği Testler

### LR ve Wald Testleri

Veri setindeki tüm olası kırılma tarihleri için LR ve Wald testlerinin hesaplanması ve bunlar arasından supremum, ortalama ya da üstel olarak seçim yapılmasıyla yapısal kırılma tarihi belirlenmektedir. LR ve Wald istatistiklerini hesaplamak için kısıtlı ve kısıtsız modeller sırasıyla:

$$Y_t = \beta_0 + \beta_1 X_t + u_t$$

$$Y_t = \beta_0 + \beta_1 X_t + \beta_1 D_t + \beta_1 D_t X_t + u_t$$

Burada D belirlenen yapısal kırılma tarihi öncesi için 0, sonrası için 1 değerini alan bir kukla değişkendir. Test istatistikleri aşağıdaki gibidir (Yerdelen Tatoğlu, 2023, s. 213):

$$LR = n \ln \left( \frac{ESS_R}{ESS_{UR}} \right)$$

$$Wald = \frac{ESS_R - ESS_{UR}}{ESS_{UR}/n}$$

Supremum test istatistiği yapısal değişiklik yoktur temel hipotezini test etmek için hesaplanan LR ya da Wald test istatistiklerinin maksimumudur.

$$\text{supremum } S_T = \sup_{t_1 \leq t \leq t_2} S_T(t)$$

Burada  $S_T(t)$ , belirlenen kırılma tarihi t'deki Wald ya da LR test istatistikleridir. Maksimum değere sahip olan tarih yapısal kırılma tarihi olarak belirlenmektedir. Test istatistiğinin ortalama ve üstel versiyonları ise sırasıyla,

$$\text{average } S_T = \frac{1}{t_2 - t_1 + 1} \sum_{t=t_1}^{t_2} S_T(t)$$

$$\text{exponential } S_T = \ln \left[ \frac{1}{t_2 - t_1 + 1} \sum_{t=t_1}^{t_2} \exp \left\{ \frac{1}{2} S_T(t) \right\} \right]$$

Her üç test istatistiğinin dağılımı da standart değildir. Andrews (1993) dağılımlarını geliştirmiş ve Hansen (1997) yaklaşık asimptotik olasılık değerlerini elde etmiştir (Yerdelen Tatoğlu, 2023, s. 213).

### 3. Yapısal Kırılma Tarihinin Dışsal Olarak Belirlendiği Testler

#### LR ve Wald Testleri

Belirlenen tarihte yapısal kırılmanın varlığını sınamak için LR ve Wald, testleri kullanılabilir. Her iki testte de temel hipotez belirlenen tarihte yapısal değişiklik olmadığını ifade etmektedir. Her iki test istatistiği de k parametre sayısına eşit serbestlik dereceli  $\chi^2$  dağılımına uymaktadır (Yerdelen Tatoğlu, 2023, s. 215).

#### Chow Testi

Eğer yapısal değişiklik varsa, model yapısal değişikliğin öncesi ve sonrası için ayrı ayrı tahmin edilip kalıntıların kareleri toplamı hesaplandığında, bu

kalıntıların kareleri toplamı tüm gözlemler için tahmin edilen modelden hesaplanan kalıntıların kareleri toplamından küçük olacaktır. Yapısal değişiklik olmadığı durumda ise eşit olacaktır. Bu bilgi kullanılarak Chow testi yapılmaktadır (Yerdelen Tatoğlu, 2023, s. 217).

Chow testi için hipotezler:

$H_0$ : yapısal değişiklik yoktur.

$H_1$ : yapısal değişiklik vardır.

Chow test istatistiği (Yerdelen Tatoğlu, 2023, s. 217):

$$F = \frac{\left[ \sum_{t=1}^T u_t^2 - \left( \sum_{t=1}^s u_{1t}^2 + \sum_{t=s+1}^T u_{2t}^2 \right) \right] / k}{\left( \sum_{t=1}^s u_{1t}^2 + \sum_{t=s+1}^T u_{2t}^2 \right) / (T - 2k)}$$

Burada s yapısal değişikliğin olduğu dönem olmakla birlikte ilk dönemde  $t_1$  ve ikinci dönemde  $t_2$  sayıda gözlem olduğu varsayılın.  $\sum_{t=1}^s u_{1t}^2$  ve  $\sum_{t=s+1}^T u_{2t}^2$  yapısal değişiklik öncesi ve sonrası için kurulan modellerden  $\sum_{t=1}^T u_t^2$  ise tüm gözlemleri içeren modelden elde edilen kalıntı kareler toplamıdır. k parametre sayısı ve T toplam gözlem sayısı olmak üzere F test istatistiği k, T-2k serbestlik dereceli F dağılımına uymaktadır.  $H_0$  hipotezi reddedilirse, belirtilen tarihte yapısal değişiklik olduğu kabul edilmektedir (Yerdelen Tatoğlu, 2023, s. 217).

Chow testinin ön koşulu yapısal değişiklik öncesi ve sonrası olarak ayrılan iki grubun varyansların eşit olmasıdır. İki grubun varyanslarının eşitliği yine F testi kullanılarak sınanabilmektedir (Yerdelen Tatoğlu, 2023, s. 218).

## Yapısal Değişikliğin Modellenmesi

### 1. Sabitte Yapısal Kırılmaya Sahip Model

Tek kırılma olduğu durumda, yapısal kırılma öncesi için 0 ve sonrası için 1 değerini alan bir kukla değişken tanımlanabilmektedir.  $D_t=1$  kırılma sonrası için olan 2. parça ve  $D_t=0$  kırılma öncesi dönem olan 1. parça olarak tanımlanırsa, model (Yerdelen Tatoğlu, 2023, s. 222-223):

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 D_t + u_t$$

D'nin aldığı değere göre modeller:

$$D_t = 0 \text{ ise; } E(Y_t) = \beta_0 + \beta_1 X_{1t}$$

$$D_t = 1 \text{ ise; } E(Y_t) = (\beta_0 + \beta_2) + \beta_1 X_{1t}$$

## 2. Eğimde Yapısal Kırılmaya Sahip Model

Tek kırılma olduğu durumda,  $D_t=1$  kırılma sonrası için olan 2. parça ve  $D_t=0$  kırılma öncesi dönem olan 1. parça olarak tanımlanırsa, eğimde değişiklik olan model (Yerdelen Tatoğlu, 2023, s. 224-225):

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{1t} D_t + u_t$$

D'nin aldığı değere göre modeller:

$$D_t = 0 \text{ ise; } E(Y_t) = \beta_0 + \beta_1 X_{1t}$$

$$D_t = 1 \text{ ise; } E(Y_t) = \beta_0 + (\beta_1 + \beta_2) X_{1t}$$

## 3. Spline Fonksiyon (Sabitte ve Eğimde Yapısal Kırılmaya Sahip Model)

Tek kırılma olduğu durumda,  $D_1=1$  kırılma sonrası için olan 2. parça ve  $D_1=0$  kırılma öncesi dönem olan 1. parça olarak tanımlanırsa spline fonksiyon (Yerdelen Tatoğlu, 2023, s. 227):

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 D_t + \beta_3 D_t X_{1t} + u_t$$

D'nin aldığı değere göre modeller:

$$D_1=0 \text{ ise; } E(Y_t) = \beta_0 + \beta_1 X_{1t}$$

$$D_1=1 \text{ ise; } E(Y_t) = (\beta_0 + \beta_2) + (\beta_1 + \beta_3) X_{1t}$$

## Parçalı (Piece-Wise) Doğrusal Regresyon Modelleri

Spline fonksiyonda birleşme noktalarında karşılaşılan sıçramaları ortadan kaldırıp parçaları birleştirmek için oluşturulan modellere parçalı (piece-wise) regresyon modelleri denilmektedir. Tek kırılmanın olduğu durumda iki parça için iki kukla değişken kullanılmakta ve model aşağıdaki gibi tanımlanmaktadır (Yerdelen Tatoğlu, 2023, s. 229):

$$Y_t = [\beta_{10} + \beta_{11}(X_t - X_0)]D_{1t} + [\beta_{20} + \beta_{21}(X_t - X_1)]D_{2t} + u_t$$

Burada  $D_1$  1. parça ve  $D_2$  2. parça için 1 değerini alacak şekilde oluşturulduğunda,

$$D_1=1 \text{ ve } D_2=0 \text{ olduğunda; } E(Y_t) = \beta_{10} + \beta_{11}(X_t - X_0)$$

$$D_1=0 \text{ ve } D_2=1 \text{ olduğunda; } E(Y_t) = \beta_{20} + \beta_{21}(X_t - X_1)$$

Bu bilgiler modelde yerine yerleştirildiğinde:

$$Y_t = \beta_{10} + \beta_{11}[X_t D_{1t} + X_1 D_{2t} - X_0] + \beta_{21}(X_t - X_1) D_{2t} + u_t$$

$X_1^* = X_t D_{1t} + X_1 D_{2t} - X_0$  ve  $X_2^* = (X_t - X_1) D_{2t}$  tanımlamaları ile model aşağıdaki gibi ifade edilebilmektedir (Yerdelen Tatoğlu, 2023, s. 229-230):

$$Y_t = \beta_{10} + \beta_{11} X_1^* + \beta_{21} X_2^* + u_t$$

## Eşik Değerli Regresyon

Eşik değerli regresyon modellerinde, bağımsız değişkenin aldığı değere göre bir eşik belirlenmekte ve model, bu eşik değer altında ve üstünde farklı katsayılarla tahmin edilmektedir. Eşik değer genelde içsel olarak belirlenmektedir (Yerdelen Tatoğlu, 2023, s. 232).

## Eşik Değerin Belirlenmesi

Hansen (1999) kalıntı kareler toplamını minimum yapan eşik değer aşağıdaki gibi elde edilebileceğini önermiştir:

$$\hat{\gamma} = \arg\min_{\gamma} S_1(\gamma)$$

Burada,  $S_1(\gamma)$  n kalıntı kareler toplamıdır.

Tek eşik değere sahip olan modellerde eşik değer istatistiksel anlamlılığını sınamak için kurulan temel hipotez eşik etkisinin olmadığını dolayısıyla modelin doğrusal olduğunu göstermektedir. Alternatif hipoteze göre ise eşik etkisi vardır, model doğrusal değildir. F test istatistiği,  $S_0$  ve  $S_1$  sırasıyla doğrusal ve eşik etkisinin bulunduğu modellerden elde edilen kalıntı kareler toplamı olmak üzere aşağıdaki gibi hesaplanmaktadır (Yerdelen Tatoğlu, 2023, s. 232-233):

$$F = \frac{(S_0 - S_1)}{\hat{\sigma}^2}$$

Burada  $\hat{\sigma}^2$  eşik değerli modelin kalıntı varyansıdır. F istatistiği için asimptotik kritik değerler bulunamadığı için bootstrap yoluyla türetilmektedir.

