

**T. C.  
İSTANBUL ÜNİVERSİTESİ  
SOSYAL BİLİMLER ENSTİTÜSÜ  
İŞLETME ANABİLİM DALI  
SAYISAL YÖNTEMLER BİLİM DALI**

**DOKTORA TEZİ**

**İLİŞKİ ANALİZİNDE VARSAYIMLARDAN  
SAPMALARIN BELİRLENMESİ VE  
ÇÖZÜMLENMESİNE YÖNELİK BİR  
BİLGİSAYAR PROGRAMI GELİŞTİRİLMESİ**

**HAZIRLAYAN:  
SONA MARDİKYAN**

**ÖĞRENCİ NUMARASI:  
2502010024**

**TEZ DANIŞMANI:  
PROF. DR. NEYRAN ORHUNBİLGE**

**İSTANBUL-2005**

**İlişki Analizinde Varsayımlardan Sapmaların Belirlenmesi ve Çözümlemesine  
Yönelik bir Bilgisayar Programı Geliştirilmesi  
Hazırlayan: Sona Mardikyan**

**ÖZ**

Günümüzde ileriye yönelik tahmin yöntemleri arasında, en önemlisi regresyon analizidir. Bu yöntemde amaç, örnek regresyon fonksiyonunu temel alarak anakütle regresyon modelini olabildiğince doğru tahmin etmektir. Tahmin için önerilen çeşitli yöntemler arasında en yaygın olarak kullanılanı En Küçük Kareler yöntemidir. Ancak bu yöntemin sonuçları belirli varsayımlar ile geçerli olmakta, varsayımlardan sapmalar tahminleyenlerin özelliklerinden bazılarının geçerli olmamalarına neden olmaktadır. Regresyon analizini uygulamak ve varsayımlarının geçerliliklerini araştırmak üzere çeşitli istatistik programlar kullanılmaktadır. Ancak, sadece regresyon analizine ve analiz varsayımlarına odaklanmış bir paket program bulunmamaktadır. Bu ihtiyaçtan hareketle bu çalışmada, regresyon analizinin varsayımlarıyla birlikte doğrudan uygulanabileceği yeni bir bilgisayar programı Java programlama diliyle geliştirilmiştir.

Çalışmanın birinci bölümünde, basit ve çoklu regresyon modellerinin önemli özellikleri ile temel varsayımlarına kısaca değinilmekte, En Küçük Kareler yöntemi, korelasyon analizi, anakütle ve örnek verileriyle tahmin ve politika belirleme konuları ele alınmaktadır. İkinci bölümde, regresyon analizinin varsayımlarının saptanmasına yönelik testler ile geçerli olmamaları durumunda önerilen çözüm yöntemleri ayrıntılarıyla incelenmektedir. Üçüncü bölümde ise, geliştirilen regresyon analizi programının tüm işlemleri çeşitli örneklerle anlatılmakta ve doğrulukları SPSS programı ile ispatlanmaktadır. Çalışmanın uygulama kısmında, işgücü verimliliğini açıklayacak regresyon modeli ve modelin varsayımları yeni program ile hızlı ve kolay bir biçimde incelenmektedir.

# **The Development of a Computer Program for Determination and Solution of the Violations of the Assumptions in the Relationship Analysis**

**Prepared by: Sona Mardikyan**

## **ABSTRACT**

Nowadays, among the forecasting methods, the most important one is the regression analysis. In this method, the aim is to estimate the population regression model as much as possible taking as basis the sample regression function. Among the various methods proposed for estimation, the most popularly used is the Least Squares method. But, the results of this method are valid under certain assumptions and the invalidity of some properties of the estimators is caused by the violations of these assumptions. Various statistics programs are used to apply the regression analysis and investigate the validity of the assumptions. A packet program concentrated only on the regression analysis and the assumptions of this analysis does not exist; so, in order to be able to apply directly the regression analysis together with its assumptions, a new computer program has been developed with the Java programming language.

In the first part of this study, the important properties of the simple and multiple regression analysis and the basic assumptions are briefly explained, the Least Squares method, correlation analysis and the estimation and the determination of the politics based on the population and sample data are studied. In the second part, the tests for determination of the validity of the assumptions in regression analysis and the solution methods for violations of the assumptions are discussed in detail. In the third part, the operations of the new regression analysis program are explained using different examples and the results are justified with the SPSS program. In the application part of the study, the regression model to explain the labor productivity and the assumptions of this model are investigated in a fast and easy way with the new program.

## ÖNSÖZ

Bilindiği gibi regresyon analizi başta ekonomi ve işletme olmak üzere pek çok bilim dalında yaygın olarak kullanılan bir tahmin ve politika belirleme yöntemidir. Aralarında sebep-sonuç bağlantısı bulunan iki veya daha fazla değişken arasındaki ilişkiyi, matematik bir model ile karakterize eden bu istatistik yöntemde regresyon modeli oluşturulduktan sonra modelin yeterli olup olmadığının kontrolü analizin en önemli bölümüdür. Bu nedenle oluşturulan modelin doğru modele yeterli derecede yaklaştığını garanti etmek ve regresyon analizinin tüm varsayımlarını sağlayıp sağlamadığını kontrol etmek büyük önem taşımaktadır.

Çalışmanın amacı, regresyon analizinde, En Küçük Kareler yöntemiyle regresyon denkleminin oluşturulmasına ve varsayımların geçerliliklerinin araştırılmasına yönelik testler ile geçerli olmamaları durumunda önerilen yöntemlerle modelin yeniden tahmin edilmesine yarayan çalışmalara odaklanmış yeni bir bilgisayar programı geliştirmektir. Tüm istatistik paket programlarının seçenekleri arasında yer alan bu analiz için ayrı bir program geliştirmeyi amaçlamanın nedeni, günümüzde tahmin ve politika belirlemenin çok önemli olması ve analizin araştırmacılar ve karar vericiler tarafından çok sıklıkla kullanılmasıdır. Diğer programlarda farklı alt seçenekler kullanılmadan gerçekleştirilmeyen bu istatistik yöntemin tek bir program ile uygulanabilmesi, konu ile ilgili olanlara büyük kolaylık ve zaman tasarrufu sağlayacağı umut ve temenni edilmektedir.

Eğitimim süresince büyük bir zevkle izlediğim dersleriyle istatistiği bana sevdiren, tez çalışmamın her aşamasında beni motive eden, desteğini hiç bir zaman esirgemeyen değerli hocam Sayın Prof. Dr. Neyran ORHUNBİLGE'ye sonsuz teşekkürlerimi sunarım. Doktora programına başlamama neden olan ve akademik kariyer gelişimimde desteğini her zaman yanımda hissettiğim Sayın Prof. Dr. Meltem ÖZTURAN'a, Java programlama dilinin kullanımında karşılaştığım sorunlarda gösterdiği özverili yardımlarıyla bana destek olan, bilgi ve birikimlerini benimle paylaşarak yol gösteren Sayın Yrd. Doç. Dr. Osman DARCAN'a,

alıřmanın geliřme ařamalarında deneyimlerine bařvurduėum Sayın Yrd. Do. Dr.  
Bertan BADUR'a teřekkürü bir bor bilirim.

# İÇİNDEKİLER

ÖZ.....	iii
ABSTRACT.....	iv
ÖNSÖZ.....	v
TABLolar LİSTESİ.....	x
ŞEKİLLER LİSTESİ.....	xi
GİRİŞ.....	1
<b>1. REGRESYON ANALİZİ</b> .....	<b>3</b>
1.1. Tanımı, Amacı ve Önemi.....	3
1.2. Türleri.....	5
1.3. Basit ve Çoklu Doğrusal Regresyon Analizi.....	7
1.3.1. Regresyon Modelinin Önemli Özellikleri.....	10
1.3.2. En Küçük Kareler Yöntemi.....	12
1.3.2.1. En Küçük Kareler Yöntemiyle Elde Edilen Regresyon Katsayılarının Özellikleri.....	16
1.3.2.1.1. Tahminleyenin Dağılımı.....	16
1.3.2.1.2. Gauss-Markov Teorisi.....	17
1.3.2.2. Regresyon Denkleminin Standart Hatası.....	19
1.4. Korelasyon Analizi.....	21
1.4.1. Basit Doğrusal Korelasyon Analizi.....	22
1.4.2. Çoklu Doğrusal Korelasyon Analizi.....	23
1.4.2.1. Çoklu Korelasyon Katsayısı.....	23
1.4.2.2. Kısmi Korelasyon Katsayıları.....	26
1.5. Anakütle Verileri ile Tahmin ve Politika Belirlenmesi.....	27
1.6. Örnek Verileri ile Doğrusal Regresyon ve Korelasyon Analizi.....	29
1.6.1. Regresyon Analizinde Testler.....	31
1.6.1.1. Regresyon Katsayılarının Testleri.....	31
1.6.1.2. Korelasyon Katsayılarının Testleri.....	34
1.6.1.3. Kısmi Korelasyon Katsayılarının Testleri.....	36
1.6.2. Örnek Verileri ile Tahmin ve Politika Belirleme.....	37
<b>2. REGRESYON ANALİZİNİN VARSAYIMLARI, VARSAYIMLARDAN SAPMALARIN BELİRLENMESİ VE ÇÖZÜMLENMESİ</b> .....	<b>41</b>
2.1. Otokorelasyonun Belirlenmesi ve Çözümleme Yöntemleri.....	42
2.1.1. Otokorelasyonun Nedenleri.....	43
2.1.2. Otokorelasyonun En Küçük Kareler Yöntemiyle Elde Edilen Tahminleyenlere Etkisi.....	44
2.1.3. Otokorelasyon Katsayısı.....	46
2.1.4. Otokorelasyonun Belirlenmesi.....	47
2.1.4.1. Grafik Yöntem.....	47
2.1.4.2. Akış Sayısı Testi (Number of Runs Test).....	49
2.1.4.3. Durbin-Watson Testi.....	51
2.1.4.4. Farebrother Testi.....	55
2.1.4.5. Von Neumann Oranı Testi.....	56
2.1.5. Otokorelasyonun Ortadan Kaldırılması için Çözüm Yolları.....	58
2.1.5.1. Otokorelasyon Katsayısı $\rho$ 'nun Bilinmesi Durumu.....	58
2.1.5.1.1. Genelleştirilmiş Farklar Yöntemi.....	58

2.1.5.1.2. İlk Farklar Yöntemi.....	60
2.1.5.2. Otokorelasyon Katsayısı $\rho$ 'nun Bilinmemesi Durumu .....	61
2.1.5.2.1. İki Aşamalı Durbin Yöntemi.....	61
2.1.5.2.2. Durbin-Watson d İstatistiğine Dayanan Tahmin Yöntemi .....	62
2.1.5.2.3. Cochrane-Orcutt Yöntemi.....	62
2.2. Farklı Varyanslılığın Belirlenmesi ve Çözümleme Yöntemleri.....	64
2.2.1. Farklı Varyanslılığın Nedenleri.....	65
2.2.2. Farklı Varyanslılığın En Küçük Kareler Yöntemiyle Elde Edilen Tahminleyenlere Etkisi .....	66
2.2.3. Farklı Varyanslılığın Belirlenmesi.....	67
2.2.3.1. Grafik Yöntem .....	67
2.2.3.2. White Testi .....	68
2.2.3.3. Goldfeld-Quandt Testi .....	71
2.2.3.4. Breusch-Pagan/Godfrey Testi .....	73
2.2.3.5. Spearman Sıra Korelasyon Testi.....	77
2.2.3.6. Park Testi .....	78
2.2.3.7. Glejser Testi .....	79
2.2.4. Farklı Varyanslılık Durumunu Çözümleyen Yöntemler.....	81
2.2.4.1. Genelleştirilmiş En Küçük Kareler Yöntemi .....	81
2.2.4.2. Ağırlıklı En Küçük Kareler Yöntemi.....	84
2.2.4.3. $\sigma_i^2$ 'nin Bilinmesi Durumu .....	85
2.2.4.4. $\sigma_i^2$ 'nin Bilinmemesi Durumu .....	87
2.3. Normallik Varsayımından Sapmaların Belirlenmesi ve Çözümleme Yöntemleri.....	93
2.3.1. Normallik Varsayımından Sapmaların En Küçük Kareler Yöntemine Etkileri.....	94
2.3.2. Normallik Varsayımından Sapmaların Belirlenmesi .....	94
2.3.2.1. Grafik Yöntem .....	95
2.3.2.2. Momentlere Dayanan Asimetri ve Basıklık Ölçülerine göre Yapılan Test .....	96
2.3.2.3. Ki-Kare Uygunluk Testi.....	99
2.3.2.4. Jarque-Bera Testi .....	100
2.3.2.5. Kolmogorov-Smirnov Testi .....	101
2.3.2.6. Shapiro-Wilk Testi.....	102
2.3.3. Normallik Varsayımından Sapmaların Bulunması Durumunda Önerilen Çözüm Yöntemleri .....	103
2.3.3.1. Güçlü Regresyon (Robust Regression).....	104
2.4. Çoklu Doğrusal Bağlantının Belirlenmesi ve Çözümleme Yöntemleri.....	111
2.4.1. Çoklu Doğrusal Bağlantının Nedenleri.....	112
2.4.2. Çoklu Doğrusal Bağlantının Sonuçları .....	112
2.4.3. Çoklu Doğrusal Bağlantının Belirlenmesi .....	113
2.4.3.1. Varyans Artış Faktörleri (VIF-Variance Inflation Factors) .....	114
2.4.3.2. Yardımcı Regresyon Modelleri.....	115
2.4.3.3. Klein Kuralı.....	116
2.4.3.4. Durum Sayısı (Condition Number).....	117
2.4.4. Çoklu Doğrusal Bağlantı için Çözüm Yolları.....	117

2.4.4.1. Bağımsız Değişkenlerin Modelden Çıkarılması-Adım Adım Regresyon (Stepwise Regression).....	119
2.4.4.2. Tarafli Tahminleyenler Yöntemi (Ridge Regression).....	121
2.4.4.3. Asal Bileşenler Yöntemi (Principal Components Method) .....	128
<b>3. REGRESYON ANALİZİ İÇİN YENİ BİR BİLGİSAYAR PROGRAMI .</b>	<b>134</b>
3.1. JAVA Programlama Dili.....	135
3.2. Regresyon Analizi Programı .....	136
3.2.1. "File" Menüsü .....	137
3.2.2. "Model" Menüsü .....	140
3.2.3. "Regression" Menüsü .....	143
3.2.4. "Assumptions" Menüsü .....	149
3.2.4.1. "Multicollinearity" Alt Başlığı .....	150
3.2.4.2. "Autocorrelation" Alt Başlığı.....	154
3.2.4.3. "Heteroscedasticity" Alt Başlığı .....	158
3.2.4.4. "Normality" Alt Başlığı.....	162
<b>4. TÜRKİYE’NİN 500 BÜYÜK SANAYİ KURULUŞU VERİLERİYLE İŞGÜCÜ VERİMLİLİĞİ ÜZERİNE BİR UYGULAMA .....</b>	<b>165</b>
4.1. İşgücü Verimliliği .....	165
4.2. Çalışmanın Değişkenleri .....	167
4.3. Çalışmanın Verileri .....	169
4.4. Yeni Program ile Regresyon Analizinin Uygulanması.....	169
<b>SONUÇ.....</b>	<b>178</b>
<b>KAYNAKÇA .....</b>	<b>182</b>
<b>EKLER.....</b>	<b>190</b>
<b>ÖZGEÇMİŞ.....</b>	<b>235</b>



## TABLolar LİSTESİ

Tablo (1.1): Çoklu Regresyon Analizinde Varyans Analizi Tablosu .....	35
Tablo (2.1): $\sigma^2$ ile $E(y)$ Arasındaki İlişkiye Göre Önerilen Dönüşümler.....	91

## ŞEKİLLER LİSTESİ

Şekil (1.1) Serpilme Diyagramları .....	8
Şekil (2.1): Hata Terimlerinin Birbirleriyle İlişisini Gösteren Grafikler .....	48
Şekil (2.2): Standardize Hataların Dağılımını Gösteren Grafikler.....	49
Şekil (2.3): Farklı Varyanslılık .....	65
Şekil (2.4): $e^2$ ile $x$ Arasındaki İlişkiyi Gösteren Grafikler.....	68
Şekil (2.5): Histogramlar.....	95
Şekil (2.6): Normal Olasılık Grafikleri .....	96
Şekil (2.7): Ridge İzi.....	126
Şekil (2.8): İki Bağımsız Değişkenli Modelde Asal Bileşenler .....	131
Şekil (3.1): Regresyon Analizi Programı'nın Ana Ekranı .....	137
Şekil (3.2): "File" Menüsü .....	138
Şekil (3.3): "New"Seçeneği .....	138
Şekil (3.4): "Show" Seçeneği.....	139
Şekil (3.5): Tek Bağımsız Değişkenli Model için "Selection" Seçeneği.....	141
Şekil (3.6): Tek Bağımsız Değişkenli Modellerde "Y" için Sunulan Seçenekler....	141
Şekil (3.7): Tek Bağımsız Değişkenli Modellerde "X" için Sunulan Seçenekler....	142
Şekil (3.8): Çoklu Regresyon Modellerinde Model Belirleme .....	143
Şekil (3.9): Basit Regresyon Modellerinde "Correlations" Seçeneği .....	144
Şekil (3.10): Çoklu Regresyon Modellerinde "Correlations" Seçeneği.....	145
Şekil (3.11): (Ek:2)'deki Verilere Uygulanan Basit Doğrusal Regresyon Analizi Sonuçları .....	146
Şekil (3.12): (Ek:2)'deki Verilere Uygulanan Eğrisel Regresyon Analizi Sonuçları .....	147
Şekil (3.13): (Ek:2)'deki Verilere Uygulanan $\ln(x)$ Dönüşümü ile Ulaşılan Regresyon Analizi Sonuçları.....	148
Şekil (3.14): Çoklu Regresyon Modeli için Doğrusal Regresyon Analizi.....	148
Şekil (3.15): "Assumptions" Menüsü .....	149
Şekil (3.16): "Multicollinearity" Alt Başlığı.....	150
Şekil (3.17): Çoklu Regresyon Modelinde VIF Değerleri.....	151
Şekil (3.18): Çoklu Regresyon Modelinde "Stepwise" Yöntemi.....	152
Şekil (3.19): Ridge Regression Yönteminde "c" Katsayısının Belirlenmesi.....	153
Şekil (3.20): Ridge Regression Yöntemiyle Modelin Belirlenmesi.....	154
Şekil (3.21): "Autocorrelation" Alt Başlığı .....	154
Şekil (3.22): Durbin-Watson Testi.....	156
Şekil (3.23): Genelleştirilmiş Farklar Yöntemi.....	157
Şekil (3.24): Çoklu Regresyon Modelinde Durbin-Watson Test Sonuçları .....	157
Şekil (3.25): "Heteroscedasticity" Alt Başlığı .....	158
Şekil (3.26): Spearman Sıra Korelasyon Katsayıları .....	160
Şekil (3.27): Ağırlıklı En Küçük Kareler Yönteminde Ağırlık Değerlerinin Seçimi .....	161
Şekil (3.28): AEKK Yöntemiyle Elde Edilen Model.....	162
Şekil (3.29): Jarque-Bera Testi .....	164
Şekil (4.1): İşgücü Verimliliği Verileri için Model Seçimi .....	170
Şekil (4.2): İşgücü Verimliliği Verileri için Korelasyon Matrisi.....	170

Şekil (4.3): İşgücü Verimliliği Verileri ile Regresyon Analizi .....	171
Şekil (4.4): İşgücü Verimliliği Verileri için VIF Değerleri .....	172
Şekil (4.5): İşgücü Verimliliği Verilerine Adım Adım Regresyon Yönteminin Uygulanması .....	173
Şekil (4.6): İşgücü Verimliliği Verilerine Uygulanan Durbin-Watson Testi Sonuçları .....	174
Şekil (4.7): Genelleştirilmiş Farklar Yönteminin Aktif Olmaması Durumu .....	174
Şekil (4.8): İşgücü Verimliliği Verilerine Spearman Sıra Korelasyon Testinin Uygulanması .....	175
Şekil (4.9): İşgücü Verimliliği Verilerine Uygulanan Jarque-Bera Testinin Sonuçları .....	176

## GİRİŞ

İnsanođlu yüzyıllardır bilinçli veya bilinçsiz olarak gösterdiği çabalarla gelecekle ilgili tahminler yapmaya çalışmıştır. Bu çabaların amacı, gelecekte karşısına çıkabilecek olayları tahmin ederek planlarını bu yönde yapmak ve böylece kendini oluşabilecek zararlardan korumak olmuştur. Tarihte kahinler, geleceği tahmin etme becerisine sahip olduklarını savunmuşlar ve bu amaçla yönetici konumunda olan kişilere hizmet vermişlerdir. Fakat zaman, bu tahminlerin daha gerçekçi ve bilimsel yöntemlerle yapılması gerektiğini göstermiştir. İstatistik tahmin yöntemleri, bu gereksinimin sonucu olarak ortaya çıkmış ve özellikle İkinci Dünya savaşından sonra programlanabilen hesap makinelerinin ortaya atılmasıyla birlikte çok hızlı bir gelişme göstermiştir. 20. yüzyılın başlarında Galton tarafından ortaya atılan regresyon kavramı bu yöntemlerin temelini oluşturmaktadır.