Eşik etkisinin varlığı için LM ve LR testleri de kullanılabilmekte hem model hem de testler iki veya daha fazla eşik bulunduğuna da genişletilebilmektedir (bkz. Chan (1993) ve Hansen (1999)) (Yerdelen Tatoğlu, 2023, s. 233).

## Eşik Değerli Regresyon Modellerinin Tahmini

Eşik değerli regresyon modeli Hansen (1996, 2000) ve Bacon ve Watts (1971) tarafından önerilen regresyonlarla tahmin edilebilmektedir (Yerdelen Tatoğlu, 2023, s. 233).

## Hansen Eşik Değerli Regresyon

Hansen (1996, 2000) tek eşik değer olması durumunda iki rejimli aşağıdaki eşik regresyon (PTR) modelinden hareket etmiştir:

$$y_t = \beta'_1 x_t I(q_t \leq \gamma) + \beta'_2 x_t I(q_t > \gamma) + u_t \quad (t=1, \dots, T)$$

Yukarıdaki modelde  $y_t$  bağımlı ve  $x_t$  bağımsız değişkenleri ve  $u_t$  hata terimini temsil etmektedir ( $u_t \sim IID(0, \sigma^2)$ ).  $q_t$  eşik değişkenini,  $I(\cdot)$  gösterge fonksiyonunu ve  $\gamma$  eşik parametresini göstermektedir. Eşik değerli model:

$$y_t = \beta' x_t(\gamma) + u_t$$

Burada:

$$x_t(\gamma) = \begin{pmatrix} x_t I(q_t \leq \gamma) \\ x_t I(q_t > \gamma) \end{pmatrix} \text{ ve } \beta' = (\beta'_1 \beta'_2)$$

Model en küçük kareler yöntemi ile tahmin edilmektedir (Yerdelen Tatoğlu, 2023, s. 233-234).

## Hockey-Stick Regresyon

Bacon ve Watts (1971) tarafından önerilen yöntemde bağımsız değişkenin eşik değeri  $\gamma$  aşağıdaki gibi ifade edilmektedir (Yerdelen Tatoğlu, 2023, s. 235-236):

$$X_t < \gamma \text{ ise } Y_t = \alpha_0 + \alpha_1 X_t + u_{1t}$$

$$X_t \geq \gamma \text{ ise } Y_t = \beta_0 + \beta_1 X_t + u_{2t}$$

Tek eşikli modellerin yanı sıra, birden fazla eşik değer içeren modeller veya ani kırılmalar yerine yumuşak geçişlere izin veren regresyon modelleri de kullanılabilir. Ayrıca Tong (1983), eşik değişkeni olarak gecikmeli bağımlı değişkenin kullanıldığı ve bağımlı değişkenin kendi gecikmelerine bağlı olarak farklı rejimlerde davranış gösterdiği eşik otoregresyon (TAR) modelini önermiştir (Yerdelen Tatoğlu, 2023, s. 233).

## NİTEL TERCİH MODELLERİ, KESİKLİ VE SANSÜRLÜ REGRESYON MODELLERİ

### Nitel Tercih Modelleri

Bir ekonometrik modelde bağımsız değişkenler gibi bağımlı değişken de nitel değişken olabilmekte, 0-1, 1-2-3 gibi kategorik değerler alabilmektedir. Bu modellere *nitel tercih modelleri*, *nitel bağımlı değişkenli modeller*, *tercih modelleri* ya da *kategorik modeller* isimleri verilmektedir. Kukla bağımlı değişken iki değer

alıyorsa ikili nitel tercih modeli, ikiden fazla değer alıyorsa çoklu nitel tercih modeli söz konusu olmaktadır (Yerdelen Tatoğlu, 2023, s. 251).

İkili tercih modellerine örnek olarak bir araba alıp almama kararı, bir koleje gidip gitmeme kararı, bir sendikaya üye olup olmama kararı ya da bir işe girip girmeme kararı verilebilir. Örneğin araba alıp almama kararı (Yerdelen Tatoğlu, 2023, s. 251):

$$Y_i = \begin{cases} 1 & \text{biyey i araba alıyorsa} \\ 0 & \text{biyey i araba almıyorsa} \end{cases}$$

İkiden fazla değer alan çoklu tercih modellerinin ise sıralı ve sıralı olmayan tercih modelleri olmak üzere iki türü vardır. Bağımlı değişkenin aldığı değerlerin doğal bir sırası olmadığı modellere örnek olarak bireylerin işe ulaşımı verilebilir (Yerdelen Tatoğlu, 2023, s. 252):

$$Y_i = \begin{cases} 1 & \text{biyey i arabası ile tek başına işe gidiyorsa} \\ 2 & \text{biyey i arabası ile en az bir biyey ile beraber işe gidiyorsa} \\ 3 & \text{biyey i metro kullanarak işe gidiyorsa} \\ 4 & \text{biyey i otobüs kullanarak işe gidiyorsa} \\ 5 & \text{biyey i yürüyor ya da başka bir şekilde gidiyorsa} \end{cases}$$

Sıralı bağımlı değişkenli modellerde ise alternatifler doğal bir sıraya sahiptir. Örnek olarak bireylerin sağlık durumu verilebilir (Yerdelen Tatoğlu, 2023, s. 252):

$$Y_i = \begin{cases} 1 & \text{i. biyeyin sağlığı kötü ise} \\ 2 & \text{i. biyeyin sağlığı iyi ise} \\ 3 & \text{i. biyeyin sağlığı çok iyi ise} \end{cases}$$

Ardışık değişkenler, sıralı değişkenlerin özel bir türüdür. Bu değişkenlerde bir kategorinin gerçekleşmesi, kendisinden önceki kategorilerin gerçekleşmesini gerektirmektedir. Örnek olarak alınan eğitimin derecesi verilebilir (Yerdelen Tatoğlu, 2023, s. 252):

$$Y_i = \begin{cases} 1 & \text{biyey i ilköğretim diplomasına sahipse} \\ 2 & \text{biyey i lise diplomasına sahipse} \\ 3 & \text{biyey i üniversite diplomasına sahipse} \\ 4 & \text{biyey i yüksek lisans diplomasına sahipse} \end{cases}$$

İkili nitel tercih modelleri aşağıdaki gibi gösterilmektedir:

$$Y_i^* = \beta'X_i + u_i$$

Burada  $u_i \sim N(0, 1)$  ve  $u_i$  ile  $u_j$  ( $i \neq j$ ) birbirinden bağımsızdır.  $Y_i$  şöyle tanımlanmaktadır:

$$Y_i = I(Y_i^* > 0)$$

$I(\cdot)$ , parantez içindeki ifadenin gerçekleşmesi durumunda 1'e, gerçekleşmemesi durumunda ise sifıra eşit olan bir gösterge değişkendir. İkili tercih değişkenine ait olasılıklar şu şekilde gösterilebilir:

$$Prob(Y_i = 1) = F(\beta' X_i)$$

$$Prob(Y_i = 0) = 1 - F(\beta' X_i)$$

Burada  $\beta$  parametreleri,  $Y$ 'nin gerçekleşme olasılığı üzerinde  $X$ 'teki değişikliklerin etkisini yansıtmaktadır (Yerdelen Tatoğlu, 2023, s. 253).

$X$  veriyken  $Y$ 'nin koşullu beklenen değeri  $E(Y_i | X_i)$ , olayın gerçekleşme olasılığı  $P(Y_i = 1 | X_i)$  olarak tanımlanmaktadır.  $Y_i = 1$  iken olayın gerçekleşme olasılığı ( $P$ ) ve  $Y_i = 0$  iken olayın gerçekleşmeme olasılığı ( $1-P$ ) ise aşağıdaki eşitlikler yazılabilmektedir:

$$E(Y_i) = 0(1-P) + 1(P) = P$$

$$E(Y_i | X_i) = \beta' X = P$$

Bu modellerde bağımlı değişken iki değer aldığı için hata terimi de iki değer almakta ve binom dağılmaktadır. Fakat büyük örneklerde dağılımın normal dağılıma yakınsadığı varsayılmaktadır (Yerdelen Tatoğlu, 2023, s. 253).

Nitel bağımlı değişkenli modellerde bağımlı değişkenin olasılığı, farklı modeller kullanılarak ifade edilebilmektedir. Bu modeller doğrusal olasılık modeli, logit, probit ve tamamlayıcı log-log model gibidir (Yerdelen Tatoğlu, 2023, s. 252).

### **Doğrusal Olasılık Modeli**

Model doğrusal olarak düşünülüp en küçük kareler yöntemi ile tahmin edilebilir, fakat tahmin aşamasında ve sonrasında birçok sorunla karşılaşmaktadır. Bunlar (Yerdelen Tatoğlu, 2023, s. 253-254):

- $u$ ,  $\beta$ 'ya bağlı olarak heteroskedastiktir.
- Bağımlı değişkenin tahmini değeri 0-1 aralığının dışına çıkmaktadır.
- Modelin  $R^2$ 'si çok düşük olacaktır.

En küçük karelerin doğurduğu dezavantajlardan dolayı doğrusal olasılık modeli, genelleştirilmiş en küçük kareler ya da tartılı en küçük kareler yöntemleri ile tahmin edilebilmekte, fakat bu durumlarda da varyans negatif olabilmekte ve bağımlı değişkenin tahmini değeri 0-1 aralığında kısıtlanmayabilmektedir. Bu dezavantajlarından dolayı, ikili tercih modellerinin tahmininde doğrusal olasılık modeli pek tercih edilmemektedir (Yerdelen Tatoğlu, 2023, s. 253-254).

### Probit Model

Literatürde normit model ismiyle de bilinen ve S-şekilli bir eğri ile ifade edilebilen probit modelde kümülatif normal dağılım fonksiyonu kullanılmaktadır. Bu fonksiyon doğrusal değildir ve dönüşümlerle doğrusallaştırılması gereklidir (Yerdelen Tatoğlu, 2023, s. 255).