Günümüzde tüm sosyal ve fen bilimlerinde yaygın olarak kullanılan ilişki analizi veya regresyon analizi, aralarında sebep-sonuç ilişkisi bulunan bir bağımlı değişkenle bir ya da birden fazla bağımsız değişken arasındaki ilişkiyi açıklama ve değerlendirmeyi amaçlamaktadır. Bilindiği gibi analizin geliştirilme nedeni, tesadüfi değişken olan bağımlı değişkenin anakütle ortalamasını, tekrarlı örneklemede sabit değerler aldığı varsayılan bağımsız değişkenler ile tahmin etmek (Regresyon Analizi) veya hedef değişken olan bağımlı değişken üzerinde politika belirlemek isteyen karar vericilere bu değişkenin hangi değişkenlere bağlı olduğunu (Korelasyon Analizi) saptama olanağı sağlamaktır. Yapılan tahminin ve belirlenen politikanın başarısı, seçilen modelin uygunluğu ve analizin varsayımlarının geçerli olmasına bağlıdır.

İleri seviyedeki birçok araştırmayı yapmaya olanak sağlayan istatistik paket programları bu önemli ve uygulanması güç olan istatistik yöntemin uygulanmasında araştırmacılara büyük kolaylık sağlamaktadır. Bunlardan SPSS, SAS, MINITAB, STATA gibi istatistik paket programları en fazla kullanılanlarından birkaçıdır. Bu programlarda regresyon analizinin temel hesaplamaları yapılabilmekte, ancak,

regresyon modelinin varsayımlarının geçerliliğine yönelik testler ve geçersizlik durumunda kullanılması gerekli çözüm yolları tek bir seçenek altında yer almamakta, farklı alt seçeneklere başvurmak gerekmektedir.

Çalışmanın amacı, hemen hemen her istatistik paket programında yer alan regresyon ve korelasyon analizini tek bir programla bir bütün olarak gerçekleştirme olanağı sağlamaktır. Çalışmanın, 1. Bölümünde basit ve çoklu regresyon analizinde En küçük Kareler yöntemi ele alınmakta daha sonra 2. Bölümde regresyon analizinin temel varsayımları, bu varsayımdan sapmaların belirlenmesinde en yaygın olarak kullanılan testler ve varsayımların geçerli olmamaları durumunda önerilen çözüm yöntemleri ayrıntılarıyla ele alınmaktadır. 3. Bölümde ise Java programlama diliyle, En Küçük Kareler yöntemiyle regresyon denkleminin oluşturulması ve anakütle için modelin geçerliliğinin araştırılmasında kullanılan testler, yöntemin varsayımlarının geçerliliklerinin saptanması, varsayımlardan sapma durumunda uygulanabilecek yöntemleri içeren bir bilgisayar programı geliştirilmektedir. Değişik örnekler ile program test edilmekte, sonuçlar SPSS paket programının sonuçlarıyla karşılaştırılarak aynı sonuçların elde edilip edilmediği araştırılmaktadır.

Çalışmanın uygulama aşaması olan 4. Bölümde Türk İmalat Sanayiinde işgücü verimliliğini etkileyen faktörleri araştırmak üzere, Sanayi Odası'nın Türkiye'nin En Büyük 500 Kuruluşu yayınlarından, Capital dergisinin benzer bir çalışmasındaki bazı sonuçlarından ve İstanbul Menkul Kıymetler Borsası'nın Web sayfasından ulaşılan bilgilerinden yararlanılarak oluşturulan 60 sanayi kuruluşu verileri ile regresyon analizi uygulanmaktadır. Geliştirilen yeni programla analiz gerçekleştirilmiş, tüm varsayımların geçerliliği de araştırılarak programın hızlı ve kolay işleyişi bir kez daha kontrol edilmiştir.

# 1. REGRESYON ANALİZİ

## 1.1. Tanımı, Amacı ve Önemi

Gelecekle ilgili tahmin ve planlar yapabilmek uzun yıllardan bu yana kazanılmak istenen becerilerden biridir. Bu amaçla, istatistik ve ekonomi alanlarında çalışan araştırmacılar, çeşitli alanlarda belirsizlikleri ölçmek ve gelecekle ilgili tahminler yapabilmek için bilimsel tahmin yöntemleri geliştirmişlerdir. *Regresyon Analizi* ve *Zaman Serileri Analizi* başlıkları altında toplanabilen bu analizlerden regresyon analizi değişik ölçüm ve değerler için tahmin ve politika belirlemek için kullanılan bir yöntemdir.<sup>1</sup> Bir diğer tanıma göre ise, “Regresyon analizi, bir değişkenin diğer bir değişken veya değişkenlerle tahmin edilebilmesi için kullanılan istatistik bir yöntemdir.”<sup>2</sup>

Regresyonun günümüzdeki yorumu ise şu şekildedir: “Regresyon çözümlenmesi, bir bağımlı değişkenin başka bağımsız değişken(ler)e olan bağımlılığını, birincinin (anakütle) ortalama değerini, ikinci(ler)in (yinelene örneklerdeki) bilinen ya da değişmeyen değerleri cinsinden tahmin etme ve/ya da kestirme amacıyla inceler.”<sup>3</sup>

Burada sözü edilen tanımlardan da anlaşılacağı gibi bu önemli istatistik tekniğin, tahmin ve politika saptama gibi iki temel işlevi vardır. Regresyon analizini diğer tahmin tekniklerinden üstün kılan en önemli özellik de, aynı zamanda politikaların saptanıp, planlamanın yapılmasında yol gösterici olmasıdır.<sup>4</sup>

---

<sup>1</sup> Thomas P. Ryan, **Modern Regression Methods**, New York, John Wiley & Sons, Inc., 1997, s. 1.

<sup>2</sup> John Netter, William Wasserman, Michael H. Kutner, **Applied Linear Regression Models**, Illinois, Richard D. Irwin, Inc., 1983, s. 23.

<sup>3</sup> Damodar N. Gujarati, **Temel Ekonometri**, Çev. Ümit Şenesen, Gülay Günlük Şenesen, İstanbul, Literatür, 1999, s. 16.

<sup>4</sup> Neyran Orhunbilge, **Uygulamalı Regresyon ve Korelasyon Analizi**, 2. bs., İstanbul, İ.Ü. İşletme Fakültesi, 2002, s. 12.

Regresyon analizindeki ilişki fonksiyonel bir ilişki olmayıp, değişkenler arasında bir sebep sonuç ilişkisi bulunduğu kullanılabilecek bir yöntemdir. Bir başka ifade ile,

“Değişkenler arasındaki istatistik ilişkilerde genellikle *rassal* ya da *olasılıklı (stokastik)*<sup>5</sup> değişkenler, yani olasılık dağılımı olan değişkenler kullanılır. Öte yandan fonksiyonel ya da kesin ilişkilerde de değişkenler kullanılır ama bunlar *rassal* ya da *olasılıklı* değildir.”<sup>6</sup>

Kesin ilişki olması durumunda, değişkenlerden birinin alacağı değer karşısında diğer değişken hep aynı değeri alır. Genellikle az sayıda değişkenin yer aldığı bu tür ilişkilere daha çok fiziksel, kimyasal olaylar gibi olaylarda karşılaşılmaktadır.

Değişkenler arasında stokastik ilişki olması durumunda ise, değişkenlerden birinin alacağı aynı değere karşılık diğer değişken farklı değerler alabilmektedir. Bu tür ilişki söz konusu olduğunda değişkenlerin aldığı değerler modelin matematik yapısı ile tam uyum sağlamazlar, değerler matematik şekilden sapmalar gösterirler. Sapmaları açıklamak için modele hata terimi eklenir. Seçilen matematik kalıp değişkenler arasındaki ilişkiyi en iyi açıklayacak şekil olarak düşünülür. İstatistik olaylarda, genellikle çok sayıda değişken tarafından açıklanabilen bu tür stokastik ilişkiler söz konusudur.

Modellerde, genellikle sayısal olarak ölçülebilen nicel değişkenler yer almalarına rağmen, sayısal olarak ölçülemeyip sadece sayısal olarak ifade edilebilen nitel değişkenlere de yer verilebilmektedir. Kukla (gölge) değişkenler adı verilen bu değişkenler iki veya daha fazla değer alabilmektedirler.

Genellikle regresyon analizine başvurulmasının üç başlıca amacı vardır.<sup>7</sup> Bunlar ilişkinin tanımlanması, ilişkinin kontrol edilebilmesi, tahmin ve politikaların belirlenmesi başlıkları altında toplanabilirler. Bazı araştırmalarda bu amaçlardan

---

<sup>5</sup> Stokastik sözcüğü Yunanca “hedefin göbeği” anlamındaki stokhos sözcüğünden gelir. Bir hedefe ok atmanın çıktısı olasılıklı (stokastik) bir süreçtir, yani hedefi tam vuramayan atışlarla dolu bir süreçtir.

<sup>6</sup> Gujarati, **a.g.e.**, s. 19-20.

<sup>7</sup> Netter, Wasserman, Kutner, **a.g.e.**, s. 30.

sadece biri dikkate alınırken, birçoğunda ise hepsine ait sonuçlar önem kazanmaktadır.

## 1.2. Türleri

Regresyon analizinde, farklı ilişkileri ortaya koyabilmek için değişik türde modeller oluşturulabilmektedir. Regresyon modellerinde, bağımlı<sup>8</sup> değişken etkilenen değişkendir ve bu değişkenin değişimi incelenmektedir. Bu değişken tek veya çok denklemlilerde eşitliğin solunda yer almaktadır. Bağımsız<sup>9</sup> değişkenler sonuç yaratan veya etkileyen değişkenlerdir. Bu değişkenlerin bağımlı değişkendeki değişmelerin nedeni oldukları düşünülmektedir.

Regresyon analizinde modeller bağımsız değişken sayısı, fonksiyon tipi ve verilerin kaynağı dikkate alınarak üç ana grupta toplanabilmektedir. Araştırmalarda, bu üç ana ayırımın alt ayırımlarının bileşimi uygulanmaktadır.<sup>10</sup>

- Bağımsız değişken sayısına göre:
  - Basit regresyon analizi
  - Çoklu regresyon analizi
- Fonksiyon tipine göre:
  - Doğrusal regresyon analizi
  - Doğrusal olmayan regresyon analizi
- Verilerin kaynağına göre:
  - Anakütle verileriyle regresyon analizi
  - Örnek verileriyle regresyon analizi
  - Zaman Serilerinde regresyon analizi

---

<sup>8</sup> Litertürde açıklanan, kestirilen, tepki değişkeni gibi terimler de yer almaktadır.

<sup>9</sup> Litertürde açıklayıcı, kestiren, kontrol değişkeni terimleri de kullanılmaktadır.

<sup>10</sup> Orhunbilge, **a.g.e.**, s.13.



Tek denklemlı regresyon modellerinde bir bağımlı deęişken yer almakta ve bir veya birden fazla bağımsız veya açıklayıcı deęişkenle bağımlı deęişkendeki deęişmeler incelenmektedir. Bu modeller, bir bağımsız deęişken olması durumunda basit regresyon, birden fazla bağımsız deęişken olması durumunda ise çoklu regresyon modelleri olarak adlandırılmaktadırlar. Basit ve çoklu regresyon modellerini oluşturmada kullanılan en yaygın yöntem En Küçük Kareler yöntemidir. Ancak, bu yöntemin sonuçları, modele ait temel varsayımların geçerli olması durumunda uygulanabilmektedir. Bu varsayımlar geçerli olmadıklarında, En küçük Kareler yöntemine alternatif olarak geliştirilen yöntemler bulunmaktadır. Kullanılan bu yöntemler arasında literatürde en sık söz edilenler: Farklı varyanslılık saptandığında, Ağırlıklı En Küçük Kareler yöntemi, bağımsız deęişkenler arasında çoklu doğrusal bağlantı bulunduğunda Ridge Regresyon (Ridge Regression) ve Asal Bileşenler (Principal Components) yöntemi, normallik varsayımının geçerli olmadığı modellerde ise Güçlü Regresyon (Robust Regression) yöntemi olarak sıralanabilir.

Regresyon modelleri parametrelerine veya deęişkenlerine göre doğrusal olmayabilirler. Doğrusal olmayan regresyon modellerinden bir kısmı deęişkenlere çeşitli dönüşümler uygulanarak doğrusallaştırılabilirler. Bu tür modellere *gerçekte doğrusal modeller* adı verilir ve bunlar doğrusal modeller gibi oluşturulurlar. Doğrusal olmayan regresyon modellerinin bir dięer bölümü ise herhangi bir şekilde doğrusallaştırılamazlar. Bu modellerin tahmini için tekrarlı (iterative) yöntemler kullanılır.<sup>11</sup>

Regresyon analizinde de, dięer istatistik yöntemlerde olduęu gibi anakütle verilerinin tümü yerine bu anakütleden seçilen örnek verileriyle analiz yapılır, daha sonra elde edilen sonuçlar anakütledeki ilişkinin tahmininde kullanılır.<sup>12</sup>

---

<sup>11</sup> Selahattin Güriş, Ebru Çaęlayan, **Ekonometri Temel Kavramlar**, İstanbul, Der Yayınevi Yayın No:282, 2000, s. 10.

<sup>12</sup> Orhunbilge, **a.g.e.**, s. 15.

Zaman serileri başlığı altında incelenen analizler ise, tahmini yapılacak değişkenin geçmişteki durumunun çeşitli yöntemlerle incelenmesi ve elde edilen bilgilerden yararlanarak gelecekteki değerlerinin tahmin edilmesine dayanmaktadır.<sup>13</sup> Zaman serileri modellerinde, değişkenlerin geçmiş devre değerleri de modele katılarak gelecekle ilgili tahminler yapılabilmektedir. Eğer model, bağımsız değişkenlerin yalnızca şimdiki değerlerini değil, ama aynı zamanda gecikmeli değerlerini de içeriyorsa, bu modele *gecikmesi dağıtılmış* model adı verilmektedir. Bağımsız değişkenleri arasında bağımlı değişkenin bir veya daha fazla gecikmeli değerlerini içeren modellere *otoregressif* (AR) model denmektedir.<sup>14</sup> *Hareketli Ortalama* (MA) modellerinde ise, bağımlı değişken hata teriminin mevcut ve geçmiş dönem değerlerinin lineer kombinasyonu ile ifade edilmektedir. *Bileşik Otoresgressif Hareketli Ortalama* (ARMA) yönteminde ise model, otoresgressif ve hareketli ortalama modellerinin bileşimiyle oluşturulmaktadır. Bazı iktisadi olayların ise birden fazla denklemle açıklanması gerekmektedir. Bu modellerden en fazla kullanılanları VAR(vektör otoresgressif) ve VARMA(vektör otoresgressif hareketli ortalama) modelleridir. Çok denklemlili bu model veya sistemlerde bazı değişkenler hem bağımlı hem de bağımsız değişken olarak yer alırlar. Böyle sistemlerin eşanlı olarak çözümleri gerektiğinden eşanlı denklem sistemi olarak adlandırılırlar.<sup>15</sup> Çok denklemlili ilişkileri açıklamak için eşanlı denklem sistemleri gibi ardışık sistem, blok-ardışık sistem, görünüşte ilişkisiz regresyon yöntemleri de kullanılabilir.

### 1.3. Basit ve Çoklu Doğrusal Regresyon Analizi

Daha önce de belirtildiği gibi, bir bağımlı değişkenin, tek bir bağımsız değişkenle açıklandığı modeller basit regresyon modelleridir. İki değişken arasındaki ilişkiyi doğrusal varsayarak oluşturulan regresyon modellerine basit doğrusal regresyon modelleri denir. Bu modellerde bağımlı Y değişkenindeki değişimler bağımsız X değişkeni tarafından açıklanmaktadır. Ancak, sosyal ve teknik bilim

---

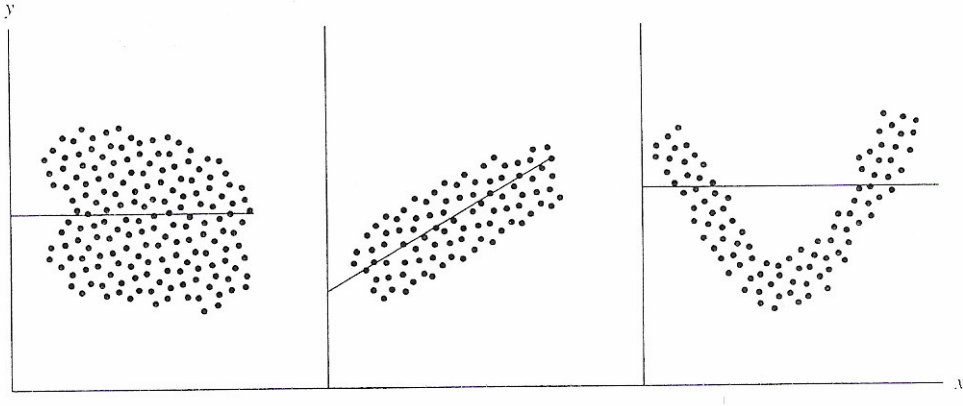
<sup>13</sup> Neyran Orhunbilge, **Zaman Serileri Analizi Tahmin ve Fiyat İndeksleri**, İstanbul, Avcıol Basım, 1999, s. 2.

<sup>14</sup> Fahamet Akın, **Ekonometri**, Bursa, Beta Basım A.Ş., 2002, s. 632.

<sup>15</sup> Güriş, Çağlayan, **a.g.e.**, s. 11-12.

dallarındaki olaylar dikkate alındığında, bir deęişkendeki deęişmelerin sadece bir deęişkenle açıklamanın yeterli olmayacağı açıktır. Bu tür olaylarda çok sayıda faktörün etkisi söz konusu olacağından, ilişki tek denklemlerle açıklanabilirse de çok sayıda bağımsız deęişkene ihtiyaç duyulacaktır. Doğrusal veya doğrusal olmayan fonksiyonlarla ifade edilebilen bu modeller, çoklu regresyon modelleridir. Basit doğrusal regresyon modellerinde olduğu gibi, deęişkenler arasında doğrusal ilişki olduğu varsayımı altında oluşturulan modeller çoklu doğrusal regresyon modelleridir.

Basit regresyon modellerinde, deęişkenler arasında ilişki olup olmadığını saptamak ve ilişkinin hangi tür fonksiyonla gösterilmesi gerektiğini araştırmak için öncelikle, X ve Y deęişkenlerinin aldıkları deęerler iki boyutlu koordinat sisteminde çizilmektedir<sup>16</sup>. Aşağıdaki serpilme diyagramları incelendiğinde ilkinde, iki deęişken arasında bir ilişkiden söz edilememekte, ikincisinde doğrusal, üçüncüsünde ise doğrusal olmayan bir ilişki gözlenmektedir.



**Şekil (1.1): Serpilme Diyagramları**

Doğrusala yakın bir ilişki gözlendiğinde, basit doğrusal regresyon modelleri aşağıdaki matematik fonksiyonla ifade edilmektedirler. Anakütle regresyon modeli,

<sup>16</sup> Bu diyagramlara Serpilme Diyagramları adı verilmektedir.

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad (1.1)$$

olarak oluşturulmaktadır.

Anakütle çoklu doğrusal regresyon modeli, açıklanacak olayda bağımlı değişken dahil p sayıda değişken olduğu varsayıldığında,

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_{p-1} X_{i,p-1} + \varepsilon_i \quad (1.2)$$

denklemini ifade edilebilir. Aynı model denklemini  $X_{i0} = 1$  alınarak,

$$Y_i = \sum_{k=0}^{p-1} \beta_k X_{ik} + \varepsilon_i$$

şeklinde de ifade edilebilir. Bu denklemlerde,  $Y_i$  bağımsız değişkenin i. gözlem değeri,  $X_{ik}$  ise k. bağımsız değişkenin i. gözlem değeridir. Modellerde yer alan  $\beta_0, \beta_1, \dots, \beta_{p-1}$  regresyon katsayıları olarak adlandırılırlar ve modelin parametreleridirler. i ise örnekleme sayısını gösteren endistir ve  $i=1, \dots, n$  şeklinde değişmektedir. Modeldeki  $\beta_0$  sabit katsayı olup, basit doğrusal model dikkate alındığında regresyon doğrusunun Y eksenini kestiği noktayı ifade eder. İstatistik açıdan ise, bağımsız değişken veya değişkenler sıfır değerini aldıklarında Y değişkeninin alacağı değeri göstermektedir. Denklemdeki  $\beta_1, \dots, \beta_{p-1}$  eğim katsayıları, çoğunlukla kısmi regresyon katsayıları olarak adlandırılmaktadırlar. Bunlardan  $\beta_k$  parametresi, diğer tüm bağımsız değişkenler sabit tutulduklarında,  $X_k$  değişkenindeki bir birimlik değişiminin Y değişkeni üzerindeki kısmi etkisini göstermektedir.<sup>17</sup> Kısmi regresyon katsayılarının sıfır olması bağımsız değişkenlerin, bağımlı değişken üzerinde hiçbir etkisinin bulunmadığını, sıfırdan farklılığın büyümesi ise bu etkinin güçlendiğini göstermektedir.<sup>18</sup>

---

<sup>17</sup> William D. Berry, **Understanding Regression Assumptions**, Newbury Park, CA, Sage Publications, Inc., 1993, s. 5.

<sup>18</sup> Orhunbilge, **Uygulamalı Regresyon ve Korelasyon Analizi**, s. 103.

Basit ve çoklu regresyon modellerinde yeralan  $\epsilon_i$  hata<sup>19</sup> terimidir. Tesadüfi bir değişken olan hata terimi  $\epsilon_i$ , gerçek  $Y_i$  değeri ile regresyon denkleminde bulunan tahmini  $Y_i'$  değeri arasındaki farkı, başka bir ifade ile tahminin hatasını göstermektedir.

Çoklu doğrusal regresyon modeli matris notasyonu kullanılarak da ifade edilebilir. Bu durumda (1.2) denklemi

$$Y = X\beta + \epsilon$$

şekline dönüşür. Burada,  $Y$  ( $n \times 1$ ) boyutunda bağımlı değişkenin gözlem değerlerini içeren bir vektör,  $X$  ( $n \times p$ ) boyutunda bağımsız değişkenlerin değerlerini içeren bir matris,  $\beta$  ( $p \times 1$ ) boyutunda bilinmeyen parametreleri içeren bir vektör ve  $\epsilon$  ise ( $n \times 1$ ) boyutunda bağımsız normal dağılım gösteren hata terimlerini içeren vektördür.

### 1.3.1. Regresyon Modelinin Önemli Özellikleri

Basit ve çoklu doğrusal regresyon modellerinde dikkate alınması gerekli temel varsayımlar bulunmaktadır.<sup>20</sup>

1.  $E(\epsilon_i) = 0$  bütün  $i=1,2,\dots,n$  için.  $X$  değişkeni  $X_i$  değerini aldığı anda  $Y_i$  değişkenin beklenen değeri  $E(Y_i|X_i) = \sum_{k=0}^{p-1} \beta_k X_{ik}$  olur.  $X_i$  değeri için  $\epsilon_i$ 'nin dağılımı  $E(Y_i|X_i) = \sum_{k=0}^{p-1} \beta_k X_{ik}$  noktası etrafında normal bir dağılım gösterir. Bütün  $X_i$  değerleri için hata terimlerinin beklenen değerleri  $E(\epsilon_i) = 0$  olur. Bu özelliğe *Normallik (Normality)* varsayımı adı verilmektedir.

<sup>19</sup> Literatürde artık, kalıntı terimleri de kullanılmaktadır.

<sup>20</sup> Tümay Ertek, **Ekonometriye Giriş**, 2. bs., İstanbul, Beta, 2000, s. 160.

2. Hata terimlerinin varyansı sabittir ve  $\sigma^2(\varepsilon_i) = \sigma^2$  olarak ifade edilebilmektedir. Regresyon analizinde *Eşit Varyanslılık (Homoscedasticity)* olarak bilinen bu varsayıma göre,  $X_i$ 'nin alacağı değerler ile oluşan  $\varepsilon_i$ 'lerin varyansı her bir  $X_i$  için aynı olacaktır. Eşit varyanslılık varsayımının geçerli olmaması durumu ise *Farklı Varyanslılık (Heteroscedasticity)* olarak adlandırılmaktadır.
3.  $E(\varepsilon_i \cdot \varepsilon_j) = 0$  varsayımı hata terimlerinin birbirinden bağımsız olduklarını, diğer bir ifade ile birbirlerinden etkilenmediklerini belirtir. Bu özelliğe, hatalar arasında *Otokorelasyon (Autocorrelation)* bulunmaması varsayımı adı verilmektedir.
4.  $E(X_i \cdot \varepsilon_i) = 0$  varsayımı  $\varepsilon_i$  değerinin  $X_i$  değerinden bağımsız olduğunu gösterir.
5.  $\varepsilon_i$  hata teriminin sabit varyansa sahip olduğu dikkate alınırsa  $Y_i$  değişkeninin varyansı,

$$\sigma^2(Y_i) = \sigma^2\left(\sum_{k=0}^{p-1} \beta_k X_{ik} + \varepsilon_i\right) = \sigma^2(\varepsilon_i)$$

olacaktır. Böylece, regresyon modeli, X değişkeninin değerine bağlı kalmaksızın, Y değişkeninin olasılık dağılımının hata terimi ile aynı sabit varyansa ( $\sigma^2$ ) sahip olduğunu varsaymaktadır<sup>21</sup>.

6. Sadece çoklu doğrusal regresyon modeli için geçerli olan bu varsayımda ise, bağımsız değişkenler arasında tam *Çoklu Doğrusal Bağlantı (Multicollinearity)*'nin bulunmadığı varsayılmaktadır. Yani bağımsız değişkenler arasında tam doğrusal ilişkiler yoktur.

---

<sup>21</sup> Netter, Wasserman, Kutner, **a.g.e.**, s. 31.

### 1.3.2. En Küçük Kareler Yöntemi

Basit regresyon analizinde çizilen serpilme diyagramından da görülebileceği gibi, çizilen noktaların arasından sonsuz sayıda doğrusal veya doğrusal olmayan fonksiyon geçirilebilecektir. Bu analizde amaç, en uygun, bir diğer ifade ile ilişkiyi en iyi temsil eden regresyon fonksiyonunu seçebilmektir. Şayet ilişki serpilme diyagramında doğrusal ise, en uygun regresyon doğrusunun bilinmeyenleri olan  $\beta_0$  ve  $\beta_1$  parametrelerinin değerini belirlemek gerekmektedir. Bilindiği üzere anakütle regresyon modeli,

basit doğrusal regresyon için,

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

çoklu doğrusal regresyon için,

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_{p-1} X_{i,p-1} + \varepsilon_i$$

denklemleriyle gösterilmektedir. Bu denklemlerin parametreleri, anakütle verilerinin tüm değerleri gözlenemediği durumda, n birimlik bir örnekten tahmin edilirler. Örnek verileri kullanıldığında denklemlerde tek değişiklik küçük harflerin kullanılmasıdır. Nitekim, bu nedenle örnek verileri ile basit doğrusal regresyon modeli,

$$y_i = b_0 + b_1 x_i + e_i$$

şeklini alır.<sup>22</sup> Burada  $b_0$  ve  $b_1$ <sup>23</sup>,  $\beta_0$  ve  $\beta_1$ 'in,  $e_i$  ise  $\varepsilon_i$ 'nin tahminleyenidirler, yani örnekten belirlenecek değerlerini ifade etmektedirler. Anakütle için tek bir  $\beta_0$  ve  $\beta_1$

---

<sup>22</sup> Matris notasyonu kullanıldığında, anakütle ve örnek modelleri için de matrisler büyük harfler ile gösterilmektedir.

<sup>23</sup> Literatürde  $b_0$  ve  $b_1$  yerine  $\hat{\beta}_0$  ve  $\hat{\beta}_1$  notasyonu da kullanılmaktadır.

değeri söz konusu iken bu anakütleden seçilebilecek bütün mümkün örneklerden herbiri için değişik  $b_0$  ve  $b_1$  değerleri elde edilebilir. Bu  $b_0$  ve  $b_1$  değerlerinin teorik dağılımları normaldir ve beklenen değerleri  $\beta_0$  ve  $\beta_1$ 'e eşittir.

Benzer şekilde, örnek verileri ile çoklu doğrusal regresyon modeli de aşağıdaki şekilde olacaktır.

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots + b_{p-1} x_{i,p-1} + e_i$$

Basit ve çoklu regresyon modellerinde anakütle verileri ile çalışıldığında, gerçek  $Y_i$  değeri, tahmini  $Y_i'$  değeri ile hata terimi  $\varepsilon_i$ 'nin toplamına eşit olacaktır.

$$Y_i = Y_i' + \varepsilon_i$$

Aynı modeller için örnek verileri ele alındığında,

$$y_i = y_i' + e_i$$

eşitliği söz konusudur. Bu durumda,  $Y$  (veya  $y$ ) gözlem değerlerine en yakın tahmini  $Y'$ (veya  $y'$ ) değerleri minimum hata ile tahmin edilmelidir. Bu amaçla, mevcut regresyon analizi yöntemleri arasında, en popüler ve güçlü olan yöntem *En Küçük Kareler Yöntemi*'dir. Bu yöntemin tahminleyenleri, bölüm (1.3.1)'de belirtilen varsayımlar altında, tarafsızlık, etkinlik, tutarlılık gibi önemli istatistik özellikler taşırlar.

En Küçük Kareler yönteminde, parametreler tüm gözlem değerleri için elde edilen hata terimlerinin karelerini minimize edecek şekilde hesaplanmaktadır. Örnek verileri ile çalışıldığında, basit doğrusal regresyon modeli için, farkların kareleri toplamı,

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2$$



olarak ifade edilirse, bu değeri en küçük yapacak  $b_0$  ve  $b_1$  değerlerini bulmak için, bu denklemin  $b_0$  ve  $b_1$  parametrelerine göre kısmi türevleri alınıp sıfıra eşitlendiğinde *normal denklemler* adı verilen denklem sistemi aşağıdaki gibi elde edilir.

$$\sum_{i=1}^n y_i = nb_0 + b_1 \sum_{i=1}^n x_i$$

$$\sum_{i=1}^n x_i y_i = b_0 \sum_{i=1}^n x_i + b_1 \sum_{i=1}^n x_i^2$$

Normal denklemlerin ortak çözümü ile  $b_0$  ve  $b_1$  değerleri elde edilirler.

$$b_1 = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}$$

$$b_0 = \frac{1}{n} (\sum y_i - b_1 \sum x_i)$$

Aynı yöntemde çoklu doğrusal regresyon modeli için, farkların kareleri,

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_{i1} - b_2 x_{i2} - \dots - b_{p-1} x_{i,p-1})^2$$

olarak, matris notasyonu kullanıldığında ise,<sup>24</sup>

$$\begin{aligned} \sum_{i=1}^n e_i^2 &= (Y - Xb)'(Y - Xb) \\ &= Y'Y - b'X'Y - Y'Xb + b'X'Xb \end{aligned} \quad (1.3)$$

<sup>24</sup> Norman R. Draper, Harry Smith, **Applied Regression Analysis**, New York, John Wiley & Sons, 1981, s. 86.

$$= Y'Y - 2b'X'Y + b'X'Xb$$

şeklinde ifade edilebilir. Burada bilindiği gibi  $b'X'Y$  (1×1) boyutunda bir matris ve  $(b'X'Y)' = Y'Xb$  dir.

En Küçük Kareler yönteminde (1.3) denkleminin  $b'$ 'ye göre türevi alınıp 0'a eşitlenmesi ile

$$(X'X)b=X'Y$$

normal denklem sistemi elde edilmektedir. Burada bulunan  $p$  doğrusal denklemin bağımsız olduğu, dolayısıyla  $X'X$ 'in tekil olmayan bir matris olduğu ve tersinin bulunduğu varsayımı vardır. Bu durumda normal denklemlerin çözümü yapıldığında, parametre tahminleyenleri,

$$b=(X'X)^{-1}X'Y$$

denklemleri ile hesaplanmaktadır.<sup>25</sup> Bu tahminleyenler  $\sum_{i=1}^n e_i^2$  değerini minimum yapan değerlerdir, çünkü bu değer  $b'$ 'ye göre ikinci dereceden kısmi türevi olan  $2X'X$  matrisi bir pozitif definit matrisidir.

---

<sup>25</sup> J. Scott Long, **Regression Models for Categorical and Limited Dependent Variables**, California, Sage Publications, Inc., 1997, s. 18.

### 1.3.2.1. En Küçük Kareler Yöntemiyle Elde Edilen Regresyon Katsayılarının Özellikleri

Regresyon modelinin doğrusallık varsayımı altında, modelin parametreleri  $\beta$  vektörünün tahminleyeni  $b$  vektörü bazı özellikler taşımaktadır. Bu özellikler aşağıdaki başlıklarla ele alınabilir.<sup>26</sup>

#### 1.3.2.1.1. Tahminleyenin Dağılımı

Bilindiği üzere, En Küçük Kareler yönteminin tahminleyeni,  $X$  sabit model matrisi ve bağımlı değişkenin gözlem değerlerini içeren  $Y$  matrisi ile ifade edildiğinde,

$$b = (X'X)^{-1}X'Y = MY \text{ olarak gösterilebilir.}$$

Burada yeni bir matris,  $M = (X'X)^{-1}X'$  olarak tanımlanabilir. Böylece,  $b$  vektörünün beklenen değeri,  $Y$ 'nin beklenen değeri ile ifade edilebilir.

$$E(b) = E(MY) = ME(Y) = (X'X)^{-1}X'X\beta = \beta$$

Bu sonuç,  $b$  vektörünün,  $\beta$  vektörünün yansız bir tahminleyeni olduğunu göstermektedir. Benzer şekilde tahminleyenin kovaryans matrisi,

$$V(b) = MV(Y)M' = [(X'X)^{-1}X'] \sigma^2 I_n [(X'X)^{-1}X']'$$

ile gösterilebilir. Bu formülde,  $\sigma^2$  sabit hata varyansını göstermektedir,  $(X'X)^{-1}$  matrisi ise simetrik bir matris olduğu için yukardaki eşitlik,

---

<sup>26</sup> John Fox, **Applied Regression Analysisi, Linear Models, and Related Methods**, California, Sage Publications, 1997, s. 216-217.

$$V(b) = \sigma^2 (X'X)^{-1} X'X(X'X)^{-1} = \sigma^2 (X'X)^{-1}$$

olarak gösterilebilir. Bu durumda, tahminleyenin kovaryans matrisi sadece model matrisi  $X'$ 'e ve hata varyansına bağlı olarak hesaplanabilmektedir.

Sonuç olarak, eğer  $Y$  normal dağılım gösteriyorsa ve doğrusal modelin tüm varsayımları sağlanıyorsa  $b$ , beklenen değeri  $\beta$  ve varyansı  $\sigma^2 (X'X)^{-1}$  olan normal bir dağılım göstermektedir.

$$b \sim N_p[\beta, \sigma^2 (X'X)^{-1}]$$

### 1.3.2.1.2. Gauss-Markov Teorisi

En Küçük Kareler yöntemi modelinde, eğer hata terimleri beklenen değeri 0 ve sabit varyanslı bağımsız bir dağılım gösteriyorlarsa, modelin tahminleyeni  $b$  vektörü de,  $\beta$  vektörünün en etkin tahminleyenidir. Başka bir ifadeyle, gözlem değerlerini doğrusal fonksiyon ile ifade edilebilen tüm yansız tahminleyenler içinde  $b$ , örnekleme varyansı en küçük olan yansız tahminleyendir. Bu yüzden, En Küçük Kareler tahminleyeni  $b$ , literatürde en iyi doğrusal yansız tahminleyen (BLUE)<sup>27</sup> olarak adlandırılır.  $\tilde{b}$  en etkin tahminleyen ise,

$$\tilde{b} = (M+A)Y$$

eşitliği yazılabilir. Burada,  $A$  matrisi, en etkin tahminleyenin dönüşüm matrisi ile, En Küçük Kareler yönteminin tahminleyenin dönüşüm matrisi arasındaki farkı göstermektedir. En Küçük Kareler tahminleyenin en iyi doğrusal yansız

---

<sup>27</sup> BLUE teriminin açılımı Best Linear Unbiased Estimator'dır.

tahminleyen olduğunu göstermek için,  $A=0$  olduğunu göstermek gerekmektedir.<sup>28</sup>  
 $\tilde{b}$ ,  $\beta$ 'nin en iyi doğrusal yansız tahminleyeni olsun. Bu durumda,

$$\beta = E(\tilde{b}) = E[(M+A)Y] = E(MY) + E(AY) = E(b) + AE(Y) = \beta + AX\beta$$

eşitliği yazılabilir. Bu eşitliğin sağlanması için  $AX\beta$  teriminin, dolayısıyla  $A$  matrisinin 0 matrisi olması gerekmektedir.