İkili değişken  $Y_i$  ile gözlenemeyen değişken  $Y_i^*$  arasında aşağıdaki gibi bir ilişki vardır:

$$Y_i^* > 0 \text{ ise } Y_i = 1$$

$$Y_i^* \leq 0 \text{ ise } Y_i = 0$$

$Y_i$ 'nin beklenen değeri:

$$\begin{aligned} E(Y_i) &= P_i = P(Y_i = 1) \\ &= P(Y_i^* > 0) = P(-u_i < \beta' X_i) \\ &= F(\beta' X_i) = F(Z_i) \end{aligned}$$

Murada,  $F(\cdot)$  standart normal kümülatif dağılım fonksiyonunu temsil etmekte ve şöyle gösterilmektedir:

$$P_i = F(Z_i) = F(\beta' X_i) = \int_{-\infty}^{\beta' X_i} f(z) dz$$

$f(z)$  ise  $[z \sim N(0, 1)]$ , yoğunluk fonksiyonunu ifade etmektedir. Bu fonksiyonu doğrusallaştırmak için, kümülatif normal yoğunluk fonksiyonunun tersi alınmaktadır:

$$Z_i = F^{-1}(P_i) = \beta' X_i$$

Probit modelin en çok olabilirlik yöntemi ile tahmini için log-olabilirlik fonksiyonu aşağıdaki gibi elde edilmektedir (Yerdelen Tatoğlu, 2023, s. 256):

$$\ln L = \sum w_i \ln \Phi(\beta' x_i) \sum w_i \ln(1 - \Phi(\beta' x_i))$$

## Logit Model

Logit modelde lojistik dağılım kullanılmakta ve aşağıdaki gibi ifade edilmektedir:

$$E(Y_i) = \frac{1}{1 + e^{-\beta' X_i}} \\ = \frac{e^{\beta' X_i}}{1 + e^{\beta' X_i}}$$

Burada  $E(Y_i) = P(Y_i=1)$ 'dir.  $\beta' X_i$ ,  $-\infty$  ve  $+\infty$  arasında ve  $F(\beta' X_i)$ , 0 ile 1 arasında değer almaktadır (Yerdelen Tatoğlu, 2023, s. 256).

$E(Y_i) = P_i$  eşitliği yazılıp,  $\beta' X_i$  için çözümlerse:

$$L_i = \ln \left[ \frac{P_i(Y_i=1)}{1 - P_i(Y_i=1)} \right] = \ln(e^{\beta' X_i}) = \beta' X_i$$

Burada  $L_i$  doğrusal, fakat olasılık  $P_i$  doğrusal değildir. Köşeli parantezin içindeki ifade  $[P_i/(1-P_i)]$  odds oranı olarak bilinmektedir ve  $Y_i=0$ 'a karşı  $Y_i=1$  olasılığının oranıdır. Eşitliğin sol tarafındaki tüm ifade ise, odds oranının logaritmasını ya da  $Y_i=1$  için olayın logitini ifade etmektedir (Yerdelen Tatoğlu, 2023, s. 257).

Eğer  $Y_i=1$  olasılığında,  $X_i$ 'deki bir değişikliğin etkisi ile ilgileniliyorsa marjinal etki şu şekilde bulunmaktadır:

$$\frac{\partial P_i}{\partial X_i} = \frac{\partial P_i}{\partial \log P_i(1 - P_i)} \times \frac{\partial \log P_i(1 - P_i)}{\partial X_i} \\ = \beta P_i(1 - P_i)$$

Logit modelin en çok olabilirlik yöntemi ile tahmini için log-olabilirlik fonksiyonu:

$$L = \prod_i \frac{e^{\beta' X_i}}{(1 + e^{\beta' X_i})} \prod_j \frac{1}{(1 + e^{\beta' X_j})}$$

Lojistik dağılım kuyrukları dışında normal dağılıma benzemekte, kuyrukları ise 7 serbestlik dereceli t dağılımını andırmaktadır (Yerdelen Tatoğlu, 2023, s. 257).

## Tamamlayıcı Log-Log Model

İkili tercih modellerinden probit ve logit modellerine alternatif bir model, literatürde gompit model olarak da bilinen tamamlayıcı log-log modeldir. Model şu şekilde ifade edilebilmektedir (Yerdelen Tatoğlu, 2023, s. 257):

$$\log(-\log(1 - P_i)) = \beta' X_i$$

Bu model genelde olayın gerçekleşme olasılığının çok küçük ya da çok büyük olduğu zamanlarda kullanılmaktadır ve logit ve probit modellerden farklı olarak, asimetrik bir dağılıma sahiptir.

Tamamlayıcı log-log modelin en çok olabilirlik yöntemi ile tahmini için log-olabilirlik fonksiyonu:

$$\ln L = \sum w_i \ln \Phi(\beta' X_i) + \sum w_i \ln(1 - \Phi(\beta' X_i))$$

Burada  $w_i$  isteğe bağlı tartıları göstermektedir (Yerdelen Tatoğlu, 2023, s. 257-258).

## Kesikli Regresyon Modeli

Kesikli regresyon modellerinde, bağımlı değişken belirli bir eşik değerin altında veya üstünde gözlemlenmemektedir. Bu nedenle bağımlı değişken yalnızca belirli bir aralıkta gözlemlenmekte ve model buna göre tahmin edilmektedir. Örneğin, bağımlı değişken ailelerin dayanıklı tüketim mallarına yaptığı harcamalar ise, hiç harcama yapmayan aileler için bağımlı değişken gözlenmezken, harcama yapan aileler için gerçekleşen harcama tutarı gözlenmektedir (Yerdelen Tatoğlu, 2023, s. 266).

$$Y_i^* = \beta' X_i + u_i$$

Kesikli model için bağımlı değişken şöyle tanımlanmaktadır:

$$\begin{array}{ll} Y_i^* \leq a \text{ ise, gözlem yok} & Y_i^* \leq a \text{ ise, } Y_i = Y_i^* \\ Y_i^* > a \text{ ise, } Y_i = Y_i^* & Y_i^* > a \text{ ise, gözlem yok} \end{array} \quad \text{ya da}$$

İlk durumda bağımlı değişken eşik değerin altında, ikinci durumda ise eşik değerin üstünde kesilmektedir. Kesilme sonucunda yalnızca belirlenen aralıkta yer alan gözlemler örnekleme kalmakta ve gözlenen bağımlı değişkenin dağılımı normal dağılımdan farklılaşmaktadır.

$Y_i^*$  değişkenin koşullu dağılımı aşağıdaki gibidir:

$$Y_i^* | X_i \sim N[\beta' X_i, \sigma^2]$$

Eşik değerin ( $a$ ) üzerinde olduğu gözlemler için  $Y_i^*$ 'nin koşullu beklenen değeri:

$$E[Y_i^* | Y_i^* > a] = \beta' X_i + \sigma \frac{\phi((a - \beta' X_i) / \sigma)}{1 - \Phi((a - \beta' X_i) / \sigma)}$$

$$E[Y_i | Y_i > a] = \beta'X + \sigma\lambda_i(\alpha_i)$$

şeklinde ifade edilmektedir. Burada,  $\alpha_i = (a - \beta'X_i) / \sigma$  'dır. Ters Mills oranı ise:

$$\lambda(\alpha_i) = \phi(\alpha_i) / (1 - \Phi(\alpha_i))$$

Modeldeki marjinal etkiler şu şekilde elde edilmektedir:

$$\begin{aligned} \frac{\partial E[Y_i | Y_i > a]}{\partial \lambda} &= \beta + \sigma \left( \frac{d\lambda_i}{d\alpha_i} \right) \frac{\partial \alpha_i}{\partial \lambda} \\ &= \beta + \sigma(\lambda_i^2 - \alpha_i \lambda_i) \left( -\frac{\beta}{\sigma} \right) \\ &= \beta(1 - \lambda_i^2 + \alpha_i \lambda_i) \\ &= \beta(1 - \delta(\alpha_i)) \end{aligned}$$

$Y_i > a$  iken varyans:

$$Var[Y_i | Y_i > a] = \sigma^2 (1 - \delta(\alpha_i))$$

Kesikli regresyon modellerin en çok olabilirlik ile tahmini için log-olabilirlik şu şekilde ifade edilebilmektedir (Yerdelen Tatoğlu, 2023, s. 266-267):

$$\ln L = -\frac{n}{2}(\ln(2\pi) + \ln \sigma^2) - \frac{1}{2\sigma^2} \sum_i (Y_i - \beta'X_i)^2 - \sum_i \ln \left[ 1 - \Phi \left( \frac{a - \beta'X_i}{\sigma} \right) \right] \quad (6.30)$$

## Tobit (Sansürlü Regresyon) Modeli

Sansürlü regresyon modellerinde, bağımlı değişkenin belirli bir eşik değerin altında veya üstünde kalan gözlemleri tamamen kaybolmamakta, eşik değerde toplanarak gözlenmektedir. Bu tür modellere “*sansürlü regresyon modeli*” veya “*Tobit modeli*” denilmektedir. Örneğin kapasiteyi aşan talep, asgari ücretin altında kalan ücretler veya belirli bir eşik altındaki işsizlik oranları, eşik değerde sansürlenerek modele dâhil edilebilmektedir (Yerdelen Tatoğlu, 2023, s. 269).

Tobit modeli için bağımlı değişken aşağıdaki eşitlikte tanımlanmaktadır:

$$\begin{array}{ll} Y_i^* \leq a \text{ ise, } Y_i = 0 & \text{ya da} \quad Y_i^* \leq a \text{ ise, } Y_i = Y_i^* \\ Y_i^* > a \text{ ise, } Y_i = Y_i^* & Y_i^* > a \text{ ise, } Y_i = 0 \end{array}$$

Sansürlü ve sansürsüz gözlemlerin hem ortalama hem de marjinal değerlerinde farklılıklar bulunmaktadır.  $Y_i^*$  ve  $Y$  için marjinal etkiler sırasıyla:

$$\frac{\partial E[Y_i^* | X_i]}{\partial X_i} = \beta$$

$$\frac{\partial E[Y_i | X_i]}{\partial X_i} = \beta \Phi\left(\frac{\beta' X_i}{\sigma}\right)$$

Bunlardan hangisinin kullanılacağı, tahminin amacına bağlı olarak belirlenebilmektedir (Yerdelen Tatoğlu, 2023, s. 270).

Tobit modelinin en çok olabilirlik ile tahmini için log-olabilirlik fonksiyonu:

$$\ln L = \sum_{Y_i} -\frac{1}{2} \left[ \ln(2\pi) + \ln \frac{1}{\sigma^2} + \frac{(Y_i - \beta' X_i)^2}{\sigma^2} \right] + \sum_{Y_i=0} \ln \left[ 1 - \Phi\left(\frac{\beta' X_i}{\sigma}\right) \right]$$

Burada verilen dağılım kesikli ve sürekli dağılımların bir karışımıdır ve bu nedenle olabilirlik fonksiyonunun standart olmayan bir türüdür. Bu eşitliğin iki bölümü sırasıyla, sansürlsüz ve sansürlü gözlemler için ilgili olasılıkları temsil etmektedir (Yerdelen Tatoğlu, 2023, s. 270-271).

## Tercih Modellerine İlişkin Test ve İstatistikler

Probit, logit, tamamlayıcı log-log, kesikli ve sansürlü regresyon modellerinde kısıtlamaların geçerliliğinin sınanması için Olabilirlik Oranı (LR) testi, açıklama gücünü ifade etmek için ise Pseudo-R<sup>2</sup> kullanılmaktadır (Yerdelen Tatoğlu, 2023, s. 272).