$\tilde{b}$  doğrusal yansız ve minimum varyanslı bir tahminleyen olduğu için,  $V(\tilde{b})$  matrisinin köşegen üzerindeki elemanları mümkün olacak en küçük değerlere eşittirler.  $\tilde{b}$ 'nin kovaryans matrisi,

$$\begin{aligned} V(\tilde{b}) &= (M+A)V(Y)(M+A)' \\ &= (M+A) \sigma^2 I_n (M+A)' \\ &= \sigma^2 (MM' + MA' + AM' + AA'), \end{aligned}$$

Daha önce gösterildiği gibi  $AX = 0$ 'dır. Bu durumda,

$AM' = AX(X'X)^{-1} = 0(X'X)^{-1} = 0$  olacaktır. Dolayısıyla, bu matrisin transpozu olan  $MA'$  matrisi de 0 olacaktır.

$$V(\tilde{b}) = \sigma^2 (MM' + AA')$$

Bilindiği gibi kovaryans matrisinin köşegendeki  $j$ 'inci elemanı, kısmi regresyon katsayısı  $\tilde{b}_j$ 'in varyansına eşittir.

$$\sigma^2(\tilde{b}_j) = \sigma^2 \left( \sum_{i=1}^n m_{ji}^2 + \sum_{i=1}^n a_{ji}^2 \right)$$

<sup>28</sup>Ronald J. Wonnacott, Thomas H. Wonnacott, **Econometrics**, New York, Wiley, 1979, s.. 428-430.

Bu eşitlikte tüm terimler negatif olamayacakları için,  $V(\tilde{b}_j)$  değerinin olabilecek en küçük değere sahip olabilmesi için tüm  $a_{ji}$  terimlerinin 0 olması gerekmektedir. Bu durum, tüm köşegendeki elemanlar için geçerlidir. Dolayısıyla, A matrisinin tüm satırlarındaki elemanların 0 olması, başka bir ifade ile  $A=0$  olması gerekmektedir. Sonuç olarak,

$$\tilde{b}=(M+0)Y=MY=b$$

olduğu gösterilmiştir. Böylece En Küçük Kareler Yönteminin tahminleyenin, en iyi doğrusal yansız tahminleyen (BLUE) olduğu ispatlanmaktadır.

### 1.3.2.2. Regresyon Denklemine Standart Hatası

Anakütle verileri ile çalışıldığında, hata terimi  $\varepsilon_i$ 'nin varyansı  $\sigma^2$ , hem Y'nin hem de regresyon denkleminin değişkenliğini (varyansını) temsil ettiği için tahmin edilmesi gerekmektedir.

Bilindiği üzere, Y'nin her bir gözlem değeri  $Y_i$ 'nin, tahmini değeri  $\hat{Y}_i$ 'den olan farkı, regresyon denkleminde olan sapmaları yani regresyon analizi ile yapılan hataları göstermektedir. Bu değerlerin kareleri toplamı, HKT<sup>29</sup>,

Anakütle verileri için;

$$AHKT = \sum_{i=1}^N (Y_i - \hat{Y}_i)^2$$

Örnek verileri için ise;

---

<sup>29</sup> HKT teriminin açılımı Hata Kareleri Toplamı'dır. Literatürde SSE (Error Sum of Squares) olarak geçmektedir.

$$\text{HKT} = \sum_{i=1}^n (y_i - y'_i)^2$$

şeklinde gösterilebilir. Bu değerin ortalama değeri ise regresyon denkleminin varyansını verecektir. Dolayısıyla, çoklu doğrusal regresyon denkleminin standart hatası,

Anakütle verileri için;

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (Y_i - Y'_i)^2}{N}}$$

Örnek verileri için ise;

$$s = \sqrt{\frac{\sum_{i=1}^n (y_i - y'_i)^2}{n - p}}$$

şeklinde olacaktır.<sup>30</sup> Anakütle verileriyle çalışırken, varyansı hesaplamak için ortalama alındığında hata karelerinin toplamının gözlem sayısı N'ye bölünmesi gerekmektedir. Örnek verileriyle çalışıldığında ise aynı terim serbestlik derecesi olarak bilinen ve çoklu doğrusal regresyon analizinde (n-p) değerine eşit olan terime [basit doğrusal regresyon analizinde (n-2) değerine] bölünmesi gerekmektedir. Buradaki p bağımsız değişken dahil tüm değişken sayısını göstermektedir. Bilindiği üzere ancak bu şekilde tahmin edilen standart hatanın beklenen değeri anakütle standart hatasına eşit olmaktadır. Sembollerle ifade edildiğinde, örnek varyanslarının beklenen değeri

---

<sup>30</sup> Draper, Smith, **a.g.e.**, s. 207.

$$E\left(\frac{\sum (y_i - y'_i)^2}{n}\right) = \frac{\sum (Y_i - Y'_i)^2}{N} \cdot \frac{n-p}{n} = \sigma^2 \cdot \frac{n-p}{n} \quad (1.4)$$

olarak hesaplandığında anakütle varyansının yanlış tahminleyeni olmakta ve bu değeri olması gerekenden daha düşük tahmin etmektedir<sup>31</sup>.

$$E\left(\frac{\sum (y_i - y'_i)^2}{n}\right) < \frac{\sum (Y_i - Y'_i)^2}{N}$$

Bu durumda (1.4) eşitliğinin her iki tarafı  $\frac{n}{n-p}$  değeri ile çarpıldığında anakütle varyansının yansız tahminleyeni aşağıdaki formülle elde edilmektedir.

$$E\left(\frac{\sum (y_i - y'_i)^2}{n} \cdot \frac{n}{n-p}\right) = E\left(\frac{\sum (y_i - y'_i)^2}{n-p}\right) = E(s^2) = \sigma^2$$

#### 1.4. Korelasyon Analizi

Regresyon analizinde değişkenler arasındaki ilişki matematik bir fonksiyonla ortaya konurken, regresyon katsayılarındaki işaretler değişkenler arasındaki ilişkinin yönünü de belirtir, fakat ilişkinin derecesi hakkında bir bilgi vermez. Bu yüzden modelde yeralan bağımsız değişken veya değişkenlerin, bağımlı değişkendeki değişimleri açıklama oranının bilinmesi, modelin uygunluğu konusunda bir ölçü oluşturulacaktır.<sup>32</sup> Korelasyon analizinin amacı “Korelasyon katsayısı” olarak bilinen bu oranın hesaplanmasıdır.

<sup>31</sup> Neyran Orhunbilge, **Örnekleme Yöntemleri ve Hipotez Testleri**, 2. bs., İstanbul, Avcıol Basım, 2000, s. 26-27.

<sup>32</sup> Güriş, Çağlayan, **a.g.e.**, s. 135.



### 1.4.1. Basit Doğrusal Korelasyon Analizi

Basit doğrusal regresyon modeli dikkate alındığında, bağımlı değişkendeki toplam değişkenlik,

$$\sum_{i=1}^N (Y_i - \bar{Y})^2 = \sum_{i=1}^N (Y'_i - \bar{Y})^2 + \sum_{i=1}^N (Y_i - Y'_i)^2$$

formülüyle elde edilebilir.<sup>33</sup> Buradan görüldüğü gibi, Y'deki toplam değişkenlik (GKT)<sup>34</sup>, regresyon denklemi ile açıklanan değişkenlik (RKT)<sup>35</sup> ve açıklanamayan değişkenlik (HKT) terimlerinin toplamıyla ifade edilmektedir. Eşitliğin her iki tarafı toplam değişkenliğe bölüldüğünde, eşitliğin sağ tarafındaki ilk terim  $\rho^2$  ile ifade edilip, “belirlilik katsayısı” olarak bilinen, ve bağımlı değişkendeki değişkenliğin, bağımsız değişken tarafından açıklanma oranını gösteren değer elde edilmektedir. Bu durumda,

$$\rho^2 = \frac{\sum_{i=1}^N (Y'_i - \bar{Y})^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2} = 1 - \frac{\sum_{i=1}^N (Y_i - Y'_i)^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2}$$

veya

$$\rho^2 = \frac{RKT}{GKT} = 1 - \frac{HKT}{GKT}$$

formülleriyle ifade edilir.<sup>36</sup> Belirlilik katsayısının karekökü alındığında, genel korelasyon katsayısı elde edilecektir.

---

<sup>33</sup> Burada  $\bar{Y}$ , Y'nin aritmetik ortalamasını ifade etmektedir.

<sup>34</sup> GKT'nin açılımı, Genel Kareler Toplamı'dır.

<sup>35</sup> RKT'nin açılımı, Regresyon Kareleri Toplamı'dır.

<sup>36</sup> Allen L. Edwards, **Introduction to Linear Regression and Correlation**, second edition, New York, W.H. Freeman and Company, 1984, s. 137.

$$\rho = \sqrt{1 - \frac{\sum_{i=1}^N (Y_i - Y_i')^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2}}$$

Tüm korelasyon katsayılarının genel formülü niteliğini taşımakta olan bu formül,  $\sum_{i=1}^n (Y_i - Y_i')$  değeri  $Y_i'$  değerinin hangi fonksiyonla elde edildiğine bağlı olarak basit doğrusal, eğrisel veya çoklu korelasyon katsayısı şekline dönüşmektedir. Regresyon modeli, bağımlı değişkendeki toplam değişkenliğin tümünü açıklıyorsa,  $\rho=1$ , hiç açıklamıyorsa  $\rho=0$  olacaktır. Bu değer 1'e yaklaşması değişkenliğin, regresyon denklemiyle iyi açıklandığını ortaya koyacaktır. Doğal olarak regresyon analizinde amaç, arzu edilen büyük  $\rho$  değerlerine ulaşmaktır.

Anakütle verileri için verilen tüm bu formüller örnek verileri için de geçerli olmakta, ancak kullanılan notasyonda genellikle örnek belirlilik katsayısı  $r^2$  ve korelasyon katsayısı ise  $r$  ile gösterilmektedir.

## 1.4.2. Çoklu Doğrusal Korelasyon Analizi

### 1.4.2.1. Çoklu Korelasyon Katsayısı

Çoklu regresyon analizinde ise belirlilik ve korelasyon katsayılarına çoklu belirlilik ve çoklu korelasyon katsayısı adı verilmektedir. Anakütle verileri dikkate alındığında ve genel model için matris notasyonu kullanıldığında, bağımlı değişkendeki toplam değişkenlik, bağımsız değişkenlerle açıklanabilen ve açıklanamayan değişkenliklerin toplamından oluşacaktır.<sup>37</sup>

<sup>37</sup> Netter, Wasserman, Kutner, **a.g.e.**, s. 239.

Toplam deęişkenlik

$$GKT = Y'Y - (1/N)Y'11'Y$$

Regresyonla açıklanan deęişkenlik

$$RKT = \beta'X'Y - (1/N)Y'11'Y$$

Açıklanamayan deęişkenlik

$$HKT = \varepsilon'\varepsilon = (Y - X\beta)'(Y - X\beta) = Y'Y - \beta'X'Y$$

Çoklu belirlilik katsayısı  $\rho^2$ , bağımlı deęişkendeki toplam deęişkenlikte, doğrusal regresyon denklemi ile açıklanabilen kısmın oranı şeklinde tanımlanmaktadır. Bu durumda,

$$\rho^2 = \frac{RKT}{GKT} = 1 - \frac{HKT}{GKT}$$

olarak ifade edilebilir. Basit korelasyon katsayısı gibi, çoklu korelasyon katsayısı da  $\rho$  ile ifade edilip  $Y$  ile  $Y'$  arasındaki ilişkiyi, yani bağımsız deęişkenlerin birlikte  $Y'$ deki etkisinin derecesini temsil etmektedir. Basit doğrusal regresyon modelindeki korelasyon katsayısı ile ilgili tüm yorumlar çoklu doğrusal regresyon modeli için de geçerlidir.

Basit doğrusal korelasyon analizinde olduğu gibi, anakütle verileri için elde edilen bu formüller örnek verileri söz konusu olduğunda da aynı şekilde hesaplanmakta, ancak  $\rho^2$  ve  $\rho$  için sırasıyla  $r^2$  ve  $r$  notasyonu kullanılmaktadır.

$$r = \sqrt{1 - \frac{\sum_{i=1}^n (y_i - y'_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

Regresyon analizinde, aynı bağımlı değişkendeki değişmelerin farklı regresyon modelleri ile açıklanabileceği düşünülmektedir. Gözlem ve değişken sayıları farklı olabilecek bu modelleri karşılaştırmada kullanılacak ölçüt korelasyon katsayısı olmaktadır. Ancak,  $r$ 'nin önemli bir özelliği modele yeni bağımsız değişkenler eklendikçe azalmadığı, her zaman artma eğiliminde olmasıdır. Bu yüzden bağımlı değişkenleri aynı, fakat bağımsız değişken sayıları farklı olan modellerin karşılaştırılmasında, bağımsız değişken sayısını da hesaba katan *düzeltilmiş korelasyon katsayısı (adjusted correlation coefficient)*  $\bar{r}$  kullanılır.

$$\bar{r} = \sqrt{1 - \frac{\sum_{i=1}^n (y_i - y_i')^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \cdot \frac{n-1}{n-p}}$$

Düzeltilmiş korelasyon katsayısı ile korelasyon katsayısı arasındaki ilişki,

$$\bar{r} = \sqrt{1 - (1 - r^2) \frac{n-1}{n-p}}$$

formülüyle gösterilmektedir.<sup>38</sup>

Düzeltilmiş korelasyon katsayısı için kullanılacak bir başka değer de Goldberger<sup>39</sup> tarafından önerilmiştir.

$$\bar{r}_G = (1 - p/n)r^2$$

Günel<sup>40</sup>, regresyon denklemlerinin başarılarının karşılaştırılmasında, belirlilik katsayısının tek başına regresyon denkleminin özelliklerini temsil edemediğini ve

<sup>38</sup> Formüllerdeki  $p$  sayısı, bağımlı dahil modeldeki toplam değişken sayısını göstermektedir.

<sup>39</sup> Arthur Goldberger, **A Course in Econometrics**, Cambridge, Harvard University Press, 1991, s. 178.

örnek büyüklüğü, denklemin eğimi ve standart hatası gibi ek kriterlere gereksinim olduğunu, düzeltilmiş belirlilik katsayısı kullanıldığında ise örnek büyüklüğünün bu değeri negatif yapan eşik değerinin bilinmesi gerektiğini irdelemiştir.

### 1.4.2.2. Kısmi Korelasyon Katsayıları

Çoklu korelasyon katsayısı, tüm bağımsız değişkenlerin bağımlı değişkeni açıklamadaki gücünü yüzde olarak göstermektedir. Ancak, özellikle politika belirlemek söz konusu olduğunda bir bağımsız değişkenin, tüm diğer bağımsız değişkenler modelde kalmak kaydıyla, tek başına Y'ye olan marjinal katkısının da bilinmesi gerekmektedir. Regresyon analizinde bu değere kısmi korelasyon katsayıları adı verilmektedir. Her bir bağımsız değişken için ayrı ayrı hesaplanabilen bu değer ait olduğu bağımsız değişkenin, diğer bağımsız değişkenler sabit kalmak şartıyla, bağımlı değişkenin hangi oranda etkilediğini derece olarak göstermektedir.

Genel olarak, (p-1) adet bağımsız değişkeni bulunan bir modelde örnek verileri ile çalışılmışsa, k'ıncı değişkenin kısmi korelasyon katsayısı  $r_{Yk.1,2,\dots,(k-1),(k+1),\dots,(p-1)}$  olarak ifade edilebilir. Bu katsayının değeri,

$$r_{Yk.1,2,\dots,(k-1),(k+1),\dots,(p-1)} = \left( \frac{r_{Y.1,2,\dots,(p-1)}^2 - r_{Y.1,2,\dots,(k-1),(k+1),\dots,(p-1)}^2}{1 - r_{Y.1,2,\dots,(k-1),(k+1),\dots,(p-1)}^2} \right)^{1/2}$$

veya,

$$= \left( \frac{HKT(x_1, x_2, \dots, x_{(k-1)}, x_{(k+1)}, \dots, x_{(p-1)}) - HKT(x_1, x_2, \dots, x_{(p-1)})}{HKT(x_1, x_2, \dots, x_{(k-1)}, x_{(k+1)}, \dots, x_{(p-1)})} \right)^{1/2}$$

---

<sup>40</sup> Alptekin Günel, "Regresyon Denkleminin Başarısını Ölçmede Kullanılan Belirleme Katsayısı ve Kritiği", *Doğuş Üniversitesi Dergisi*, C.IV, 2003, s. 133-140.

formülleri ile hesaplanabilmektedir.<sup>41</sup>

Buradaki  $r_{Y.1,2,\dots,(p-1)}$  terimi çoklu korelasyon katsayısı  $r$ 'yi ve  $r_{Y.1,2,\dots,(k-1),(k+1),\dots,(p-1)}$  terimi ise  $k$ . değişken hariç, diğer tüm değişkenler modele dahil edildiğinde hesaplanan çoklu korelasyon katsayısını göstermektedir.

## 1.5. Anakütle Verileri ile Tahmin ve Politikaların Belirlenmesi

Basit doğrusal regresyon analizi anakütle verileri ile yapılmış ise, En Küçük Kareler yöntemi ile regresyon doğrusunun parametreleri hesaplanır. Elde edilen bu denklem ile  $Y$ 'nin tahmini değerleri ve bu değerlerden hareketle hata terimleri elde edilir. Bu bulgular ile öncelikle modelde, bölüm(1.3.1)'de sözü edilen normallik, eşit varyanslılık ve otokorelasyon bulunmaması varsayımlarının geçerli olup olmadıkları uygun testler ile araştırılır. Bu varsayımlardan sapmalar söz konusu değilse, öncelikle modelin korelasyon katsayısı hesaplanır. Korelasyon katsayısı 1'e yakınsa  $Y$ 'yi açıklamakta  $X$ 'in yeterli olduğuna ve regresyon denkleminin  $Y$ 'nin tahmini için kullanılabileceğine karar verilir. İlgilenilen konu sadece tahmin ise,  $X_0$ 'ın verilmiş bir  $X$  değeri olduğu varsayımı ile  $Y_0$ 'ın tahmini değeri,

$$Y_0' = \beta_0 + \beta_1 X_0$$

regresyon denklemi ile elde edilir. *Nokta tahmini* adı verilen bu tahmini değer En Küçük Kareler yöntemi ile minimize edilmiş bir de standart hatası vardır ve bu hatanın da tahmine ilave edilmesi gerekmektedir.<sup>42</sup> Bu durumda, belirli bir olasılıkla  $Y$ 'nin güven aralığı aşağıdaki şekilde yazılabilir.

$$\hat{Y} = Y_0' \pm Z \sigma$$

<sup>41</sup> Netter, Wasserman, Kutner, **a.g.e.**, s. 287.

<sup>42</sup> Orhunbilge, **Uygulamalı Regresyon ve Korelasyon Analizi**, s. 31.

Bilindiği üzere regresyon analizinde amaç, sadece tahmin değil aynı zamanda bağımlı değişken üzerinde gelecekle ilgili planlar yapabilmektir. Bu durumda analizde ilgi, sadece Y üzerinde politika belirlemek ise, korelasyon katsayısının hesaplanması yeterli olacaktır. Bu katsayının 1'e yakın olması durumunda, Y üzerinde saptanacak politikalarda X'deki değişimler etkili olacaktır.

Çoklu doğrusal regresyon analizinde ise, tüm anakütle değerleri gözlenmiş ve analiz bu veriler ile yapılmışsa regresyon denklemi yazıldıktan sonra bölüm (1.3.1)'de sözü edilmiş tüm varsayımların geçerlilikleri araştırılır. Hepsinin geçerli olması durumunda, modelin çoklu ve kısmi korelasyon katsayıları hesaplanır. Tüm korelasyon katsayıları 1'e yakınsa, regresyon denklemi Y'nin tahmininde kullanılabilir. Verilen  $X_{01}, X_{02}, \dots, X_{0(p-1)}$  değerleri için nokta tahmini  $Y_0'$ , regresyon denkleminde bu değerlerin yerleştirilmesi ile hesaplanır.

$$Y_0' = \beta_0 + \beta_1 X_{01} + \beta_2 X_{02} + \dots + \beta_{(p-1)} X_{0(p-1)}$$

Basit regresyon analizine benzer biçimde tahmine standart hatanın eklenmesi ile belirli bir olasılık değeri seçilerek aralık tahmini yapılır.

$$\hat{Y} = Y_0' \pm Z \sigma$$

Tahmin dışında, Y üzerinde saptanacak politikaların belirlenmesinde, hesaplanan çoklu ve kısmi korelasyon katsayılarının değerleri önem kazanmaktadır. Tüm bu katsayıların 1'e yakınlığı, Y'nin gelecekle ilgili planlarında, tüm bağımsız değişkenlerdeki değişmelerin etkili olacağını belirlemektedir. Geliştirilecek politikalarda öncelik, doğal olarak kısmi korelasyon katsayısı en yüksek olan bağımsız değişkenden başlanarak sırasıyla diğerlerine verilecektir.

## 1.6. Örnek Verileri ile Doğrusal Regresyon ve Korelasyon Analizi

Bilindiği üzere, anakütleyi oluşturan tüm birimler gözlenemediği durumda, tesadüfi olarak seçilen n birimlik örnekten, regresyon denkleminin ve korelasyon katsayısının tahmini değerleri hesaplanmaktadır. Fakat amaç sadece bu değerleri tahmin etmek değil, aynı zamanda bu tahmini değerlerin anakütledeki gerçek değerlere ne kadar yakın olduklarını bilmektir. Bu durumda öncelikle tahminleyenlerin doğruluk veya güvenilirliklerinin bir ölçüsüne yani standart hata değerlerine ihtiyaç vardır.

Bölüm (1.3.2.1.2)'de ispatlandığı gibi En Küçük Kareler yöntemiyle elde edilen regresyon katsayılarının tahminleyenleri, en iyi doğrusal yansız tahminleyenlerdir. Bu durumda, basit doğrusal regresyon modeli ele alındığında,

$$E(b_0) = \beta_0 \quad \text{ve} \quad E(b_1) = \beta_1$$

olmakta ve tahminleyenlerin standart hata değerleri,

$$s^2(b_0) = E(b_0 - E(b_0))^2 = E(b_0 - \beta_0)^2 \quad \text{açılımıyla}$$

$$s(b_0) = s \sqrt{\left( \frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2} \right)}$$

ve

$$s^2(b_1) = E(b_1 - E(b_1))^2 = E(b_1 - \beta_1)^2 \quad \text{açılımıyla}$$

$$s(b_1) = s \sqrt{\left( \frac{1}{\sum (x_i - \bar{x})^2} \right)}$$



olarak elde edilmektedirler.<sup>43</sup> Bu formüllerdeki  $s$ , regresyon denkleminin tahmini standart hatasıdır.

Çoklu doğrusal regresyon modellerinde ise, bölüm (1.3.2.1.1)'de gösterildiği gibi tahminleyen kovaryans matrisi,

$$V(b) = s^2 (X'X)^{-1} \quad (1.5)$$

olarak elde edilir. Bu notasyonla,  $k$ . kısmi regresyon katsayısının standart hatası,

$$s(b_k) = s \sqrt{c_{kk}}$$

formülüyle hesaplanabilir. Burada  $c_{kk}$  sayısı,  $(X'X)^{-1}$  matrisinin  $b_k$  parametresine ait köşegendeki elemanıdır. Katsayılar arasındaki kovaryanslar ise, aynı matriste, köşegen üzerinde bulunmayan elemanlar seçilip  $s^2$  ile çarpılarak elde edilirler.

$$\text{Cov}(b_i, b_j) = s^2 c_{ij}$$

Basit ve çoklu doğrusal regresyon modellerinde, örnek verileri ile bulunan korelasyon katsayısının beklenen değeri anakütle korelasyon katsayısını verecektir.

$$E(r) = \rho$$

Aynı katsayının standart hatası ise,

$$s(r) = \sqrt{\frac{1-r^2}{n-p}}$$

formülüyle hesaplanmaktadır.

---

<sup>43</sup> Netter, Wasserman, Kutner, **a.g.e.**, s. 62-69.

## 1.6.1. Regresyon Analizinde Testler

Bilindiği üzere, anakütle parametreleri hakkında bir varsayımın belirli bir anlamlılık düzeyinde geçerliliğinin, örnek istatistiklerden hareketle araştırılmasına *Hipotez Testleri* adı verilir.<sup>44</sup> Regresyon analizinde, örnek verilerinden elde edilen tahmini değerlerle, anakütle değerleri arasındaki farkın anlamlı olup olmadığının, başka bir ifade ile tahmini değerlerin anakütle değerlerini ne derece iyi temsil ettiklerinin bilinmesi gerekmektedir. Bu amaçla, basit doğrusal regresyon modelinde, değişkenler arasındaki ilişkiyi gösteren  $b_1$  parametresi ile basit korelasyon katsayısının, çoklu modellerde ise çoklu korelasyon katsayısının yanısıra kısmi regresyon katsayılarının veya kısmi korelasyon katsayılarının anakütlede geçerliliğinin araştırılması için Hipotez testleri uygulanmaktadır.

### 1.6.1.1. Regresyon Katsayılarının Testleri

Basit doğrusal regresyon modelinde  $b_0$  katsayısının değeri regresyon doğrusunun Y eksenini kestiği noktayı belirlemekte, ancak değişkenlerin ilişkisi hakkında bir fikir vermemektedir. Bu yüzden anlamlılığının araştırılması gerekmemektedir. Anakütlede, iki değişken arasındaki ilişkinin güçlü olup olmadığının göstermek için  $b_1$  katsayısı test edilmektedir. Temel ve alternatif hipotezler aşağıdaki şekilde oluşturulurlar:

$$H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

Temel hipotez, iki değişken arasında ilişki olmadığını, çift taraflı olarak oluşturulan alternatif hipotez bu ilişkinin anakütlede anlamlı olduğunu ifade etmektedir.

---

<sup>44</sup> Orhunbilge, *Örnekleme Yöntemleri ve Hipotez Testleri*, s. 131.

Tablo deęerinin bulunabilmesi için öncelikle anlamlılık seviyesi  $\alpha$ 'nın belirlenmesi gerekmektedir. Bilindięi gibi  $\alpha$  hatası, temel hipotez  $H_0$  doęru iken reddedilmesiyle,  $\beta$  hatası ise yanlış olan  $H_0$ 'ın kabulüyle ortaya çıkmakta ve bu hatalardan biri belirlendięinde dięeri de belirlenmiř olmaktadır. Uygulamalarda, genellikle daha kritik sonuçlara neden olan hata küçük tutulmakta, her ikisinin de önemli olması durumunda örnek sayısı büyük alınmaktadır. Regresyon analizinde de doęru olan  $H_0$  hipotezinin reddedilmesiyle, yani anakütlede baęımlı ile baęımsız deęişkenler arasında anlamlı bir iliřki bulunmadıęı halde, iliřkinin anlamlı bulunmasıyla yapılan  $\alpha$  hatası önem taşımakta ve bu yüzden yaygın olarak %5 veya tercihen %1 olarak seçilmektedir.

Bilindięi üzere,  $b_1$  ortalaması  $\beta_1=0$  etrafında normal daęılım gösterir. Bu varsayımdan hareketle, standart deęişken,  $n \geq 30$  için,

$$\frac{b_1 - \beta_1}{s(b_1)} \approx Z(0,1)$$

ve  $n < 30$  için

$$\frac{b_1 - \beta_1}{s(b_1)} \approx t_{n-2}$$

daęılımlarını gösterecektir.<sup>45</sup>  $H_0$  hipotezinde  $\beta_1=0$  olduęu için, test istatistięi

$$Z^* \text{ veya } t^* = \frac{b_1}{s(b_1)}$$

olarak hesaplanmakta ve karar

$$|Z^*| \leq Z_{\alpha/2} \text{ veya } |t^*| \leq t_{(\alpha/2;n-2)} \text{ ise } H_0 \text{ kabul}$$

---

<sup>45</sup> Orhunbilge, **Uygulamalı Regresyon ve Korelasyon Analizi**, s. 35-36.

$|Z^*| > Z_{\alpha/2}$  veya  $|t^*| > t_{(\alpha/2;n-2)}$  ise  $H_0$  red şeklinde alınmaktadır.

Çoklu doğrusal regresyon modelinde, diğer bağımsız değişkenler modelde sabit tutulmak kaydıyla, temsil ettikleri bağımsız değişkenin bağımlı değişken üzerindeki kısmi etkisini ifade eden kısmi regresyon katsayılarının anlamlılık testlerinin yapılması gerekmektedir. Bu amaçla, sabit katsayı hariç herbir parametre için hipotezler,

$$H_0: \beta_k=0$$

$$H_1: \beta_k \neq 0$$

şeklinde oluşturulmaktadır.<sup>46</sup>

Basit doğrusal regresyon analizinde olduğu gibi,  $\frac{b_k - \beta_k}{s(b_k)}$  standart değişkeni,

Z veya (n-p) serbestlik dereceli  $t$  dağılımı göstermektedir.

Test istatistiği,

$$Z^* \text{ veya } t^* = \frac{b_k}{s(b_k)}$$

formülüyle hesaplanır. Karar,

$|Z^*| \leq Z_{\alpha/2}$  veya  $|t^*| \leq t_{(\alpha/2;n-p)}$  ise  $H_0$  kabul

$|Z^*| > Z_{\alpha/2}$  veya  $|t^*| > t_{(\alpha/2;n-p)}$  ise  $H_0$  red şeklinde alınmaktadır.

$\beta_k$  parametresinin güven aralığı ise;

$b_k \mp t_{(\alpha/2;n-p)} s(b_k)$  şeklinde hesaplanabilmektedir.

---

<sup>46</sup> Burada  $k=1, \dots, (p-1)$  aralığında değişmektedir.

### 1.6.1.2. Korelasyon Katsayılarının Testleri

Basit korelasyon katsayısı için t, Z veya F testi uygulanabileceği halde çoklu korelasyon katsayısı için sadece F testi kullanılabilir. <sup>47</sup>

Basit korelasyon katsayısının Z ve t testinde, regresyon katsayısı  $b_1$ 'in hipotez testindeki benzer aşamalar uygulanmaktadır. Örnek verileri ile elde edilen r değerinin anlamlılığını araştırmak için, Hipotezler,

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

şeklinde kurulur. Anakütlede iki değişken arasında ilişki olmadığında ( $\rho = 0$ ), örnek korelasyon katsayıları ortalamaları olan (0) etrafında  $n \geq 30$  olduğunda Z,  $n < 30$  olduğunda ise (n-2) serbestlik dereceli t dağılımı göstermektedir. <sup>48</sup> Buradan hareketle test istatistiği,

$$Z^* \text{ veya } t^* = \frac{r}{s_r} = r \cdot \sqrt{\frac{n-2}{1-r^2}}$$

şeklinde yazılabilir. Karar,

$$|Z^*| \leq Z_{\alpha/2} \text{ veya } |t^*| \leq t_{(\alpha/2; n-2)} \text{ ise } H_0 \text{ kabul}$$

$$|Z^*| > Z_{\alpha/2} \text{ veya } |t^*| > t_{(\alpha/2; n-2)} \text{ ise } H_0 \text{ red olarak verilecektir.}$$

Çoklu regresyon modellerinde, çoklu korelasyon katsayısı testi için sadece F testi kullanılmaktadır. Aynı test  $p=2$  alınarak uygulandığında, basit korelasyon

---

<sup>47</sup> t ve Z testleri sadece bir veya iki değişken söz konusu olduğunda kullanıldıkları için, çoklu korelasyon katsayısının testinde yetersiz kalmaktadırlar.

<sup>48</sup> Orhunbilge, **Uygulamalı Regresyon Korelasyon Analizi**, s. 37.

katsayısının geçerliliğini test edecek ve doğal olarak, Z veya t testi ile aynı sonucu verecektir.

Çoklu korelasyon katsayısının F testinde, temel hipotez anakütlede bağımlı ile bağımsız değişkenler arasında ilişki olmadığını, alternatif hipotez ise anlamlı bir ilişki bulunduğunu ifade eder.

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

F testinde (varyans analizi) açıklanan varyans, açıklanamayan varyansa bölünmektedir. Regresyon analizinde, varyans tablosu matris notasyonu kullanıldığında aşağıdaki şekilde oluşturulmaktadır.<sup>49</sup>

**Tablo (1.1): Çoklu Regresyon Analizinde Varyans Analizi Tablosu**

Değişkenlik	Kareler Toplamı	Serbestlik Derecesi	Kareler Ortalaması
Regresyon	$RKT = b'X'Y - (1/n)Y'11'Y$	p-1	$RKT/(p-1)$
Hata	$HKT = Y'Y - b'X'Y$	n-p	$HKT/(n-p)$
Toplam	$GKT = Y'Y - (1/n)Y'11'Y$	n-1	

Test istatistiği,

$$F^* = \frac{\sum (y' - \bar{y})^2 / (p - 1)}{\sum (y - y')^2 / (n - p)} = \frac{RKT / (p - 1)}{HKT / (n - p)}$$

formülüyle hesaplanmakta ve karar

$$F^* \leq F_{(\alpha; p-1; n-p)} \text{ ise } H_0 \text{ kabul}$$

<sup>49</sup> Netter, Wasserman, Kutner, **a.g.e.**, s. 240.

$$F^* > F_{(\alpha;p-1;n-p)} \text{ ise } H_0 \text{ red}$$

olarak alınmaktadır.

### 1.6.1.3. Kısmi Korelasyon Katsayılarının Testleri

Kısmi regresyon katsayılarına uygulanan testlere benzer testler ile, kısmi korelasyon katsayılarının anlamlı olup olmadığı ortaya çıkmaktadır. Bu analiz sonucunda, kısmi regresyon testlerine paralel olarak hangi bağımsız değişkenlerin model dışında kalması gerektiği, hangilerinin modele dahil edilmesi gerektiği belirlenebilmektedir. Aynı ilişkiyi biri Y cinsinden, diğeri ise derece olarak gösterdiği için kısmi regresyon katsayıları testleri ile kısmi korelasyon katsayıları testleri aynı sonuçları vermektedirler.

Her bir kısmi korelasyon katsayısının anlamlı olup olmadığı,  $F$  veya  $n \geq 30$  ise  $Z$ ,  $n < 30$  durumunda  $t$  testleri ile araştırılmaktadır.

$k$ 'inci bağımsız değişken için hipotez,

$$H_0: \rho_{Yk.1,2,\dots,(k-1),(k+1),\dots,(p-1)} = 0$$

$$H_1: \rho_{Yk.1,2,\dots,(k-1),(k+1),\dots,(p-1)} \neq 0$$

şeklinde kurulur. Böylece, temel hipotez, anakütlerde modeldeki diğer bağımsız değişkenler sabit kalmak kaydıyla,  $k$ 'inci değişkenin  $Y$  üzerindeki marjinal etkisinin anlamlı olmadığını, alternatif hipotez ise bu etkinin anlamlı olduğunu dolayısıyla bu değişkenin modele dahil edilmesi gerektiğini ileri sürmektedir.

$F$  testi kullanıldığında test istatistiği,

$$F^* = \frac{r_{Y.1,2,\dots,(p-1)}^2 - r_{Y.1,2,\dots,(k-1),(k+1),\dots,(p-1)}^2}{(1 - r_{Y.1,2,\dots,(p-1)}^2)/(n-p)}$$

eşitliği ile hesaplanabilir.<sup>50</sup> Karar,

$F^* \leq F_{(\alpha;1;n-p)}$  ise  $H_0$  kabul

$F^* > F_{(\alpha;1;n-p)}$  ise  $H_0$  red şeklinde alınır.

Aynı katsayıların testinde Z veya t testi kullanıldığında ise,

$$t^* = \left( \frac{r_{Y.1,2,\dots,(p-1)}^2 - r_{Y.1,2,\dots,(k-1),(k+1),\dots,(p-1)}^2}{(1 - r_{Y.1,2,\dots,(p-1)}^2)/(n-p)} \right)^{1/2}$$

şeklinde yazılabilir.<sup>51</sup> Karar,

$|t^*| \leq t_{(\alpha/2;n-p)}$  ise  $H_0$  kabul

$|t^*| > t_{(\alpha/2;n-p)}$  ise  $H_0$  red olarak verilecektir.

## 1.6.2. Örnek Verileri ile Tahmin ve Politika Belirleme

Bilindiği gibi, analiz anakütleden tesadüfi olarak seçilen örnek verileri ile yapılmışsa, regresyon ve korelasyon katsayılarının test edilmeleri gerekmektedir.

Basit doğrusal regresyon analizinde, regresyon katsayısı için uygulanan test sonucunda katsayı anlamlı bulunmuş, yani anaküttelede Y ile X arasındaki ilişki geçerli ise, daha önceki bölümlerde bahsedilen normallik, eşit varyanslılık ve otokorelasyonun bulunmaması varsayımlarından sapmaların bulunup bulunmadığının araştırılması gerekmektedir. Bu varsayımların varlığı saptanmış ise, örnek regresyon

<sup>50</sup> Edwards, **a.g.e.**, s.107.