### Olabilirlik Oranı (LR) Testi

Eğim katsayılarının sıfıra eşit olduğu hipotezini sınamak için kullanılan LR testi, regresyondaki eğim katsayılarının tümünün sıfır olduğu hipotezini sınavan F testine benzeyen birleşik bir testtir. Logit model için kısıtsız ve kısıtlı modeller sırasıyla (Yerdelen Tatoğlu, 2023, s. 272):

$$\ln \left[ \frac{P(D_i = 1)}{1 - P(D_i = 1)} \right] = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i$$

$$\ln \left[ \frac{P(D_i = 1)}{1 - P(D_i = 1)} \right] = \beta_0 + u_i$$

Test, kısıtsız modelin kısıtlı modele göre performansını ölçmektedir. Öncelikle hem kısıtlı hem de kısıtsız modele ait olabilirlik fonksiyonu elde edilmekte ve sonra LR istatistiği aşağıdaki gibi hesaplanmaktadır:

$$LR = -2 \left[ \ln \hat{L}_R - \ln \hat{L}_{UR} \right]$$

burada,  $\hat{L}_R$  ve  $\hat{L}_{UR}$  sırasıyla kısıtlı ve kısıtsız modeller için log-olabilirlik fonksiyonlarıdır. LR test istatistiği, k serbestlik derecesi ile  $\chi^2$  dağılmaktadır. Hesaplanan değer tablo değerinden büyükse  $H_0$  hipotezi reddedilmekte ve modelin anlamlı olduğuna karar verilmektedir (Yerdelen Tatoğlu, 2023, s. 272-273).

### Modelin Açıklama Gücü

Doğrusal modellerde kullanılan  $R^2$ , tercih modellerinde olduğundan düşük çıkacağından uygun bir ölçüt değildir, bu modellerde pseudo- $R^2$  kullanılabilmektedir (Yerdelen Tatoğlu, 2023, s. 274). Çeşitli araştırmacılar tarafından türetilen Pseudo- $R^2$ 'ler vardır.

- Mc-Fadden'in önerdiği pseudo- $R^2$  (Yerdelen Tatoğlu, 2023, s. 274):

$$Pseudo - R^2 = 1 - \left( \frac{\log L_{UR}}{\log L_R} \right)$$

- Cox –Snell tarafından önerilen pseudo- $R^2$  (Yerdelen Tatoğlu, 2023, s. 274):

$$Pseudo - R^2 = 1 - \left( \frac{L_{UR}}{L_R} \right)^{2/n} = 1 - \exp(-LR/n)$$

- Estella tarafından önerilen pseudo- $R^2$  (Yerdelen Tatoğlu, 2023, s. 275):

$$Pseudo - R^2 = \left( 1 - \frac{\log L_{UR}}{\log L_R} \right)^{-(2/n) \log L_R}$$

- Cragg ve Uhler tarafından önerilen pseudo- $R^2$  (Yerdelen Tatoğlu, 2023, s. 275):

$$Pseudo - R^2 = \left( \frac{L_{UR}^{2/n} - L_R^{2/n}}{(1 - L_R^{2/n}) L_{UR}^{2/n}} \right)$$

Hepsinde  $L_{UR}$ : kısıtsız,  $L_R$ : ise sadece sabit terimi içeren kısıtlı modelin log-olabilirliğidir.

Bu pseudo- $R^2$ 'lerin hepsi birbirine yakın sonuçlar vermekle birlikte, hesaplanması daha kolay olduğu için Mc-Fadden ve Estella'nın pseudo- $R^2$ 'leri daha çok kullanılmaktadır. Pseudo- $R^2$  de,  $R^2$  gibi 0-1 arasında değer almaktadır (Yerdelen Tatoğlu, 2023, s. 275).

## MODEL SEÇİMİ VE SPESİFİKASYON HATALARI

### Model Seçimi

Alternatif modellerde bağımlı değişken aynı olmakla birlikte bağımsız değişkenlerin birbirlerinin alt kümesi olduğu modellere *yuvalanmış model* ismi verilmektedir. Yuvalanmamış modellerde ise bağımlı değişken ya da bağımsız değişkenler farklı ölçü birimine sahip olabilmektedir. Örneğin bir model doğrusalken diğer model yarı ya da tam logaritmik olabilmektedir. Model seçimi yuvalanmış ve yuvalanmamış modellerde farklı olmaktadır (Yerdelen Tatoğlu, 2023, s. 285-286).

### Yuvalanmış Modellerde Seçim

#### Düzeltilmiş Belirginlik Katsayısı

Belirginlik katsayısı ( $R^2$ ) bağımsız değişkeni fazla olan modellerde her zaman için daha yüksek olacağından model seçimlerinde kullanılacak iyi bir ölçüt değildir. Düzeltilmiş  $R^2$  ( $\bar{R}^2$ ) daha iyi bir ölçüttür:

$$\bar{R}^2 = 1 - \left(1 - R^2\right) \frac{n-1}{n-k}$$

Burada  $n$  gözlem ve  $k$  parametre sayısıdır. Yuvalanmış modellerden  $\bar{R}^2$ 'si yüksek olan model seçilmektedir (Yerdelen Tatoğlu, 2023, s. 286-287).

#### Kalıntı Kareler Toplamı ve Kalıntı Varyansı

Yuvalanmış modellerden bağımlı değişkeni daha iyi açıklayan modelin kalıntı kareler toplamı ve kalıntı varyansı daha küçük olacaktır. Bu nedenle yuvalanmış olan modellerden kalıntı varyansı düşük olan model seçilmektedir. Kalıntı varyansı aşağıdaki gibidir (Yerdelen Tatoğlu, 2023, s. 288):

$$\hat{\sigma}_u^2 = \frac{ESS}{n-k} \quad \left(ESS = \sum (Y_i - \hat{Y}_i)^2\right)$$

### R<sup>2</sup>'nin Maksimizasyonuna Dayanan Kriterler

Ortalama hata karesini minimize eden bu kriterler belirginlik katsayısını maksimize edeceklerdir. Amemiya'nın PC kriteri, Mallows'un Cp ve Hockins'in Sp kriteri bu başlık altında ele alınabilmektedir. Bu kriterler sırasıyla (Yerdelen Tatoğlu, 2023, s. 288-289):

$$PC = \sigma_{u,i}^2 \left(1 + \frac{k_i}{n}\right) = \left(\frac{n+k_i}{n-k_i}\right) (1 - R_i^2) \left(\frac{TSS}{n}\right)$$

$$Cp = \frac{(n - k_i) \sigma_{u,i}^2}{\sigma_u^2} + (2k_i - n) = \left( \frac{(n - k_i)(1 - \bar{R}_i^2)}{(1 - \bar{R}^2)} \right) + (2k_i - n)$$

$$Sp = \sigma_{u,i}^2 \left( 1 + \frac{k_i}{n - k_i - 1} \right)$$

Bu eşitliklerde  $\sigma_{u,i}^2$ ,  $R_i^2$ ,  $\bar{R}_i^2$  ve  $k_i$  sırasıyla ilgilenilen modelin kalıntı varyansı, belirginlik katsayısı, düzeltilmiş belirginlik katsayısı ve bağımsız değişken sayısıdır.  $n$  ise gözlem sayısıdır (Yerdelen Tatoğlu, 2023, s. 289). PC, Cp ve Sp kriterlerinde en küçük değere sahip model tercih edilmektedir.

### Bilgi Kriterleri

Bilgi kriterleri, yuvalanmış modeller arasından seçim yapmak amacıyla kullanılmakta olup, en küçük bilgi kriteri değerine sahip model tercih edilmektedir. En yaygın kullanılan bilgi kriterleri Akaike Bilgi Kriteri (AIC), Schwarz Bilgi Kriteri (SC) veya Bayesyen Bilgi Kriteri (BIC, SBC ya da SBIC) ile Hannan–Quinn Bilgi Kriteri (HQ ya da HQIC)'dir (Yerdelen Tatoğlu, 2023, s. 290).

AIC, AICc, BIC ve HQ bilgi kriterleri aşağıdaki gibidir:

$$AIC = \ln \hat{\sigma}_i^2 + 2 \left( \frac{k_i}{n} \right)$$

$$AIC_c = \ln \hat{\sigma}_i^2 + \left( \frac{n + k_i}{n - k_i - 2} \right)$$

$$SC = \ln \hat{\sigma}_i^2 + \left( \frac{k_i \ln(n)}{n} \right)$$

$$HQ = \ln \hat{\sigma}_i^2 + \frac{2k_i \ln(\ln(n))}{n}$$

burada  $\hat{\sigma}_i^2$ , en çok olabilirlik yöntemi ile tahmin edilmektedir (Yerdelen Tatoğlu, 2023, s. 290). ( $\hat{\sigma}_u^2 = ESS/n$ ):

### Yuvalanmamış Modellerde Seçim

#### Cox Testi

Aşağıdaki yuvalanmamış modeller,

$$Y = \beta X + u_1$$

$$Y = \alpha Z + u_2$$

arasında seçim yapmak için kullanılan Cox testinde  $H_0$  hipotezi ilk modelin ve  $H_1$  hipotezi ikinci modelin doğru model olduğu şeklinde kurulmaktadır.  $H_0$  hipotezinin geçerli olduğu varsayımı altında Cox test istatistiği aşağıdaki gibi hesaplanmaktadır (Yerdelen Tatoğlu, 2023, s. 293):

$$q = \frac{C_{12}}{\sqrt{\text{Var}(C_{12})}}$$

Burada,

$$C_{12} = \frac{n}{2} \ln \left( \frac{\hat{\sigma}_2^2}{\hat{\sigma}_{21}^2} \right), \quad \hat{\sigma}_2^2 = \frac{u_2' u_2}{n}, \quad \hat{\sigma}_{21}^2 = \hat{\sigma}_1^2 + \frac{1}{n} \hat{\beta}' X' [I - Z(Z'Z)^{-1} Z'] X$$

$\hat{\sigma}_1^2$  de benzer şekilde  $\hat{\sigma}_1^2 = u_1' u_1 / n$  'dir.  $\sqrt{\text{Var}(C_{12})}$  ise:

$$\sqrt{\text{Var}(C_{12})} = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_{21}^2} \hat{\beta}' X' [I - Z(Z'Z)^{-1} Z'] [I - X(X'X)^{-1} X'] [I - Z(Z'Z)^{-1} Z'] X \hat{\beta}$$

q test istatistiği normal dağılım tablosu ile karşılaştırılmakta  $q > Z_{\alpha/2}$  ise  $H_1$ ,  $q < Z_{\alpha/2}$  ise  $H_0$  hipotezi reddedilmektedir (Yerdelen Tatoğlu, 2023, s. 293-294).

## J Testi

Davidson ve MacKinnon J testinde yuvalanmamış modeller bir arada yazılmakta:

$$Y = (1-\lambda)\beta X + \lambda\alpha Z + u$$

$\lambda=0$  ise birinci model,  $\lambda=1$  ise ikinci model seçilmektedir. Öncelikle  $Y = \alpha Z + u_2$  modeli tahmin edilmekte ve bağımlı değişkenin tahmini değerleri elde edilmektedir. Daha sonra aşağıdaki model tahmin edilmektedir:

$$Y = (1-\lambda)\beta X + \lambda(\hat{Y}) + u$$

Bu modelde  $\lambda$  parametresinin anlamlılığı t testi kullanılarak sınanmaktadır. t istatistiği tablo değeri ile karşılaştırıldığında eğer  $t > t_{\alpha/2}$  ise  $H_1$ ,  $q < t_{\alpha/2}$  ise  $H_0$  hipotezi reddedilmektedir (Yerdelen Tatoğlu, 2023, s. 294).

## Fonksiyonel Şeklin Belirlenmesi

### Bera-McAleer BM Testi

Bera-McAleer (1987) yuvalanmamış modellerin karşılaştırması için bir test önermiştir. Aşağıdaki gibi fonksiyonel şekli farklı olan modeller ele alındığında:

$$Y = \beta_0 + \beta_1 X + u_1$$

$$\log Y = \beta_0 + \beta_1 X + u_2$$

Öncelikle her iki model ayrı ayrı tahmin edilip bağımlı değişkenin tahmini değerleri elde edilir, sırasıyla  $\hat{Y}$  ve  $\log Y$ . Daha sonra  $\hat{Y}$  'nin logaritması ve  $\log Y$  'nin antilogaritması alınır. İkinci aşamada aşağıdaki modeller tahmin edilir:

$$\log(\hat{Y}) = \beta_0 + \beta_1 X + \varepsilon_1$$

$$e^{\log Y} = \beta_0 + \beta_1 X + \varepsilon_2$$

Daha sonra bu iki modelden elde edilen kalıntılar yukarıdaki modellere bağımsız değişkenler olarak eklenip modeller tekrar tahmin edilmektedir:

$$Y = \beta_0 + \beta_1 X + \theta_1 \hat{\varepsilon}_1 + v_1$$

$$\log Y = \beta_0 + \beta_1 X + \theta_2 \hat{\varepsilon}_2 + v_2$$

Bu iki modelde sırasıyla  $H_0: \theta_1 = 0$  ve  $H_0: \theta_2 = 0$  hipotezleri t testi yardımıyla sınanmakta, hangi hipotez anlamsızsa o model seçilmektedir (Yerdelen Tatoğlu, 2023, s. 296).

### PE Testi

PE testinin ilk aşaması BM testi ile aynıdır.  $\hat{Y}$  ve  $\log Y$  elde edildikten sonra ikinci aşamada aşağıdaki modeller tahmin edilmektedir:

$$Y = \beta_0 + \beta_1 X + \theta_1 (\hat{Y} - e^{\log Y}) + v_1$$

$$\log Y = \beta_0 + \beta_1 X + \theta_2 (\log Y - \log(\hat{Y})) + v_2$$

Bu iki modelde  $H_0: \theta_1 = 0$  ve  $H_0: \theta_2 = 0$  hipotezleri t testi yardımıyla sınanmakta, hangi hipotez anlamsızsa o model seçilmektedir (Yerdelen Tatoğlu, 2023, s. 296).

### Box-Cox Testi

Farklı fonksiyonel biçime sahip olan değişkenlerin aşağıdaki gibi Box-Cox dönüşümü ile doğrusallaştırılması sonucunda değişkenler standart hale getirilebilmekte ve karşılaştırılabilmektedir. Box-Cox dönüşümü:

$$\frac{Y^\lambda - 1}{\lambda}$$

Aşağıdaki model ele alındığında,

$$Y^{\lambda_1} = \beta_0 + \beta_1 X^{\lambda_2} + u$$

Burada  $\lambda_1$  ve  $\lambda_2$ 'nin alacağı değerlere bağlı olarak model farklı fonksiyonel şekillere sahip olacaktır.

$\lambda_1 = \lambda_2 = 0$  ise tam logaritmik model

$\lambda_1 = \lambda_2 = 1$  ise doğrusal model

$\lambda_1 = 1, \lambda_2 = 0$  ise doğrusal logaritmik model

$\lambda_1 = 0, \lambda_2 = 1$  ise logaritmik doğrusal model

$\lambda_1 = 1, \lambda_2 = -1$  ise ters model

$\lambda_1 = 0, \lambda_2 = -1$  ise logaritmik ters model (Yerdelen Tatoğlu, 2023, s. 297).

## Spesifikasyon Hataları

Spesifikasyon hatası, ekonometrik modelin kurulması aşamasında yapılan modelleme hatalarını ifade etmektedir. Bu hatalar ölçme hataları ve tanımlama hataları olmak üzere iki grupta incelenmektedir. Ölçme hataları bağımlı ve/veya bağımsız değişkenlerde ortaya çıkarken; tanımlama hataları modele gerekli bağımsız değişkenlerin dâhil edilmemesi, gereksiz bağımsız değişkenlerin dâhil edilmesi veya fonksiyonel şeklin yanlış belirlenmesi şeklinde ortaya çıkmaktadır (Yerdelen Tatoğlu, 2023, s. 299)

### Ölçme Hataları

#### Bağımsız Değişkende Ölçme Hatası Olması Durumu

Bağımsız değişkende ölçme hatası olduğu durumda parametre tahminleri sapmalı olmaktadır (Yerdelen Tatoğlu, 2023, s. 300).

#### Bağımlı Değişkende Ölçme Hatası Olması Durumu

Bağımlı değişkende ölçme hatası olduğu durumda en küçük kareler tahminçileri sapmasız olmasına rağmen etkin değildir (Yerdelen Tatoğlu, 2023, s. 300).

#### Bağımlı ve Bağımsız Değişkenlerde Ölçme Hatası Olması Durumu

Hem bağımlı hem de bağımsız değişkenlerde ölçme hatası olduğu durumda en küçük kareler tahminçileri sapmalıdır (Yerdelen Tatoğlu, 2023, s. 300-301).

### Tanımlama Hataları

#### Gerekli Bağımsız Değişkenin Modelde Yer Almaması

Önemli bir bağımsız değişkenin modelde yer almaması, çoğunlukla kuramsal çerçevenin yeterince oluşturulamaması veya literatürün yeterince incelenmemesi

nedeniyle ortaya çıkmaktadır. Bunun yanında, bazı değişkenler teorik olarak önemli olmakla birlikte ölçülememesi veya verilerine ulaşılamaması nedeniyle modele dâhil edilemeyebilmektedir. Nedeni ne olursa olsun, önemli bir bağımsız değişkenin model dışında bırakılması parametre tahminlerinde sapmaya ve model sonuçlarının güvenilirliğinin azalmasına yol açmaktadır (Yerdelen Tatoğlu, 2023, s. 301).

Doğru model aşağıdaki gibi iken,

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i$$

modelin eksik değişkenle tahmin edilmesi durumunda,

$$Y_i = \beta_0 + \beta_1 X_{1i} + v_i \quad v_i = \beta_2 X_{2i} + u_i$$

hata terimi modelden dışlanan değişkeni de içerecektir. Burada iki durum ortaya çıkabilmektedir:

- Dışlanan değişken ile modelde yer alan bağımsız değişkenler arasında ilişki bulunmuyorsa ( $X_{1i}$  dışsalsa), parametre tahminleri sapmasız olmakla birlikte etkinlik azalabilmektedir.
- Dışlanan değişken ile modelde yer alan bağımsız değişkenler arasında ilişki bulunuyorsa ( $X_{1i}$  içselse), parametre tahminleri sapmalı ve tutarsız olmaktadır. Sapmanın yönü, dışlanan değişkenin katsayısının işareti ile modelde yer alan bağımsız değişkenlerle olan korelasyonun yönüne bağlı olarak değişmektedir (Yerdelen Tatoğlu, 2023, s. 301–302).

### **Gereksiz Bağımsız Değişkenin Modelde Yer Alması**

Gerekli bir değişkenin modelde yer almaması gibi gereksiz değişkenlerin de modelde yer alması istenmeyen bir durumdur. Gereksiz değişkenlere modelde yer vermek serbestlik derecesi kaybı yaratmaktadır (Yerdelen Tatoğlu, 2023, s. 302).

### **Fonksiyonel Şeklin Hatalı Belirlenmesi**

Ekonometrik modelin matematiksel formunun yanlış seçilmesi de önemli bir problem olarak karşımıza çıkmaktadır. Bu durumda en küçük kareler tahmincileri sapmalıdır ve etkin değildir (Yerdelen Tatoğlu, 2023, s. 303).

### **Spesifikasyon Hatalarının Testi**

#### **Gerekli ve Gereksiz Değişkenlerin F Testi (Kısıtlı F Testi)**

Bir değişkenin modelde yer almama kararı kısıtlı F testi ile verilebilmektedir. Örneğin kısıtlı model (R),

$$Y_i = \beta_0 + \beta_1 X_{1i} + u_i$$

ve kısıtsız model (UR),

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i$$

şeklinde iken ilk model ikincinin içerisine yuvalanmıştır (Yerdelen Tatoğlu, 2023, s. 303). F testinde ilk önce her iki model de en küçük kareler yöntemi ile tahmin edilmekte ve  $R^2$  değerleri hesaplanmaktadır. F istatistiği:

$$F = \frac{(R_{UR}^2 - R_R^2)/h}{(1 - R_R^2)/(n - k)}$$

Burada  $R_{UR}^2$ : kısıtsız ve  $R_R^2$ : kısıtlı modellerin  $R^2$ 'si, h: kısıtsız modelde yer almayan değişken sayısı (örnekte 1'dir) ve k: kısıtsız modeldeki parametre sayısıdır.

Temel hipotez  $H_0: \beta_2=0$  şeklinde kurulmakta ve F test istatistiği, h, n-k serbestlik derecesi ile F dağılımına uymaktadır. Hesaplanan değer tablo değerinden büyükse  $H_0$  hipotezi reddedilmektedir. Bu durumda X değişkeni gereklidir ve modelde yer almalıdır; aksi halde gereksizdir (Yerdelen Tatoğlu, 2023, s. 304).

### **Gerekli ve Gereksiz Değişkenlerin Olabilirlik Oranı (LR) Testi**

Bir değişkenin modelde yer almama kararına LR testi ile de verilebilmektedir. Kısıtlı model (R) ve kısıtsız model (UR) en çok olabilirlik yöntemi ile tahmin edilmekte ve log-olabilirlik değerleri hesaplanmaktadır. Daha sonra LR testi aşağıdaki gibi yapılmaktadır (Yerdelen Tatoğlu, 2023, s. 305):

$$LR = -2[l_R - l_{UR}]$$

Burada  $l_{UR}$ : kısıtsız ve  $l_R$ : kısıtlı modelin log-olabilirlik değerleridir. LR test istatistiği, h (kısıt sayısı) serbestlik derecesi ile  $\chi^2$  dağılımına uymaktadır. Hesaplanan değer tablo değerinden büyükse kısıtlı model doğrudur şeklinde kurulan  $H_0$  hipotezi reddedilmektedir. Bu durumda X değişkeni gereklidir ve modelde yer almaktadır (Yerdelen Tatoğlu, 2023, s. 306).