<sup>51</sup> F testinde payın serbestlik derecesi 1 olduğu için, aynı katsayının t testindeki t\* değeri F\* değerinin kareköküne eşittir.



denklemini tahminlerde kullanılabilir. Verilmiş bir  $x_0$  değerine karşılık olan tekil bir  $y_0$  değerinin tahmini için örnek regresyon denklemi en iyi doğrusal yansız tahminleyeni verecektir.

$$y_0' = b_0 + b_1 x_0$$

$Y_0$ 'ın gerçek değeri ile tahmin edilen değeri  $y_0'$  arasındaki fark *öngörü veya ileriye yönelik tahmin hatası* olarak adlandırılır. Bu farkın kaynağı  $Y_0$ 'ın değerinin  $E(Y_0)$ 'a eşit olmayıp regresyon doğrusu üzerinde yer almaması ve örnekleme hatası olabilir. Öngörü hatası aşağıdaki eşitlikle ifade edilir.

$$Y_0 - y_0' = (\beta_0 + \beta_1 X_0 + \varepsilon_0) - (b_0 + b_1 x_0)$$

Burada bağımsız ve normal dağılmış tesadüfi değişkenlerin dağılımı söz konusu olduğu için tahmin hatası da normal dağılmış tesadüfi bir değişkendir. Bu değişkenin ortalaması

$$E(Y_0 - y_0') = \beta_0 + \beta_1 X_0 + E(\varepsilon_0) - E(b_0) - E(b_1) x_0 = 0$$

olup, varyansı hata terimlerinin varyansı ve tahmin varyansının toplamından oluşmaktadır.<sup>52</sup>

$$s^2(Y_0 - y_0') = \sigma^2 + \sigma^2 \left[ \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right]$$

$\sigma$  bilinmediğinden yerine regresyon doğrusunun standart hatası  $s$  kullanıldığında tahmin hatasının standart hatası

---

<sup>52</sup> Güriş, Çağlayan, **a.g.e.**, s. 162-164.

$$s(Y_0 - y_0') = s \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

şeklinde yazılabilir. Bu durumda  $y_0'$  için aralık tahmini belirli bir olasılıkla,

$$\hat{y}_0 = y_0' \pm (t \text{ veya } Z) \cdot s \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

formülüyle yapılır. Formülden de anlaşılacağı gibi örnek birim sayısının artması veya  $x_i$ 'lerin ortalamadan farklarının karelerinin büyümesi tahmin hatasının standart hatasını küçültür. Ayrıca regresyon doğrusunun standart hatasının küçük olması da tahmin aralığını daraltmakta ve böylece örnek regresyon doğrusunun tahmin yeteneğini arttırmaktadır.

Anakütle verileri ile yapılan regresyon analizine paralel olarak, amaç politika belirlemek olduğunda korelasyon katsayısı önem kazanmaktadır. Korelasyon katsayısının testinde temel hipotez reddedilmişse, ilişki anlamlıdır ve bağımsız değişken, bağımlı değişken üzerinde geliştirilecek politikalarda kullanılabilir.

Çoklu doğrusal regresyon analizinde ise, regresyon modelinin anlamlılık düzeyi F testinde olumlu bulursa dahi, başka bir ifade ile çoklu korelasyon katsayısı anlamlı bulursa dahi, modeldeki tüm kısmi regresyon katsayıları için anlamlılık testleri yapılmalıdır. Modele dahil edilecek bağımsız değişkenlerin kombinasyonu ancak bu testler sonucunda belirlenebilmektedir. Nokta tahmininden önce, basit doğrusal regresyon analizinde olduğu gibi, temel varsayımlar olan normallik, otokorelasyonun bulunmaması ve eşit varsayanlılığa ek olarak modelde çoklu doğrusal bağlantının bulunmayışının da gösterilmesi gerekmektedir. Bu adımdan sonra, bağımlı değişkeni anakütlede anlamlı bir şekilde etkileyen bağımsız değişkenlerle regresyon modeli kurularak, verilen  $x_{01}, x_{02}, \dots, x_{0(p-1)}$  değerleri için nokta tahmini  $y_0'$

$$y_0' = b_0 + b_1 x_{01} + b_2 x_{02} + \dots + b_k x_{0(p-1)}$$

olarak yapılacaktır.

İleriye yönelik tahminin aralık tahminin yapılabilmesi için tahmin hatasının dağılımı incelenmelidir.  $(Y_0 - y_0')$  normal dağılmıştır ve ortalaması  $E(Y_0 - y_0') = 0$  olup, standart hatası matris notasyonu ile,

$$s(Y_0 - y_0') = s \sqrt{1 + \frac{1}{n} + X_0' (X'X)^{-1} X_0}$$

şeklinde gösterilebilir.<sup>53</sup>

İleriye yönelik tahminin aralık tahmini, örnekleme dağılımına göre belirli bir hata payı ve Z veya t değerleri ile,

$$\hat{y}_0 = y_0' \pm (t \text{ veya } Z) s \sqrt{1 + \frac{1}{n} + X_0' (X'X)^{-1} X_0}$$

olarak elde edilecektir.<sup>54</sup>

Amaç sadece politika belirleme olduğunda, çoklu korelasyon katsayısı ile birlikte kısmi korelasyon katsayıları da dikkate alınmalıdır. Tüm bu katsayıların testine temel hipotez reddedilmişse, tüm değişkenler bağımlı değişkenin ileriye yönelik planlarında kullanılabilir. Öncelik anakütle verileriyle yapılan analize paralel olarak, kısmi korelasyon katsayısı en yüksek olan bağımsız değişkene verilecektir.

<sup>53</sup> Burada  $X_0' = [1 \ x_{01} \ x_{02} \ \dots \ x_{0(p-1)}]$  vektörüdür.

<sup>54</sup> Orhunbilge, **Uygulamalı Regresyon ve Korelasyon Analizi**, s. 149.

## 2. REGRESYON ANALİZİNİN VARSAYIMLARI, VARSAYIMLARDAN SAPMALARIN BELİRLENMESİ VE ÇÖZÜMLENMESİ

Basit ve çoklu regresyon analizlerinde elde edilen sonuçlar, birinci bölümde sözü edilen varsayımların geçerli olmaları durumunda anlam kazanmakta, ileriye yönelik tahmin ve politika belirlenmesi için kullanılmaktadır. Birden fazla varsayımdan sapma aynı uygulamada geçerli olmakta fakat bu problemlerin hepsini aynı anda saptayabilecek tek bir teste literatürde rastlanmamaktadır.<sup>55</sup> Bu yüzden, bu bölümde, bu varsayımlar ayrı ayrı ele alınmakta ve öncelikle geçerliliklerini belirleyen testlerden yaygın olarak kullanılanları ayrıntılarıyla verilmektedir. Varsayımlardan sapmaların belirlenmesi durumunda En Küçük Kareler yönteminin tahminleyenleri istenen istatistik özellikleri taşımadıkları için kullanılamamakta ve bu yöntem yerine başka alternatif yaklaşımlar sunulmaktadır. Literatürde en çok sözü edilen alternatif yöntemler, her bir varsayım için ele alınıp incelenmektedir.

Dört grupta toplanabilen bu varsayımlar aşağıdaki sırada ele alınabilirler. Bilindiği üzere, ilk üç varsayım hem basit hem de çoklu regresyon modelleri için, sonuncusu ise sadece çoklu regresyon modellerinde söz konusu olabilmektedir.

**Varsayım 1: Hatalar Arasında Otokorelasyon (Autocorrelation) Bulunmaması:** Hata terimleri arasında ilişki olması, başka bir ifade ile hata terimlerinin kendisinden önceki veya sonraki hata terimlerini etkilemesi otokorelasyon olarak tanımlanmaktadır. En Küçük Kareler yönteminin temel varsayımlarından biri, hata terimleri arasında otokorelasyon bulunmamasıdır.

---

<sup>55</sup> Anil K. Bera, Carlos M. Jarque, "Efficient Tests for Normality, Heteroscedasticity and Serial Independence of Regression Residuals: Monte Carlo Evidence", **Economic Letters**, Vol.VII, 1981, s. 313-318.

**Varsayım 2: Eşit Varyanslılık (Homoscedasticity):** Çeşitli X değerlerine karşılık gelen hata terimlerinin aynı varyansa sahip olması durumu eşit varyanslılık olarak adlandırılır. Bu varsayıma göre  $X_i$ 'nin alacağı değerler ile oluşan  $\varepsilon_i$ 'lerin varyansı her bir  $X_i$  için aynı olacaktır. Fakat uygulamalarda bağımsız değişkenler birbirlerinden farklı değerler aldıklarında, eşit varyans varsayımı geçerli olmamaktadır. Bu durum farklı varyanslılık (heteroscedasticity) olarak bilinmektedir.

**Varsayım 3: Normallik (Normality):** Bu varsayıma göre bağımsız değişken veya değişkenlerin tahmini değerlerine ait hata terimlerinin normal dağılması gerekmektedir.

**Varsayım 4: Çoklu Doğrusal Bağlantının (Multicollinearity) Olmaması:** Bağımsız değişkenler arasında yüksek doğrusal ilişkiler bulunması durumu, çoklu doğrusal bağlantı olarak tanımlanmakta ve doğal olarak çoklu regresyon analizinde ortaya çıkmaktadır. Çoklu regresyon analizinde elde edilen sonuçların anlamlı olabilmeleri için, modelde çoklu doğrusal bağlantının bulunmadığının da doğrulanması gerekmektedir.

## 2.1. Otokorelasyonun Belirlenmesi ve Çözümleme Yöntemleri

Hata terimleri arasındaki korelasyon ardışık bir şekilde sıralandığında “otokorelasyon” olarak adlandırılmaktadır.<sup>56</sup> Başka bir tanıma göre ise otokorelasyon terimi, “[zaman serisi verilerindeki gibi] zaman içinde ya da [kesit verilerinde olduğu gibi] uzam<sup>57</sup> (mekan) içinde sıralanan gözlem dizilerinin birimleri arasındaki ilişkidir.”<sup>58</sup>

---

<sup>56</sup> Samprit Chatterjee, Bertram Price, **Regression Analysis By Example**, second edition, New York, John Wiley & Sons, Inc., 1991, s. 151.

<sup>57</sup> “Space” kelimesinin Türkçe karşılığı olarak kullanılmıştır.

<sup>58</sup> Maurice G. Kendall, William R. Buckland, **A Dictionary of Statistical Terms**, New York, Hafner Publishing Company, 1971, s. 8.

Regresyon modeli,  $\varepsilon_i$  hata terimleri arasında otokorelasyon bulunmadığını varsayar. Bu durumda,  $E(\varepsilon_i) = 0$  varsayımıyla

$$E(\varepsilon_i \cdot \varepsilon_j) = E(\varepsilon_i) \cdot E(\varepsilon_j) = 0 \quad i \neq j$$

değerini almaktadır.<sup>59</sup> Bu değer sıfırdan farklı olması durumu ise hata terimleri arasında otokorelasyonun bulunduğunu göstermektedir

### 2.1.1. Otokorelasyonun Nedenleri

Otokorelasyonun varlığı değişik sebeplerden kaynaklanabilir. Başlıca nedenleri:<sup>60</sup>

1. **Süredurum (Inertia):** Zaman serilerinin çoğunda bir hızlanma güdüsü (yükseliş) vardır. Bu yükseliş sırasında, bir serinin herhangi bir dönemdeki değeri bir öncekinden yukarıdadır. Bu durumda, zaman serisi verilerine ilişkin regresyon analizlerinde ardışık gözlemlerin birbirlerini etkilemeleri olasılığı yüksektir. Bu durum genellikle pozitif otokorelasyon olarak ortaya çıkmaktadır.
2. **Dahil Edilmeyen Değişkenler:** Uygulamada diğerlerine göre daha sık karşılaşılan bu durum, araştırmacının çeşitli nedenlerle modele dahil etmesi gerekli bazı açıklayıcı değişken veya değişkenleri ihmal etmesiyle ortaya çıkmaktadır. Bu ihmalin nedeni, modelde yer alacak bağımsız değişken listesinin zaman, ülke, konu, amaç gibi etkenlere göre değişebileceği ve doğal olarak teoriye uygun kesin listelerin bulunmayışıdır. Bazen de ihmalin sebebi gerekli ve yeterli verinin elde edilememesi de olabilmektedir. Dışlanmış değişkenlerin modele dahil edilmesiyle hata terimleri arasında gözlenen ilişki çoğunlukla kaybolur.

<sup>59</sup> Bilindiği gibi hata terimleri bağımsız şans değişkenleri iseler,  $E(\varepsilon_i \cdot \varepsilon_j) = E(\varepsilon_i) \cdot E(\varepsilon_j)$  olarak yazılabilmektedir.

<sup>60</sup> Gujarati, **a.g.e.**, s. 403.

3. **Yanlış Fonksiyon Kalıbı:** Otokorelasyonun nedenlerinden bir diğeri de modelin fonksiyonel şeklinin doğru belirlenmemesidir. Fonksiyonel ilişkinin belirlenmesinde arařtırmacının grř ve tecrbeleri ve konu ile ilgili teoriler etkili olmaktadır. Fakat deęiřken seęiminde olduęu gibi, burada da her zaman kesinlik sz konusu olamayacaęından modelin fonksiyon tipi yanlış seęilebilmektedir.
4. **lme Hataları:** Bir diğeri nemli neden de verilerle ilgili lme hatalarıdır. Uygulamalarda kullanılan veriler, istatistikiler veya ekonometristler tarafından dzenlenirken eřitli nedenlerle lme veya dzenleme hataları yapılabilmektedir. Bu durum, modelin hata terimlerinde sapmalara ve dolayısıyla otokorelasyona neden olabilmektedir.

### **2.1.2. Otokorelasyonun En Kk Kareler Yntemiyle Elde Edilen Tahminleyenlere Etkisi**

Basit ve oklu regresyon analizinde, diğeri tm varsayımlar geerliliklerini srdrrken, otokorelasyon varsayımının geerli olmaması durumunda En Kk Kareler yntemi ile elde edilen tahminleyenlerin ve bunların varyanslarının otokorelasyondan nasıl etkilendiklerinin incelenmesi gerekmektedir. Yapılan arařtırmalarda elde edilen sonular ve bunların regresyon modellerine etkileri ařaęıdaki gibi zetlenebilir

- En Kk Kareler tahminleyenleri, tutarlı ve sapmasızdırlar ancak etkin (en kk varyanslı) ve asimtotik etkin deęillerdir.<sup>61</sup> Bu nedenle doęrusal olmalarına karřın en iyi doęrusal yansız tahminleyen (BLUE) deęillerdir.

---

<sup>61</sup> İspat iin bkz.: Griř, aęlayan, **a.g.e.**, s. 407-427.

- Hata terimlerinin varyansının tahminleneni sapmalıdır. Bu durum parametre varyanslarının tahminleyenlerinin de sapmalı olmasına neden olacaktır.<sup>62</sup> Bu sapma genellikle pozitif otokorelasyona sahip olan zaman serilerinde negatif olacak ve bu seriler için varyanslar gerçek değerlerinden küçük tahmin edileceklerdir.
- Parametre tahminleyenlerinin varyanslarının gerçek değerlerinden farklı olması yapılacak t, F testleri ile aralık tahminlerini etkilerler. Bu testler, otokorelasyon gözardı edilerek uygulandıklarında, büyük bir olasılıkla tahmin edilen regresyon katsayılarının istatistik anlamlılığı hakkında yanlış sonuçlar vereceklerdir. İktisadi zaman serilerinde genellikle parametre varyansları gerçek değerlerinden küçük tahmin edildiklerinden t ve F test istatistikleri gerçek değerlerinden daha büyük hesaplanabilmektedir. Bu ise doğru olmamasına rağmen alternatif hipotezin kabulüne neden olmaktadır. Parametre tahminleyenlerinin aralık tahmini de yukarıda sözü edilen nedenlerle gerçek değerlerini yansıtmayacaklardır. Parametre varyanslarının gerçek değerlerinden küçük hesaplanması tahmin aralığının daralmasına neden olacaktır. Bu bulgu Kmenta'nın gösterdiği gibi örnek sayısı sonsuza giderken bile olasıdır.<sup>63</sup>
- Otokorelasyon durumunda hata terimlerinin varyansı gibi kareleri toplamı da gerçeği yansıtmayacak ve bu durum belirlilik katsayısını da etkileyecektir. Zaman serilerinde, hata karelerinin gerçek değerinden sapmanın negatif olması, belirlilik katsayısının gerçek değerinden büyük hesaplanmasına neden olmaktadır.

### 2.1.3. Otokorelasyon Katsayısı

---

<sup>62</sup> Bilindiği gibi parametre varyanslarının tahminleyenlerinin belirlenmesi için, hata terimlerinin varyansı kullanılmaktadır.

<sup>63</sup> Jan Kmenta, **Elements of Econometrics**, New York, Macmillan Publishing, 1971, s. 274-275.



Otokorelasyon katsayısı herhangi bir deęişken için hesaplanıp, bu deęişkenin gözlem deęerlerinin kendinden önceki veya sonraki gözlem deęerleri ile ilişkisi olup olmadığını, varsa derecesini bildirir.<sup>64</sup> Korelasyon katsayısı gibi, otokorelasyon katsayısı da kovaryansın, deęişkenlerin standart sapmasına bölünmesi ile hesaplanmakta ve  $\pm 1$  arasında deęerler alabilmektedir. Bu katsayının, pozitif deęer olarak 1'e yaklaşması verilerde pozitif otokorelasyon, -1'e yaklaşması ise negatif otokorelasyon bulunduęunun göstergesidir.

Regresyon modelinde, hata terimleri için kovaryans, otokorelasyon bulunmaması durumunda

$$\text{Cov}(\varepsilon_i, \varepsilon_j) = 0 \quad i \neq j$$

olacaktır.<sup>65</sup>  $i = j$  olması durumunda ise, kovaryans varyansa eşittir.

$$\text{Cov}(\varepsilon_i, \varepsilon_i) = \frac{\sum (\varepsilon_i - \bar{\varepsilon})(\varepsilon_i - \bar{\varepsilon})}{N} = \frac{\sum (\varepsilon_i - \bar{\varepsilon})^2}{N} = \sigma_i^2$$

Hata terimleri için anakütle otokorelasyon katsayısı  $\rho$  ile gösterilip,  $\text{Cov}(\varepsilon_t, \varepsilon_{t-1})$  teriminin  $\varepsilon_t$  ve  $\varepsilon_{t-1}$  terimlerinin standart sapmalarına bölünmesiyle elde edilir. Bu katsayı örnek verileriyle,

$$\hat{\rho} = \frac{\sum_{t=2}^n (e_t - \bar{e}_t)(e_{t-1} - \bar{e}_{t-1})}{\sqrt{\sum_{t=2}^n (e_t - \bar{e}_t)^2 \sum_{t=2}^n (e_{t-1} - \bar{e}_{t-1})^2}}$$

formülüyle tahmin edilir. Hata terimlerinin ortalaması sıfır olduęu ve

$$\sum_{t=2}^n (e_t)^2 \cong \sum_{t=2}^n (e_{t-1})^2$$

<sup>64</sup> Güriş, Çaęlayan, **a.g.e.**, s. 396

<sup>65</sup> Bilidięi gibi kovaryans, deęişkenler arasında ilişki bulunup bulunmadığını göstermektedir.

olarak kabul edilebileceği için, daha basit bir formülle otokorelasyon katsayısı aşağıdaki gibi hesaplanmaktadır.

$$\hat{\rho} = \frac{\sum_{t=2}^n e_t e_{t-1}}{\sum_{t=2}^n e_t^2}$$

#### 2.1.4. Otokorelasyonun Belirlenmesi

Otokorelasyonun bulunup bulunmadığının saptanması için kullanılan en yaygın testler aşağıdaki alt başlıklarda ele alınmaktadır.