### **Ramsey Reset Testi**

Ramsey (1969), model spesifikasyonunun doğru olup olmadığının belirlenmesi RESET testini önermiştir, bu test literatürde spesifikasyon hatasının belirlenmesi için en çok kullanılan testtir (Yerdelen Tatoğlu, 2023, s. 308). Aşağıdaki model ele alındığında:

$$Y = X\beta + u$$

Burada  $X$ ,  $NT \times k$  boyutludur ve  $u$ 'nın sıfır ortalama ile normal dağıldığı varsayılmaktadır. Modelde spesifikasyon hatası varsa,  $E[u|X] = \xi \neq 0$ 'dır. Ramsey RESET testinde  $\xi$ ,  $Z\theta$  olarak düşünülmekte ve aşağıdaki modelde  $H_0: \theta$  hipotezi sınanmaktadır:

$$Y = X\beta + Z\theta + u$$

$b$ ,  $\beta$ 'nın en küçük kareler tahminçisi olmak üzere genelde  $Z = [(Xb)^2, (Xb)^3, \dots, (Xb)^{p+1}]$  olarak kullanılmaktadır; bu da  $Z = (\hat{Y}^2, \hat{Y}^3, \dots, \hat{Y}^{p+1})$ 'yi ifade etmektedir. Ramsey ve Gilbert (1972) ve Thursby (1989)  $p=1, 2, 3$  alınabileceğini söylemişlerdir. DeBenedictis ve Giles (1998) ise  $p=3$ 'ü önermişlerdir (Yerdelen Tatoğlu, 2023, s. 308).

Temel hipotez:

$$H_0: \theta = 0 \quad (\text{model spesifikasyon doğrudur})$$

şeklinde kurulmaktadır. İlk model kısıtlı, ikinci model kısıtsız model iken spesifikasyon hatasını sınamak için aşağıdaki gibi kısıtlı F testi yapılmaktadır:

$$F = \frac{(R_{UR}^2 - R_R^2)/p}{(1 - R_{UR}^2)/(NT - k - p)}$$

Burada  $R_{UR}^2$  kısıtsız modelin;  $R_R^2$  ise kısıtlı modelin belirginlik katsayısını ifade etmektedir. Hesaplanan F değeri  $p$ ,  $NT-k-p$  serbestlik dereceli F dağılımına uyumaktadır.  $H_0$  hipotezi reddedilirse, spesifikasyon hatasının olduğuna karar verilmektedir (Yerdelen Tatoğlu, 2023, s. 308).

### DeBenedictis ve Giles Reset Testleri

Ramsey RESET testinin gücü modele eklenen polinom terimlerine bağlıdır ve bazı durumlarda testin gücü düşük olabilmektedir. Bu **nedenle** DeBenedictis ve Giles (1998), Taylor polinomları yerine Fourier yaklaşımını kullanarak RESET testinin gücünü artıran modifiye bir RESET testi önermiştir (Yerdelen Tatoğlu, 2023, s. 310). DeBenedictis ve Giles (1998),  $Z = [\sin(w), \cos(w), \sin(2w), \cos(2w), \dots, \sin(p'w), \cos(p'w)]$  önermişlerdir (Yerdelen Tatoğlu, 2023, s. 310).  $w$ , aşağıdaki gibi doğrusal dönüşüme (Mitchell ve Onvural, 1996),

$$w = \pi \frac{2(Xb) - [(Xb)_{\max} - (Xb)_{\min}]}{(Xb)_{\max} - (Xb)_{\min}}$$

veya sinüs dönüşümüne (Box, 1966),

$$w = 2\pi \sin^2(Xb) - \pi$$

sahip olacak şekilde türetilmektedir. İlkinde RESET S, ikincisinde ise RESET L testi adını almaktadır. Her iki test de Milliken ve Graybill (1970) koşullarının sağlandığı varsayımıyla temel hipotez altında  $2p'$  ve  $(Nt-k-2p')$  serbestlik dereceleriyle F dağılımına sahiptir (Yerdelen Tatoğlu, 2023, s. 310).

### Fonksiyonel Form İçin White Lagrange Çarpımı Testi

Godfrey ve Wickens (1981) modelin fonksiyonel formunun doğru belirlenip belirlenmediğini Box-Cox dönüşümüne dayanan White Lagrange Çarpımı (LM) tipi bir test yardımıyla sınamayı önermişlerdir (Yerdelen Tatoğlu, 2023, s. 310).

$\lambda$ , skaler bir parametre iken  $X$ 'in Box-Cox dönüşümü aşağıdaki gibidir:

$$X(\lambda) = \frac{(X^\lambda - 1)}{\lambda}$$

Bu eşitlik  $\lambda=0$  için tanımlı değildir, fakat  $X>0$  iken:

$$\lim_{\lambda \rightarrow 0} X(\lambda) = \ln(X)$$

$X(0)$ ,  $\ln(X)$ 'e eşittir. Aşağıdaki regresyon modeli ele alındığında,

$$Y(\lambda) = \sum_{i=1}^{K_1} \beta_i X_i(\lambda) + \sum_{i=K_1+1}^K \beta_i Z_i + u$$

burada  $Z$  modelin değişkenlerini temsil etmektedir.  $\lambda=1$  için doğrusal ve  $\lambda=0$  için logaritmik doğrusal model elde edilmektedir. Doğrusal ve logaritmik doğrusal modeller için temel hipotezler sırasıyla  $H_0: \lambda=1$  ve  $H_0: \lambda=0$  şeklinde kurulmaktadır (Yerdelen Tatoğlu, 2023, s. 310-311).  $u$ 'nun  $N(0, \sigma^2)$  olduğu varsayımı altında logaritmik olabilirlik fonksiyonu aşağıdaki gibidir:

$$\log L(\theta) = \frac{(\ln 2\pi)}{2} - \frac{(\ln \sigma^2)}{2} + (\lambda - 1) \ln Y - \frac{\left[ Y(\lambda) - \sum_{i=1}^{K_1} \beta_i X_i(\lambda) - \sum_{i=K_1+1}^K \beta_i Z_i \right]^2}{2\sigma^2}$$

Burada  $\theta = [\beta', \sigma^2, \lambda]'$  'dir.  $(\lambda - 1) \ln(Y)$  ise,  $Y$ 'yi  $Y(\lambda)$  'ye dönüştüren fonksiyonun türevinin logaritmasıdır (Yerdelen Tatoğlu, 2023, s. 311).

$H_0: \lambda=1$  hipotezini sınamak için  $\theta$ 'nın kısıtlı en çok olabilirlik tahmini  $\left( \hat{\theta} = [\hat{\beta}', \hat{\sigma}^2, 1]' \right)$  elde edilmektedir. Burada,  $\hat{\beta}$  ve  $\hat{\sigma}^2$  doğrusal modelden elde edilmektedir.

$H_0: \lambda=0$  hipotezini sınamak için  $\theta$ 'nın kısıtlı en çok olabilirlik tahmini  $\left( \hat{\theta} = [\hat{\beta}', \hat{\sigma}^2, 0]' \right)$  elde edilmektedir. Burada,  $\hat{\beta}$  ve  $\hat{\sigma}^2$  logaritmik doğrusal modelden elde edilen tahminlerdir.

LM test istatistiği aşağıdaki gibi hesaplanmaktadır:

$$\begin{aligned} \xi_{LM} &= i'W(\hat{\theta}) \left[ W(\hat{\theta})' W(\hat{\theta}) \right]^{-1} W(\hat{\theta})' i \\ &= T \left[ i'W(\hat{\theta}) \left[ W(\hat{\theta})' W(\hat{\theta}) \right]^{-1} W(\hat{\theta})' i / i'i \right] \\ &= TR_{i, \bar{W}}^2 \end{aligned}$$

LM test istatistiği 1 serbestlik dereceli  $\chi^2$  dağılımına uymaktadır.  $H_0$  hipotezi reddedildiğinde, modelde fonksiyonel form hatası olduğundan bahsedilmektedir (Yerdelen Tatoğlu, 2023, s. 311-312).

## EŞANLI DENKLEM MODELLERİ VE GÖRÜNÜRDE İLİŞKİSİZ REGRESYON

### Eşanlı Denklem Modelleri

İktisadi değişkenler arasında bazen tek yönlü neden-sonuç ilişkisi yerine bağımlı değişken ile bazı bağımsız değişkenler arasında karşılıklı ilişki ortaya çıkmaktadır (Yerdelen Tatoğlu, 2023, s. 327). Örnek olarak aşağıdaki tüketim (C) modeli ve gelir (Y) özdeşliği verilebilir:

Eşanlı denklem sisteminde içsel ve dışsal değişkenler bulunmaktadır. *İçsel değişken* model içerisinde belirlenen, denklem sisteminde ortak olan değişken olarak tanımlanabilirken *dışsal değişken* model dışında belirlenen değişkendir.

Eşanlı denklem sisteminin iktisadi teoriyi yansıtan ilk haline *yapısal biçim*, tahmin edilebilmesi için dönüştürüldüğü biçime ise *daraltılmış biçim* denilmektedir. Daraltılmış biçimde, diğer değişkenler tarafından eşanlı olarak belirlenen değişkenler kümesi bir araya toplanmakta ve dolayısıyla her bir eşitlikte sol tarafta tek bir içsel değişken yer almaktadır (Yerdelen Tatoğlu, 2023, s. 328).

## Örnek:

$$\left. \begin{array}{l} C_t = \beta_1 + \beta_2 Y_t + u_t \\ Y_t = C_t + I_t (= S_t) \end{array} \right\} \text{yapısal biçim}$$

Yapısal biçimi daraltılmış biçime getirmek için birinci denklemde  $Y_t$ 'nin yerine ikinci denklemdeki eşiti konulmakta ve daha sonra aynı işlemler ikinci eşitlik için de yapılmaktadır (Yerdelen Tatoğlu, 2023, s. 328-329):

$$\left. \begin{array}{l} C_t = \frac{\beta_1}{1-\beta_2} + \frac{\beta_2}{1-\beta_2} I_t + \frac{1}{1-\beta_2} u_t \\ Y_t = \frac{\beta_1}{1-\beta_2} + \frac{1}{1-\beta_2} I_t + \frac{1}{1-\beta_2} u_t \end{array} \right\} \text{daraltılmış biçim}$$

## Eşanlı Denklem Modellerinin Tahmini

### En Küçük Kareler Yöntemi

Eşanlı denklem sisteminin her bir eşitliği ayrı ayrı en küçük kareler yöntemi ile tahmin edilebilmektedir. Ancak eşitliklerde yer alan içsel değişkenler hata terimi ile ilişkili olduğundan, en küçük kareler yöntemi sapmalı ve tutarsız parametre tahminleri vermektedir (Yerdelen Tatoğlu, 2023, s. 329).