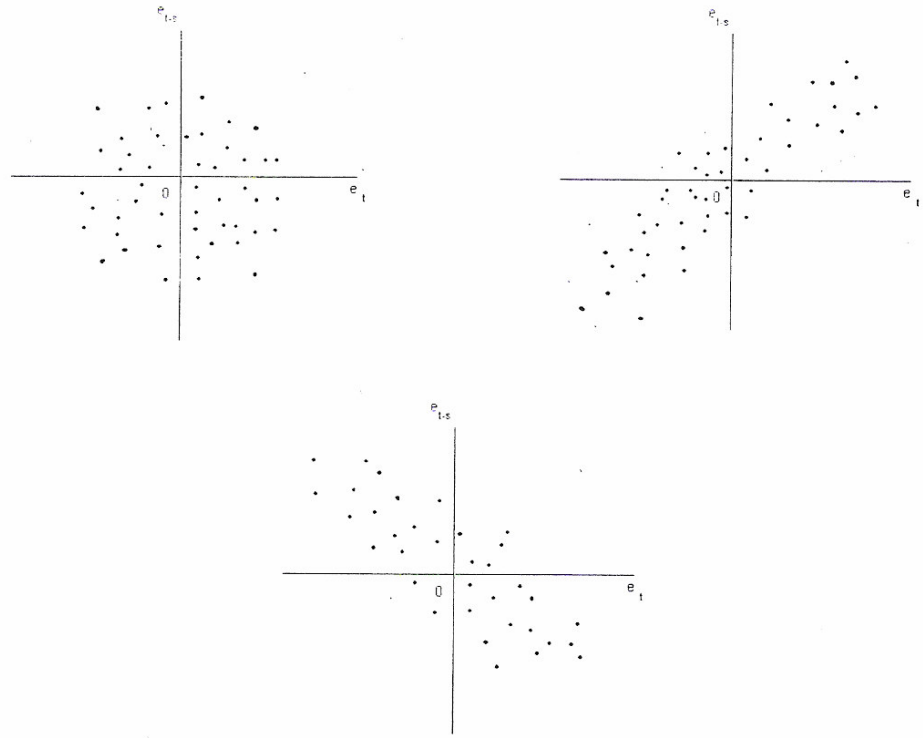
##### 2.1.4.1. Grafik Yöntem

Klasik modelin otokorelasyon yoktur varsayımını incelemek için, anakütle  $\varepsilon_i$  hata terimleri yerine, En Küçük Kareler yöntemiyle örnek kütlede elde edilen  $e_i$  terimlerinin görsel olarak incelenmesi çoğu zaman bu varsayımın varlığı konusunda bir fikir vermektedir.<sup>66</sup> En kolay yöntemlerden biri olmasına karşın, grafiklerin ölçeklendirilmesinden doğabilecek hatalar araştırmacıyı yanıltmakta ve kesin kararlar verilememesine neden olmaktadır.

Tahmin edilen model için hata terimlerinin birbirleri ile ilişkilerini gösteren grafiklerle otokorelasyonun varlığı incelenmektedir. Aşağıdaki şekilde ilk grafik otokorelasyonun bulunmadığını, ikinci grafik pozitif, üçüncüsü ise negatif otokorelasyonun varlıklarını açıkça göstermektedirler.

---

<sup>66</sup> G. S. Maddala, **Introduction to Econometrics**, second edition, New York, Macmillan, 1992, s. 480.

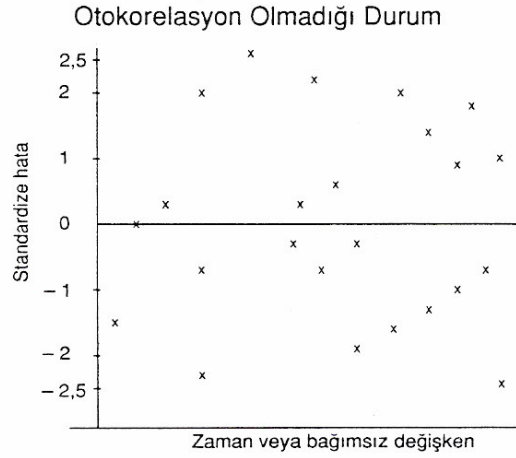
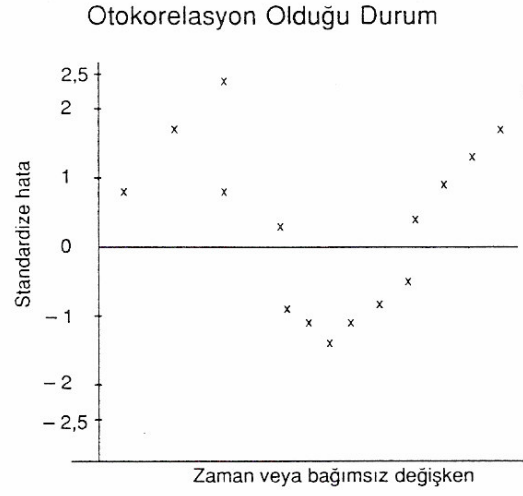


**Şekil (2.2):Hata Terimlerinin Birbirleriyle İlişisini Gösteren Grafikler**

Ayrıca, hata terimlerinin her birinin modelin standart hatasına bölünmesiyle elde edilen standardize hataların bağımsız değişken veya zamana göre çizilen grafiklerle de bu varsayımın varlığı saptanabilmektedir.<sup>67</sup> Bu grafiklerde pozitif hatalar pozitif hataları, negatif hatalar negatif hataları takip ediyorsa pozitif otokorelasyondan, pozitif hatalar negatif hataları takip ediyorlarsa negatif otokorelasyondan söz edilebilmekte, hataların tesadüfi dağılımı ise otokorelasyonun bulunmadığını göstermektedir.<sup>68</sup>

<sup>67</sup> Orhunbilge, **Uygulamalı Regresyon ve Korelasyon Analizi**, s. 216-217.

<sup>68</sup> Draper, Smith, **a.g.e.**, s. 153-156.



**Şekil (2.3): Standardize Hataların Dağılımını Gösteren Grafikler**

### 2.1.4.2. Akış Sayısı Testi (Number of Runs Test)

Akış sayısı testi bazen de Geary<sup>69</sup> testi adı verilen parametrik olmayan bu test, tesadüfiliğin araştırılması için kullanılmaktadır. Regresyon analizinde de, modelden tahmin edilen hata terimleri arasındaki ilişkinin tesadüfi olup olmadığı bu test ile incelenebilmekte ve otokorelasyonun bulunup bulunmadığına karar verilebilmektedir.

<sup>69</sup> Roy C. Geary, "Relative Efficiency of Count of Sign Changes for Assessing Residual Autocorrelation in Least Regression", **Biometrika**, Vol.LVII, 1970, s. 123-127.

Bu testte, hata terimlerinin sayısal değerleri değil, işaretleri dikkate alınarak, aynı işaretli terimlerin birbirlerini izlemelerinin tesadüfi olup olmadığı araştırılmaktadır. Bu amaçla, hata terimlerinin işaretleri bir dizi oluşturacak şekilde sıralanmakta, dizideki her bir işaret değişimi<sup>70</sup> sayılarak *sıra sayısı* adı verilen değişkenin değeri belirlenmektedir. Sıra sayısının tesadüflüğünün olasılığı ile dizilişin tesadüfi olup olmadığı belirlenmektedir. Tesadüflük varsayımı altında sıra sayısı  $k$ ,

ortalaması

$$\mu_k = \frac{2n_1n_2}{n_1 + n_2} + 1$$

ve varyansı

$$\sigma_k^2 = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}$$

olan normal bir dağılım gösterir.<sup>71</sup> Bu formülde,

$n_1$ =dizideki (+) işaretli terimlerin sayısını

$n_2$ =dizideki (-) işaretli terimlerin sayısını

göstermektedir.

Regresyon analizinde de, otokorelasyonun varlığını araştırmak için Akış Sayısı testi (Runs Test) kullanıldığında temel hipotez, hata terimlerinin birbirlerini etkilemediklerini yani otokorelasyonun bulunmadığını ve hata terimlerin işaretlerinin sıralamasının tesadüfi olduğunu ifade ederken, alternatif hipotez tesadüfi olmadığını savunmaktadır. Hipotezler,

$$H_0: \rho=0$$

$$H_1: \rho \neq 0$$

<sup>70</sup> (+)'dan (-)'ye veya (-)'den (+)'ya

<sup>71</sup> burada  $n_1 > 10$  ve  $n_2 > 10$  varsayımı vardır.

olarak oluşturulmaktadır. Test istatistiği

$$Z^* = \frac{k - \mu_k}{\sigma_k}$$

formülüyle hesaplanmakta ve karar

$|Z^*| \leq Z_{\alpha/2}$  ise  $H_0$  kabul

$|Z^*| > Z_{\alpha/2}$  ise  $H_0$  red

şeklinde alınmaktadır.

Parametrik olmayan bu testte hata terimlerinin büyüklükleri, ilişkinin ölçüsünü belirlerken dikkate alınmamakta ve bu yüzden elde edilen sonuçların doğruluğu tam olarak saptanamamaktadır.

### 2.1.4.3. Durbin-Watson Testi

Otokorelasyonu sınamada en yaygın kullanılan test, Durbin-Watson tarafından geliştirilen testtir.<sup>72, 73</sup>

Bu testte de, Akış Sayısı testinde olduğu gibi temel hipotez otokorelasyonun bulunmamasını ( $H_0 : \rho=0$ ) ifade etmektedir. Test istatistiği d, En Küçük Kareler hata terimleri  $e_t$ 'ler ile aşağıdaki gibi,

---

<sup>72</sup> James Durbin, Graham Watson, "Testing for Serial Correlation in Least Squares Regression-I", **Biometrika**, Vol. XXXVII, 1950, s. 409-428.

<sup>73</sup> James Durbin, Graham Watson, "Testing for Serial Correlation in Least Squares Regression-II", **Biometrika**, Vol. XXXVIII, 1951, s. 159-178.

$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} \quad (2.1)$$

hesaplanmaktadır. Bu testin diğer testlere göre en büyük üstünlüğü test istatistiği d'nin, En Küçük Kareler yöntemiyle zaten hesaplanmış olan hata terimlerine bağlı olması, dolayısıyla kolay hesaplanmasıdır. Kullanımındaki kolaylığının yanısıra, d istatistiğinin dayandığı varsayımlara da dikkat edilmesi gerekmektedir. Bu varsayımlar,<sup>74</sup>

- regresyon modeli sabit terim içermektedir
- hata terimleri birinci dereceden otoregressif bir seri ile ifade edilmektedir<sup>75</sup>

$$\varepsilon_t = \rho\varepsilon_{t-1} - u_t$$

- açıklayıcı değişkenler olasılıklı değillerdir, ya da yinelenen örnek verilerinde değişmezler
  - regresyon modeli, bağımlı değişkenin gecikmeli değer(ler)ini bağımsız değişken olarak almaz
  - verilerde eksik gözlem yoktur
- olarak sıralanabilirler.

(2.1) formülüyle hesaplanan d test istatistiği, otokorelasyon katsayısı ile de ifade edilmektedir.<sup>76</sup>

<sup>74</sup> Gujarati, **a.g.e.**, s. 421.

<sup>75</sup> Burada  $u_t$  normal dağılan tesadüfi bir terimdir.

<sup>76</sup> (2.1) formülünde payın karesi açıldığında,  $d = \frac{\sum_{t=2}^n e_t^2 + \sum_{t=2}^n e_{t-1}^2 - 2\sum_{t=2}^n e_t e_{t-1}}{\sum_{t=1}^n e_t^2}$  olarak ifade

edilebilir. Payın ilk iki teriminde sadece bir gözlemlik fark olduğu dikkate alınırsa, bu iki terimin

$$d \cong 2(1 - \hat{\rho})$$

Bilindiği gibi  $\hat{\rho}$ 'nun alt sınırı -1, üst sınırı ise +1'dir. Bu durumda otokorelasyon bulunmadığında  $d=2$  olmakta, tam pozitif otokorelasyon bulunduğu  $d=0$ , tam negatif otokorelasyon durumunda ise  $d=4$  değerlerini almakta ve böylece  $d$ 'nin geçerli olduğu değer aralığı  $0 \leq d \leq 4$  olarak belirlenmektedir.

Test istatistiği  $d$ , bağımsız değişken sayısı, gözlem sayısı ve bağımsız değişkenlerin değerlerine bağlı olarak hesaplanmaktadır. Bu nedenle  $d$ 'nin dağılımı örnekten örneğe değişmekte,  $t$  ve  $F$  testlerinde olduğu gibi tam bir dağılım belirlenmemekte, ancak Durbin ile Watson tarafından geliştirilen  $d_L$  alt sınır ve  $d_U$  üst sınır değerleri ile otokorelasyonun varlığına dair kararlar verilmektedir. Genellikle %1 ve %5 anlamlılık seviyeleri için oluşturulan tablolarda<sup>77</sup> bu kritik değerler, bağımsız değişken veya değişkenlerin değerinden bağımsız olarak, sadece gözlem sayısı ve bağımsız değişken sayısı ile belirlenmektedirler.<sup>78</sup>

Dubin-Watson testinde, temel hipotez otokorelasyonun bulunmadığını ifade ederken, alternatif hipotez tek veya çift taraflı olarak oluşturulmaktadır.

$$H_0: \rho=0$$

$$H_1: \rho \neq 0 \text{ veya } \rho > 0 \text{ (pozitif otokorelasyon) veya } \rho < 0 \text{ (negatif otokorelasyon)}$$

Modelde hesaplanan hata terimleriyle  $d$  test istatistiği hesaplandıktan sonra belirlenen anlamlılık seviyesiyle tablodan kritik alt ve üst değerler ( $d_L$  ve  $d_U$ )

---

yaklaşık olarak birbirine eşit olduğu varsayılabilir. Bu durumda  $d \cong 2(1 - \frac{\sum e_t e_{t-1}}{\sum e_t^2})$  olacak, ve

ikinci terim otokorelasyon katsayısına eşit olduğu için  $d \cong 2(1 - \hat{\rho})$  yazılabilecektir.

<sup>77</sup> Eugene Savin, Kenneth White, "The Durbin-Watson Test for Serial Correlation with Extreme Sample Sizes or Many Regressors", *Econometrica*, Vol.XXIX, 1977, s.1989-1996.

<sup>78</sup> Durbin-Watson tabloları  $n > 14$  için düzenlenmiştir.  $n < 15$  durumunda bu test uygulanamamaktadır.



belirlenmekte ve karar hipotezin tek yönlü veya çift yönlü olmasına bağlı olarak alınmaktadır.<sup>79</sup>

a) Tek Taraflı Testler:

1. Pozitif Otokorelasyon Testi: ( $H_0: \rho=0$ ,  $H_1: \rho>0$ )

$d \leq d_L$  ise  $H_0$  red, pozitif otokorelasyon vardır.

$d \geq d_U$  ise  $H_0$  kabul, pozitif otokorelasyon yoktur.

$d_L < d < d_U$  ise bu konuda karar verilemez.

2. Negatif Otokorelasyon Testi: ( $H_0: \rho=0$ ,  $H_1: \rho<0$ )

$d \geq 4 - d_L$  ise  $H_0$  red, negatif otokorelasyon vardır.

$d \leq 4 - d_U$  ise  $H_0$  kabul, negatif otokorelasyon yoktur.

$4 - d_U < d < 4 - d_L$  ise bu konuda karar verilemez.

b) Çift Taraflı Test: ( $H_0: \rho=0$ ,  $H_1: \rho \neq 0$ )

$d \leq d_L$  veya  $d \geq 4 - d_L$  ise  $H_0$  red, otokorelasyon vardır.

$d_U \leq d \leq 4 - d_U$  ise  $H_0$  kabul, otokorelasyon yoktur.

$d_L < d < d_U$  veya  $4 - d_U < d < 4 - d_L$  ise bu konuda karar verilemez.

Durbin-Watson testi çok yaygın olarak kullanılmasına karşın, d test istatistiği kararsızlık bölgesine düştüğünde otokorelasyonun varlığı hakkında kesin bir karar alınamaz. Bu sorunu çözmek için araştırmacılar d sınavına çeşitli uyarlamalar geliştirmişlerdir. Bu çalışmaların büyük bir çoğunluğunda  $d_U$  üst sınırının yaklaşık olarak gerçek anlamlılık sınırı olduğu bulunmuştur.<sup>80</sup> Dolayısıyla hesaplanan d

<sup>79</sup> Ertek, a.g.e., s. 256-258.

<sup>80</sup> Henri Theil, *Principles of Econometrics*, New York, John Wiley & Sons, 1971, s. 201.

değeri kararsızlık bölgesine düşerse, aşağıdaki şekilde *uyarlanmış d sınaması* kullanılabilir.<sup>81</sup> Belirlenen  $\alpha$  anlamlılık seviyesinde,

- 1)  $H_0: \rho=0, H_1: \rho>0: d < d_U$  ise  $H_0$  red, pozitif otokorelasyon vardır.
- 2)  $H_0: \rho=0, H_1: \rho<0: d > 4-d_U$  ise  $H_0$  red, negatif otokorelasyon vardır.
- 3)  $H_0: \rho=0, H_1: \rho\neq 0: d < d_U$  veya  $d > 4-d_U$  ise  $H_0$  red, otokorelasyon vardır.

#### 2.1.4.4. Farebrother Testi

Bilindiği gibi Durbin-Watson testinin varsayımlarından biri de modelde sabit terimin bulunması ile ilgilidir. Modelde sabit terim bulunmaması durumunda Farebrother<sup>82</sup> tarafından geliştirilen bu test kullanılabilir.

Bu testte hipotezler çift yönlü olarak kurulmakta ve test istatistiği  $d$ , Durbin-Watson testindeki formül (2.1) ile hesaplanmaktadır.

$$H_0: \rho=0$$

$$H_1: \rho\neq 0$$

Otokorelasyonun varlığına karar verebilmek için belirlenen  $\alpha$  anlamlılık seviyesinde, Durbin-Watson tablosundan  $d_U$  değeri belirlenir.  $d_L$  değeri için ise Farebrother tablosu'na başvurularak, aynı anlamlılık seviyesi için bir  $d_M$ <sup>83</sup> değeri belirlenir. Böylece, kritik değerlerden üst sınır değişmemekte, ancak alt sınır için daha küçük bir değer elde edilmektedir. Kararlar, Durbin-Watson testinde olduğu gibi alınmaktadır.

---

<sup>81</sup> Akın, **a.g.e.**, s. 472.

<sup>82</sup> Richard W. Farebrother, "The Durbin-Watson Test For Serial Correlation When There Is No Intercept in the Regression", **Econometrica**, Vol.XLVIII, 1980, s. 1553-1563.

<sup>83</sup>  $d_M < d_L$  olarak belirlenmektedir.

a) Tek Taraflı Testler:

1. Pozitif Otokorelasyon Testi:  $(H_0: \rho=0, H_1: \rho>0)$

$d \leq d_M$  ise  $H_0$  red, pozitif otokorelasyon vardır.

$d \geq d_U$  ise  $H_0$  kabul, pozitif otokorelasyon yoktur.

$d_M < d < d_U$  ise bu konuda karar verilemez.

2. Negatif Otokorelasyon Testi:  $(H_0: \rho=0, H_1: \rho<0)$

$d \geq 4 - d_M$  ise  $H_0$  red, negatif otokorelasyon vardır.

$d \leq 4 - d_U$  ise  $H_0$  kabul, negatif otokorelasyon yoktur.

$4 - d_U < d < 4 - d_M$  ise bu konuda karar verilemez.

b) Çift Taraflı Test:  $(H_0: \rho=0, H_1: \rho \neq 0)$

$d < d_M$  veya  $d > 4 - d_M$  ise  $H_0$  red, otokorelasyon vardır.

$d_U < d < 4 - d_U$  ise  $H_0$  kabul, otokorelasyon yoktur.

$d_M < d < d_U$  veya  $4 - d_U < d < 4 - d_M$  ise bu konuda karar verilemez.