### Tanımlama (Belirleme) Koşulları:

#### 1. Sayma Koşulu:

M tane denklemi olan bir eşanlı modelde bir denklemin tanımlanabilmesi için, o denklemde yer alan değişken sayısının tüm denklem sisteminde yer alan değişken sayısından farkı, sistemdeki toplam denklem sayısının (içsel değişken sayısının) bir eksiğinden az olmamalıdır (Yerdelen Tatoğlu, 2023, s. 331).

Sayma koşulu gerekli bir koşul olmakla birlikte yeterli değildir, sayma koşulu sağlansa bile bir denklem tanımlanamayabilir. Bir denklemin tanımlanabilmesi için rank koşulunu da sağlaması gerekmektedir.

#### Rank Koşulu

M tane içsel değişkeni olan bir modelde, bir denklem ancak modelin öteki denklemlerinde yer alan fakat incelenen denklemden dışlanmış olan değişkenlerin katsayılarından oluşturulabilen  $(M-1)*(M-1)$  boyutlu en az bir tane sıfır olmayan determinantı varsa tanımlanabilmektedir (Yerdelen Tatoğlu, 2023, s. 332).

## Dolaylı En Küçük Kareler Yöntemi

Dolaylı en küçük kareler yöntemi sadece tam tanımlanmış denklemlerin tahmin edilmesinde kullanılmaktadır. İlk aşamada denklemler daraltılmış biçime çevrilmekte ve daraltılmış biçimin her bir eşitliği en küçük kareler yöntemi ile tahmin edilmektedir. Daha sonra daraltılmış biçim parametrelerinden yapısal biçim parametreleri elde edilmektedir (Yerdelen Tatoğlu, 2023, s. 333).

Dolaylı en küçük kareler yönteminden elde edilen parametreler sapmasızdır. Fakat dolaylı en küçük kareler yöntemi, yapısal biçimden daraltılmış biçime geçişin özellikle bazı karmaşık modellerde zor olması ve geçiş yapılsa bile daraltılmış biçim parametrelerinden yapısal biçim parametrelerinin elde edilmesinde karşılaşılan güçlükler nedenleriyle uygulamada pek tercih edilmemektedir (Yerdelen Tatoğlu, 2023, s. 333-334).

## İki Aşamalı En Küçük Kareler Yöntemi

İki aşamalı en küçük kareler yöntemi tam tanımlanmış denklemlerde dolaylı en küçük kareler yöntemine alternatif olarak kullanılmakta ve onunla aynı sonucu vermektedir, ilaveten bu tahmin yöntemi aşırı tanımlanmış denklemlerin tahmini için de kullanılabilir. Araç değişkenler tahmincisi olarak da bilinmektedir. Öncelikle denklemler daraltılmış biçime çevrilmekte ve daraltılmış biçimin her bir eşitliği en küçük kareler yöntemi ile tahmin edilmektedir. İkinci aşamada ise, yapısal biçimdeki gerçek değerler yerine ilk aşamada elde edilen içsel değişkenin tahmini değerleri yerleştirilerek denklemler tekrar en küçük kareler yöntemi ile tahmin edilmektedir (Yerdelen Tatoğlu, 2023, s. 336).

Eşanlı denklem sisteminin birinci eşitliği ele alındığında:

$$Y_1 = Z_1 \delta_1 + u_1$$

Burada,  $Z_1 = [Y_1 \ X_1]$  ve  $\delta_1' = (\beta_1', \gamma_1')$  'dir.  $Y_1$  modelin sağ tarafında yer alan içsel değişkenleri ifade etmektedir,  $g_1$  sayıdadır ve  $X_1$  modelin dışsal değişkenlerini ifade etmektedir,  $k_1$  sayıdadır.  $X = [X_1, X_2]$ , sistemdeki tüm dışsal değişkenleri ifade etmektedir ve  $k_2$  sayıdadır. Araç değişkenler tahmincisi:

$$d_{IV} = (W'Z)^{-1} W'Y$$

Asimptotik varyans tahminci:

$$\text{asyvar}(d_{IV}) = \sigma^2 (W'Z)^{-1} (W'W)(Z'W)^{-1}$$

Burada  $\hat{Y}_1$ ,  $Y_1$ 'in araç değişken seti olmak üzere aşağıdaki eşitlikler vardır:

$$Z = Z_1 = [Y_1 \ X_1]$$

$$W = [\hat{Y}_1 \ X_1]$$

Araç değişkenler tahmincisi şöyle yazılabilmektedir:

$$\begin{bmatrix} \hat{Y}_1'Y_1 & \hat{Y}_1'X_1 \\ X_1'Y_1 & X_1'X_1 \end{bmatrix} \begin{bmatrix} b_{IV} \\ c_{IV} \end{bmatrix} = \begin{bmatrix} \hat{Y}_1'y \\ X_1'y \end{bmatrix}$$

İki aşamalı en küçük kareler tahminlerinin tutarlı olabilmesi için üç koşulun sağlanması gerekmektedir:

1.  $plim\left(\frac{1}{n}W'W\right)$  sonlu simetrik pozitif tanımlı bir matristir.

2.  $plim\left(\frac{1}{n}W'Z\right)$  sonlu tekil olmayan bir matristir.

$$3. \quad plim\left(\frac{1}{n}W'u\right) = \begin{bmatrix} plim\left(\frac{1}{n}\hat{Y}_1'u\right) \\ plim\left(\frac{1}{n}X_1'u\right) \end{bmatrix} = 0$$

İki aşamalı en küçük kareler yönteminden elde edilen parametreler bu üç koşulun sağlanması halinde sapmasızdır. Uygulanması kolay olduğundan uygulamada sıklıkla tercih edilmektedir (Yerdelen Tatoğlu, 2023, s. 337-338).

### Üç Aşamalı En Küçük Kareler Yöntemi

Üç aşamalı en küçük kareler yöntemi, tam ve aşırı tanımlanmış denklemlerin tahmininde kullanılmaktadır. İlk iki aşaması iki aşamalı en küçük kareler yöntemi ile aynı olup, son aşamada denklemlerin hata terimleri arasındaki kovaryansı dikkate alan genelleştirilmiş en küçük kareler yöntemi uygulanmaktadır. Böylece denklemler arasındaki hata terimi korelasyonu da tahmin sürecine dâhil edilmekte ve sistem bir bütün olarak tahmin edilmektedir. Bu nedenle üç aşamalı en küçük kareler yöntemi sistem tahmincisi olarak da bilinmektedir (Yerdelen Tatoğlu, 2023, s. 344).

Eşanlı denklem sisteminin i. eşitliği ele alındığında;

$$Y_i = Z_i\delta_i + u_i$$

burada  $Z_i = [Y_i \ X_i]$  ve  $\delta'_i = (\beta'_i, \gamma'_i)$  'dir.  $Y_i$  içsel ve  $X_i$  dışsal değişkenlerdir. Model  $X'$  ile sağdan çarpıldığında aşağıdaki eşitlik elde edilmektedir:

$$X'Y_i = X'Z_i\delta_i + X'u_i$$

Bu eşitlik için varyans kovaryans matrisi ( $E(u_i u'_i) = \sigma_{ii} I$  olması koşuluyla):

$$E(X'u_i u'_i X) = \sigma_{ii} X'X$$

Böylece genelleştirilmiş en küçük kareler tahmini aşağıdaki gibidir (Yerdelen Tatoğlu, 2023, s. 344).:

$$d_i = [Z'_i X (X'X)^{-1} X' Z_i]^{-1} Z'_i X (X'X)^{-1} X' Y_i$$

## Görünürde İlişkisiz Regresyon Modelleri

Bazen birden fazla denklemin ortak değişkeni bulunmamakta ve birbirinden bağımsız görünmektedir, bununla birlikte hata terimleri birbirleri ile ilişkili olabilmektedir. Bu durumda görünürde ilişkisiz regresyon söz konusu olmaktadır (Yerdelen Tatoğlu, 2023, s. 347).

$$Y_t = \beta_1 + \beta_2 L_t + u_t$$

$$X_t = \alpha_1 + \alpha_2 K_t + v_t$$

Yukarıdaki denklemlerde  $u_t$  ve  $v_t$  hata terimleri arasında korelasyon bulunması durumunda görünürde ilişkisiz regresyon modeli söz konusudur. Tahminde ilk aşamada her bir denklem en küçük kareler yöntemi ile tahmin edilerek kalıntılar elde edilmekte, daha sonra bu kalıntılar yardımıyla hata terimlerinin varyans-kovaryans matrisi tahmin edilmektedir:

$$\Omega = \begin{bmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{bmatrix}$$

Elde edilen varyans-kovaryans matrisi kullanılarak sistem, Genelleştirilmiş En Küçük Kareler yöntemi ile tahmin edilmektedir (Yerdelen Tatoğlu, 2023, s. 347).

$$\hat{\beta} = (X'\Omega^{-1}X)^{-1} (X'\Omega^{-1}Y)$$

$\Omega$  varyans-kovaryans matrisinde, denklemlerin hata terimleri arasında ilişki bulunmuyorsa kovaryans terimi  $\sigma_{uv} = 0$  olmaktadır. Bu durumda  $\Omega$ , köşegen dışı elemanları sıfır olan köşegen (dolayısıyla simetrik) bir varyans-kovaryans matrisine dönüşmektedir. Breusch-Pagan LM testinde temel hipotez, denklemlerin

hata terimleri arasındaki korelasyonun sıfır olduğu şeklinde kurulmaktadır (Yerdelen Tatoğlu, 2023, s. 348).

Lagrange Çarpanı (LM) test istatistiği aşağıdaki gibi hesaplanmaktadır:

$$\lambda_{LM} = T \sum_{i=1}^N \sum_{j=1}^N \hat{\rho}_{uv}^2$$

Burada  $\hat{\rho}_{uv}^2$  : u, v kalıntıları arasındaki korelasyon katsayısıdır:

$$\hat{\rho}_{uv} = \hat{\rho}_{vu} = \frac{\sum_{i=1}^N \hat{u}_i \hat{v}_i}{\left( \sum_{i=1}^N \hat{u}_i^2 \right)^{1/2} \left( \sum_{i=1}^N \hat{v}_i^2 \right)^{1/2}}$$

LM test istatistiği, M denklem sayısı olmak üzere M-1 serbestlik derecesi ile  $\chi^2$  dağılmaktadır. Temel hipotezi  $H_0: \rho_{ij}=0$  şeklinde kurulmakta ve test istatistiği, kritik değerden büyükse denklemlerin hata terimleri arasında korelasyon vardır sonucuna ulaşılmaktadır. Bu durumda her iki denklemin en küçük karelerle tahmin edilmesi uygun değildir, görünürde ilişkisiz regresyona ilişkin tahmin yöntemleri kullanılmalıdır (Yerdelen Tatoğlu, 2023, s. 348).

## DİNAMİK MODELLER

Çoğu iktisadi davranış geçmiş deneyimlerin etkisinde olduğundan bağımlı ya da bağımsız değişkenler arasında gecikmeli değişkenler yer alabilmektedir. Değişkenler arasındaki gecikmeli ilişkileri dikkate alan modellere *dinamik modeller* denilmektedir. Dinamik modeller iki başlık altında incelenir: otoregresif modeller ve dağıtılmış gecikmeli modeller. Bağımlı değişkenin gecikmeli değerlerini içeren modeller *otoregresif modeller*, bağımsız değişkenlerin gecikmeli değerlerini içeren modeller ise *dağıtılmış gecikmeli modeller* olarak adlandırılmaktadır (Yerdelen Tatoğlu, 2023, s. 367).

### Dağıtılmış Gecikmeli Modeller

Dağıtılmış gecikmeli modeller, bağımsız değişkenlerin cari (t) dönemdeki değerlerinin yanında gecikmeli değerlerinin de yer aldığı modellerdir. k gecikmeli dağıtılmış gecikmeli model, K maksimum gecikme olmak üzere aşağıdaki gibi ifade edilmektedir:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \dots + \beta_k X_{t-k} + u_t$$

$$Y_t = \alpha + \sum_{k=0}^K \beta_k X_{t-k} + u_t$$

Tek gecikmeli model ise aşağıdaki gibidir (Yerdelen Tatoğlu, 2023, s. 367).:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + u_t$$

## Dağıtılmış Gecikmeli Modellerin Tahmin Yöntemleri

### En Küçük Kareler Yöntemi

Dağıtılmış gecikmeli modeller en küçük kareler yöntemi ile tahmin edilebilmekle birlikte, gecikmeli bağımsız değişkenler arasındaki çoklu doğrusal bağlantı nedeniyle parametre tahminleri güvenilir olmamaktadır. Ayrıca gecikme uzunluğu arttıkça gözlem sayısı ve serbestlik derecesi azaldığından alternatif tahmin yöntemleri geliştirilmiştir (Yerdelen Tatoğlu, 2023, s. 368).

### Koyck Modeli

Koyck modelinde gecikme katsayılarının tümünün aynı işarete sahip olduğu ve geometrik dizi şeklinde azaldığı varsayılmaktadır. Dolayısıyla bağımsız değişkenin bugüne yakın değerlerinin daha eski değerlerine göre bağımlı değişkene etkisi daha fazladır (Yerdelen Tatoğlu, 2023, s. 369).

Aşağıdaki k gecikmeli dağıtılmış gecikmeli model ele alındığında:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \dots + u_t$$

Koyck modeli aşağıdaki gibi elde edilmektedir:

$$Y_t = \alpha + \beta_0 X_t + \lambda \beta_0 X_{t-1} + \lambda^2 \beta_0 X_{t-2} + \dots + u_t$$

$1 < \lambda < 1$ 'dir, böylece her  $\beta$  katsayısı kendisinden önceki  $\beta$  katsayısından daha küçük olmaktadır (Yerdelen Tatoğlu, 2023, s. 369-370).

Koyck dönüşümü için öncelikle modelinin bir gecikmesi alınmakta,  $\lambda$  ile çarpılmakta ve ilk modelden çıkarılmaktadır:

$$Y_t = \alpha(1 - \lambda) + \beta_0 X_t + \lambda Y_{t-1} + v_t \quad v_t = (u_t - \lambda u_{t-1})$$

Bu modele Koyck modeli denilmektedir.  $(1 - \lambda)$  uyarılama hızıdır,  $\lambda$  ne kadar büyüse gecikme katsayıları o kadar hızlı azalmaktadır. Tahmin en küçük kareler yöntemi ile yapılmakta ve tahmin edilen  $\alpha$ ,  $\beta_0$  ve  $\lambda$  parametrelerinden hareketle ana modelinin parametreleri hesaplanmaktadır. Gecikme sayısı istenilen noktada kesilebilmektedir (Yerdelen Tatoğlu, 2023, s. 370).

### Almon Modeli

**Almon modeli**, gecikme sayısının fazla olduğu durumlarda ortaya çıkan çoklu doğrusal bağlantı problemini azaltmak amacıyla geliştirilmiştir. Bu yöntemde

gecikmeli değişkenlerin doğrusal bileşimlerine ait parametreler en küçük kareler yöntemi ile tahmin edilmekte, gecikmeli modelin parametreleri ise dolaylı olarak elde edilmektedir (Yerdelen Tatoğlu, 2023, s. 372).

Almon modeli, gecikme etkilerinin polinomial bir yapı izlediği durumlar için uygundur. Bu nedenle polinomial dağıtılmış gecikmeli model olarak da bilinmektedir. Polinom derecesi, gecikme yapısındaki dönüm noktalarının sayısına göre belirlenmektedir. Dönüm noktası bulunmuyorsa birinci dereceden ( $m = 1$ ), bir dönüm noktası bulunuyorsa ikinci dereceden ( $m = 2$ ) polinom kullanılmaktadır (Yerdelen Tatoğlu, 2023, s. 372).

Polinomun derecesi  $m$  ise parametreler aşağıdaki gibidir:

$$\beta_i = \gamma_0 + \gamma_1 i + \gamma_2 i^2 + \gamma_3 i^3 + \dots + \gamma_m i^m \quad i=0,1,\dots,m$$

Örneğin polinom 2. dereceden ise model:

$$Y_t = \alpha + \gamma_0 X_t + (\gamma_0 + \gamma_1 + \gamma_2) X_{t-1} + (\gamma_0 + 2\gamma_1 + 4\gamma_2) X_{t-2} + u_t$$

$$Y_t = \alpha + \gamma_0 (X_t + X_{t-1} + X_{t-2}) + \gamma_1 (X_{t-1} + 2X_{t-2}) + \gamma_2 (X_{t-1} + 4X_{t-2}) + u_t$$

$$Y_t = \alpha + \gamma_0 Z_0 + \gamma_1 Z_1 + \gamma_2 Z_2 + u_t$$

Gecikme uzunluğu ve polinomun derecesi bilgi kriterleri yardımıyla belirlenebilmektedir (Yerdelen Tatoğlu, 2023, s. 372-373).

## Otoregresif Modeller

Bağımlı değişkenin gecikmeli değerlerinin modelin sağ tarafında yer aldığı modellere otoregresif model denilmektedir (Yerdelen Tatoğlu, 2023, s. 379).  $k$  gecikmeli otoregresif model:

$$Y_t = \alpha + \delta X_t + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_k Y_{t-k} + u_t$$

## Otoregresif Modellerin Tahmin Yöntemleri

### En Küçük Kareler Yöntemi

Otoregresif modeller en küçük kareler yöntemi ile tahmin edilebilmekle birlikte, hata terimlerindeki otokorelasyon ve gecikmeli bağımlı değişkenin hata terimi ile ilişkili olması nedeniyle dışsalılık varsayımı ihlal edilmekte, bu da sapmalı ve tutarsız parametre tahminlerine yol açmaktadır (Yerdelen Tatoğlu, 2023, s. 380).

## İki Aşamalı En Küçük Kareler Yöntemi

Otoregresif modellerde ortaya çıkan içsellik probleminin giderilebilmesi için araç değişkenler yöntemi olarak da bilinen iki aşamalı en küçük kareler (2EKK) yöntemi kullanılabilmektedir (Yerdelen Tatoğlu, 2023, s. 381).

İlk aşamada gecikmeli bağımlı değişken uygun araç değişkenler yardımıyla tahmin edilmektedir. Kullanılacak araç değişkenin hata terimi ile ilişkisiz, yerine geçtiği içsel değişken ile ise yüksek korelasyona sahip olması gerekmektedir. Uygulamada araç değişken olarak genellikle  $Y_{t-2}$  veya  $\Delta Y_{t-2}$  kullanılmaktadır. İkinci aşamada ise içsel değişken yerine ilk aşamada elde edilen tahmini değerler kullanılarak model tekrar en küçük kareler yöntemi ile tahmin edilmektedir (Yerdelen Tatoğlu, 2023, s. 381–382).

Tahmin sonrasında kullanılan araç değişkenlerin geçerliliği sınanmalı; gerekli koşullar sağlanmıyorsa farklı araç değişkenler tercih edilmelidir (Yerdelen Tatoğlu, 2023, s. 382).

## Otoregresif Modellerde Testler

### İçsellik İçin Hausman Testi

Otoregresif modellerde, gecikmeli bağımlı değişken nedeniyle ortaya çıkan içsellik probleminin sınanmasında Hausman testi kullanılmaktadır. Testte en küçük kareler ve iki aşamalı en küçük kareler tahmincileri karşılaştırılmaktadır. Temel hipotez altında dışsallık varsayımı sağlanmakta ve en küçük kareler tahmincisi geçerli olmaktadır. Hipotezin reddedilmesi durumunda ise içsellik bulunduğu kabul edilmekte ve modelin iki aşamalı en küçük kareler yöntemi ile tahmin edilmesi gerekmektedir (Yerdelen Tatoğlu, 2023, s. 383–384).

$$H_0: \text{Cov}(Z_t, u_t) = 0$$

$$H_1: \text{Cov}(Z_t, u_t) \neq 0$$

Hausman testi aşağıdaki gibi hesaplanmaktadır:

$$H = (\hat{\beta}_{TSLs} - \hat{\beta}_{OLS})' \left[ \text{Var}(\hat{\beta}_{TSLs}) - \text{Var}(\hat{\beta}_{OLS}) \right]^{-1} (\hat{\beta}_{TSLs} - \hat{\beta}_{OLS})$$

İki aşamalı en küçük kareler tahmininden sonra modelde otokorelasyon, heteroskedastisite, çoklu doğrusal bağlantı ve normal dağılım varsayımları ayrıca sınanmalıdır (Yerdelen Tatoğlu, 2023, s. 386).

### **İçsellik İçin Durbin Score ve Wu-Hausman Testleri**

Durbin Score ve Wu-Hausman testleri, araç değişkenlerin dışsallık varsayımını sağlayıp sağlamadığının sınanmasında kullanılmaktadır. Temel hipotez araç değişkenlerin dışsal olduğu şeklindedir. Hipotezin reddedilmesi durumunda araç değişkenlerin geçerli olmadığı kabul edilmekte ve farklı araç değişkenlerin kullanılması gerekmektedir (Yerdelen Tatoğlu, 2023, s. 385–386). (1974)), Hausman (1978)) testleri dışsallığın sınanması için kullanılmaktadır (Yerdelen Tatoğlu, 2023, s. 385-386).

### **Varsayımların Testleri**

İki ve üç aşamalı en küçük kareler tahmininden sonra modelde otokorelasyon, heteroskedastisite, çoklu doğrusal bağlantı ve normal dağılım varsayımları ayrıca sınanmalıdır (Yerdelen Tatoğlu, 2023, s. 386).