### 2.1.4.5. Von Neumann Oranı Testi

Durbin-Watson testi kararsızlık bölgesine düştüğünde alternatif test olarak kullanılabilen bu testte, test istatistiği ardışık hata terimleri ile hesaplanmakta ve aşağıdaki formülle gösterilmektedir.<sup>84</sup>

$$VN = \frac{\sum_{t=p+2}^n (e_t - e_{t-1})^2 / (n - p - 1)}{\sum_{t=p+1}^n e_t^2 / (n - p)}$$

<sup>84</sup> Güriş, Çağlayan, a.g.e., s. 469-471.

Aynı formül ardışık hata terimleriyle hesaplanan d istatistiği ile de ifade edilebilir. Bu durumda, d test istatistiği,

$$d_{VN} = \frac{\sum_{t=p+2}^n (e_t - e_{t-1})^2}{\sum_{t=p+1}^n e_t^2}$$

olacağından, Von Neumann Oranı test istatistiği,

$$VN = d_{VN} \frac{n-p}{n-p-1}$$

formülüyle hesaplanabilir. Regresyon modellerinde ardışık hata terimleri hesaplandığında modelin tahmin edilen parametresi p kadar gözlem kaybolacaktır. Bu nedenle, yukarıdaki formüllerde payda (n-p-1) ve paydada ise (n-p) kadar bağımsız gözlem yer almaktadır.

Buradan hesaplanan test istatistik değeri Von Neumann Oranı tablosundan bulunan alt ( $VN_L$ ) ve üst ( $VN_U$ ) değerler ile karşılaştırılır<sup>85</sup> ve kararlar aşağıdaki şekilde alınır:

$VN \leq VN_L$  ise pozitif otokorelasyon vardır.

$VN \geq VN_U$  ise negatif otokorelasyon vardır.

$VN_L < VN < VN_U$  ise otokorelasyon yoktur.

---

<sup>85</sup>  $n \leq 60$  olduğunda Von Neumann Oranı tablosu,  $n > 60$  olduğunda ise dağılım normale yaklaşacağından standart normal dağılım tablosu kullanılmaktadır.

## 2.1.5. Otokorelasyonun Ortadan Kaldırılması için Çözüm Yolları

Otokorelasyonun varlığına bir önceki bölümde bahsedilen testlerden biriyle karar verilmişse, regresyon analiziyle elde edilen tahminleyenler yansız, fakat etkin değildirler. Bu durumda, değişkenlere çeşitli dönüştürme yöntemleri uygulanarak otokorelasyondan arındırılmış regresyon modelleri elde edilmektedir. Bu yöntemler, otokorelasyonun yapısının bilinip bilinmemesine bağlı olarak iki gruba ayrılmaktadır.

### 2.1.5.1. Otokorelasyon Katsayısı $\rho$ 'nun Bilinmesi Durumu

#### 2.1.5.1.1. Genelleştirilmiş Farklar Yöntemi

Regresyon modelinde otokorelasyon katsayısının bilindiği varsayımına dayanan bu yöntemde değişkenlere, kendi değerlerinden, bir önceki gözleme ait değerlerinin otokorelasyon katsayısı ile çarpılıp çıkarılması ile dönüşümler uygulanmaktadır.<sup>86</sup>

Regresyon modeli<sup>87</sup> i. ve (i-1). gözlem değerleri için,

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots + b_{p-1} x_{i,p-1} + e_i$$

ve

$$y_{(i-1)} = b_0 + b_1 x_{(i-1),1} + b_2 x_{(i-1),2} + \dots + b_{p-1} x_{(i-1),p-1} + e_i$$

olarak yazılmaktadır. İkinci denklem  $\rho$  ile çarpılıp birinci denklemden çıkartıldığında aşağıdaki denklem elde edilecektir.

---

<sup>86</sup> Gujarati, **a.g.e.**, s. 427.

<sup>87</sup> Burada örnek verileri ile çalışıldığı varsayılmaktadır. Anakütle verileri için de tüm sonuçlar geçerlidir.

$$y_i - \rho y_{(i-1)} = b_0(1 - \rho) + b_1(x_{i1} - \rho x_{(i-1),1}) + b_2(x_{i2} - \rho x_{(i-1),2}) + \dots + b_{p-1}(x_{i,p-1} - \rho x_{(i-1),p-1}) + (e_i - \rho e_{i-1}) \quad (2.2)$$

Genelleştirilmiş fark denklemi<sup>88</sup> olarak bilinen bu denklemde değişkenler,

$$y_i^* = y_i - \rho y_{(i-1)}$$

$$b_0^* = b_0(1 - \rho)$$

$$x_{i1}^* = x_{i1} - \rho x_{(i-1),1}$$

$$x_{i2}^* = x_{i2} - \rho x_{(i-1),2}$$

.

.

$$x_{i,p-1}^* = x_{i,p-1} - \rho x_{(i-1),p-1}$$

$$e_i^* = e_i - \rho e_{i-1}$$

olarak tanımlanırsa, dönüştürülmüş değişkenler ile denklem

$$y_i^* = b_0^* + b_1^* x_{i1}^* + b_2^* x_{i2}^* + \dots + b_{p-1}^* x_{i,p-1}^* + e_i^*$$

şeklinde yeniden yazılabilir. Bir gözlem kaybı<sup>89</sup> söz konusu olan bu denklemden elde edilecek tahminleyenler en iyi doğrusal yansız tahminleyenlerdir. Büyük örnek verileri için fazla sorun yaratmayan gözlem kaybı, küçük örneklerde bulguları önemli ölçüde farklılaştırmaktadır<sup>90</sup>. Kaybolan ilk gözlem değerleri için Prais ve Winston<sup>91</sup> tarafından geliştirilen dönüşümler uygulanmaktadır. Bu dönüşümlere göre bağımlı ve bağımsız değişkenlerin ilk gözlem modelleri aşağıdaki şekilde modele dahil edilmektedirler.

<sup>88</sup> Yaklaşık denklem adı da verilmektedir.

<sup>89</sup> İlk gözlem değerleri kaybolmaktadır.

<sup>90</sup> Jack Johnston, **Econometric Methods**, 3. edition, McGraw-Hill, New York, 1984, s. 321-323.

<sup>91</sup> S. Prais, C. Winston, "Trend Estimation and Serial Correlation", **Discussion Paper 383**, Cowles Commission, Chicago, 1954.

$$\begin{aligned}
y_1^* &= y_1 \sqrt{1-\rho^2} \\
x_{11}^* &= x_{11} \sqrt{1-\rho^2} \\
x_{12}^* &= x_{12} \sqrt{1-\rho^2} \\
&\cdot \\
&\cdot \\
x_{1,(p-1)}^* &= x_{1,p-1} \sqrt{1-\rho^2}
\end{aligned}$$

Genelleştirilmiş farklar yönteminin uygulanması ile regresyon modelinde uygun testler ile varlığı saptanmış olan otokorelasyonun ortadan kaldırılması beklenmektedir. Ancak otokorelasyonun ortadan kaldırıldığına emin olmak, için dönüştürülmüş denklemler ile kurulan yeni modelin de test edilmesi gerekmektedir.

### 2.1.5.1.2. İlk Farklar Yöntemi

İlk farklar yöntemi, genelleştirilmiş fark yönteminde  $\rho=1$  varsayımını yapmaktadır. Bu durumda, otokorelasyonun etkisini ortadan kaldırmak için değişkenlere, şimdiki değerlerinden bir önceki gözlem değerleri çıkarılarak dönüşümler uygulanmaktadır. Bir başka ifade ile yeni değişkenler,

$$\begin{aligned}
y_i^* &= y_i - y_{(i-1)} \\
x_{i1}^* &= x_{i1} - x_{(i-1),1} \\
x_{i2}^* &= x_{i2} - x_{(i-1),2} \\
&\cdot \\
&\cdot \\
x_{i,p-1}^* &= x_{i,p-1} - x_{(i-1),p-1} \\
e_i^* &= e_i - e_{i-1}
\end{aligned}$$

olarak tanımlanmaktadır. İlk farklar yöntemi ile model

$$y_i^* = b_1^* x_{i1}^* + b_2^* x_{i2}^* + \dots + b_{p-1}^* x_{i,p-1}^* + e_i^*$$

denklemlerle tahmin edilir. Genelleştirilmiş farklar yönteminde olduğu gibi, bu modelde de otokorelasyonun varlığının tekrar araştırılması gerekmektedir.,

### 2.1.5.2. Otokorelasyon Katsayısı $\rho$ 'nun Bilinmemesi Durumu

Uygulamalarda, otokorelasyon katsayısının bilinmesi durumuna ender rastlanmakta ve bu yüzden tahmin edilmesi gerekmektedir. Otokorelasyonun bulunması durumunda, En Küçük Kareler yöntemiyle bulunan hata terimleriyle hesaplanan  $\rho$  yanlış olacak ve gerçek değeri vermeyecektir. Bu durumda genellikle uygulanan yöntem,  $\rho$ 'nun aşağıda söz edilen yöntemlerden biriyle tahmin edilmesi ve bu değerle genelleştirilmiş farklar yönteminin uygulanmasıdır.

#### 2.1.5.2.1. İki Aşamalı Durbin Yöntemi

Bu yöntemde<sup>92</sup> otokorelasyon katsayısı, genelleştirilmiş fark denkleminde yararlanılarak tahmin edilir. (2.2) denkleminde  $y_i$  terimi sol tarafta yalnız bırakıldığında aşağıdaki eşdeğer denklem elde edilir.

$$y_i = b_0(1-\rho) + \rho y_{(i-1)} + b_1(x_{i1} - \rho x_{(i-1),1}) + b_2(x_{i2} - \rho x_{(i-1),2}) + \dots + b_{p-1}(x_{i,p-1} - \rho x_{(i-1),p-1}) + (e_i - \rho e_{i-1})$$

Bu model çoklu regresyon modeli olarak ele alınıp parametreleri tahmin edildiğinde,  $y_{(i-1)}$  değişkeninin katsayısı olan  $\rho$ 'nun tahmini değeri  $\hat{\rho}$  elde edilecektir. Bu, sapmalı olmasına karşın tutarlı bir tahmindir.

<sup>92</sup> J. Durbin, "Estimation of Parameters in Time-Series Regression Models", *Journal of the Royal Statistical Society*, Vol.XXII, 1960, s. 139-153.



### 2.1.5.2.2. Durbin-Watson d İstatistiğine Dayanan Tahmin Yöntemi

Bilindiği gibi, Durbin-Watson testinde, d test istatistiği otokorelasyon katsayısına bağlı olarak da ifade edilmektedir.

$$d \cong 2(1 - \hat{\rho})$$

Bu eşitlikten hareketle otokorelasyon katsayısı, tahmin edilmiş d istatistik değeriyle hesaplanabilmektedir.

$$\hat{\rho} \cong 1 - \frac{d}{2}$$

Ancak bu eşitlik büyük örnek verileri için yaklaşık bir değer olarak kabul edilmekte, küçük örnekler ile çalışıldığında ise daha iyi bir sonuç elde etmek için Theil-Nagar<sup>93</sup> formülü kullanılmaktadır.

$$\hat{\rho} = \frac{n^2(1 - \frac{1}{2}d) + p^2}{n^2 - p^2}$$

### 2.1.5.2.3. Cochrane-Orcutt Yöntemi

Yaygın olarak kullanılan bu yöntemde, otokorelasyon katsayısı iterasyon yöntemiyle elde edilmektedir. Çoklu regresyon modeli örnek verileriyle ele alınıp,

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots + b_{p-1} x_{i,p-1} + e_i$$

---

<sup>93</sup> Henri Theil, A. Nagar, "Testing the Independence of Regression Disturbances" **Journal of the American Statistical Association**", Vol.LII, 1961, s. 793-806.

denklemleriyle ifade edildiğinde, En Küçük Kareler yöntemiyle tahmin edilen hata terimlerinin birinci dereceden otoregressif bir seriden türediği varsayılmaktadır.

$$e_t = \hat{\rho} e_{t-1} - u_t \quad (2.3)$$

Yöntemin ilk aşamasında, hata terimleri En Küçük Kareler yöntemiyle elde edilmektedir. Daha sonra (2.3) ile belirtilen regresyon denkleminde bu terimler veri olarak kullanılarak otokorelasyon katsayısının tahmini değeri  $\hat{\rho}$  elde edilmektedir. Bilindiği gibi bu değer,

$$\hat{\rho} = \frac{\sum_{t=2}^n e_t e_{t-1}}{\sum_{t=2}^n e_t^2}$$

formülüyle hesaplanabilir.

Bu tahmini değer genelleştirilmiş farklar denkleminde kullanılarak, modelin yeni parametreleri tahmin edilmektedir.

$$y_i - \hat{\rho} y_{(i-1)} = b_0(1 - \hat{\rho}) + b_1(x_{i1} - \hat{\rho} x_{(i-1),1}) + b_2(x_{i2} - \hat{\rho} x_{(i-1),2}) + \dots + b_{p-1}(x_{i,p-1} - \hat{\rho} x_{(i-1),p-1}) + (e_i - \hat{\rho} e_{i-1}) \quad (2.4)$$

veya

$$y_i^* = b_0^* + b_1^* x_{i1}^* + b_2^* x_{i2}^* + \dots + b_{p-1}^* x_{i,p-1}^* + e_i^*$$

Bu aşamaya kadar uygulanan yöntem *İki Aşamalı Cochrane-Orcutt* yöntemi denilmektedir. Ancak tahmin edilen  $\hat{\rho}$  parametresinin,  $\rho$ 'nun en iyi tahminleyeni olup olmadığı bilinmediğinden, yöntem çok aşamalı olarak uygulanabilmektedir. Çok Aşamalı Cochrane-Orcutt yöntemi adı verilen bu yöntemde genelleştirilmiş farklar yöntemiyle elde edilen yeni hata terimleri arasında da (2.3) denklemleriyle

gösterilen ilişkinin varlığı varsayılmakta ve bu değerler veri olarak kullanılarak  $\hat{\rho}$  için yeni bir tahmini değer  $\hat{\hat{\rho}}$  elde edilmektedir. Bu yeni değer (2.4)'deki genelleştirilmiş farklar denkleminde  $\hat{\rho}$  yerine kullanılmakta ve parametreler tahmin edilmektedir. Aynı aşamalar yinelenerek en iyi tahminleyene erişilinceye kadar devam edilmektedir. En iyi tahminleyene ulaşıldığına karar vermek için,  $\rho$ 'nun art arda tahminleyenleri arasındaki fark hesaplanmakta ve bu değer 0,01 veya 0,001 gibi bir sayıdan küçük olana kadar iterasyona devam edilmektedir.

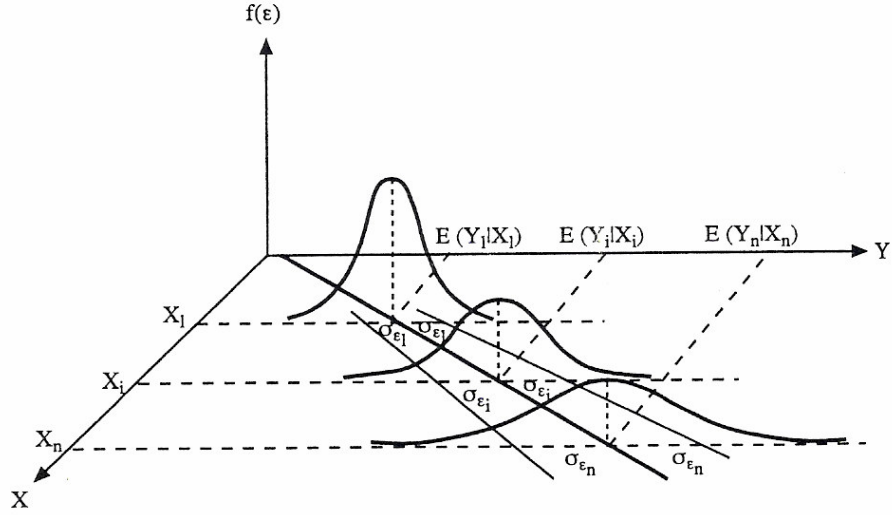
## 2.2. Farklı Varyanslılığın Belirlenmesi ve Çözümleme Yöntemleri

Basit ve çoklu regresyon modellerinin önemli varsayımlardan bir diğeri de, bağımsız değişkenlerin her bir seçilmiş değerine karşı gelen hata teriminin varyansının (aynı zamanda Y'nin varyansının) sabit  $\sigma^2$ 'ye eşit olmasıdır ve bu duruma *Eşit Varyanslılık (Homoscedasticity)* adı verilmektedir. Bazı uygulamalarda, özellikle değişkenlerin aldıkları değerlerin çok yaygın bir aralıkta değişmeleri durumunda, hata teriminin varyansı buna bağlı olarak farklı değerler almakta yani eşit varyanslılık varsayımı geçerli olmamaktadır. Bu durum ise, *Farklı Varyanslılık (Heteroscedasticity)* olarak tanımlanmaktadır.

Farklı varyanslılık durumu, görsel olarak basit regresyon modelinde aşağıdaki şekilde<sup>94</sup> olduğu gibi gösterilebilir. Burada, X değişkeninin değeri büyüdükçe Y'nin koşullu varyanslarının da büyüdüğü gözlenmektedir.

---

<sup>94</sup> Ertek, a.g.e., s. 238.



**Şekil (2.4): Farklı Varyanslılık**

### 2.2.1. Farklı Varyanslılığın Nedenleri

Farklı varyanslılık, seçilen modelin yapısından veya kullanılan verilerden kaynaklanmaktadır ve genellikle zaman serilerinden çok kesit analizinde ortaya çıkmaktadır. Başlıca nedenleri aşağıdaki gibi sıralanmaktadır:<sup>95</sup>

- Verilerin toplanmasında yapılan ölçme hataları
- Dahil edilmesi gerekli değişkenlerin, modele dahil edilmemiş olması
- Katsayılardan bir veya birkaçının zaman serisi verileriyle çalışılıyorsa zamana, kesit verileriyle çalışılıyorsa birimlere göre değişim göstermesi
- Verilerde aşırı -uç- (outliers) değerlerin bulunması

<sup>95</sup> Güriş, Çağlayan, a.g.e., s. 510-512.

## 2.2.2. Farklı Varyanslılığın En Küçük Kareler Yöntemiyle Elde Edilen Tahminleyenlere Etkisi

Regresyon modellerinde eşit varyanslılık varsayımının geçerli olmaması durumunda, En Küçük Kareler yöntemiyle elde edilen tahminleyenlerin hangi istatistik özellikleri taşıdıklarının incelenmesi gerekmektedir. Diğer tüm varsayımların modelde geçerli olmaları durumunda, farklı varyanslılığın regresyon analizine etkileri aşağıdaki gibi özetlenmektedir:

- Parametre tahminleyenleri tutarlı, sapmasız ve asimtotik sapmasız tahminleyenlerdir. Ancak etkin (en küçük varyanslı) ve asimtotik etkin değillerdir. Bu yüzden en iyi doğrusal yansız (BLUE) tahminleyen olma özelliğini taşımazlar.<sup>96</sup>
- Örnek verileriyle elde edilen  $s^2$ ,  $\sigma^2$ 'nin sapmalı tahminleyenidir. Bu durumda, parametre tahminleyenlerinin varyanslarını hesaplamak için kullanılan En Küçük Kareler yönteminin formülleri de sapmalı olmaktadır. Bu formüllerle elde edilen varyanslar, sapma pozitif ise gerçek değerlerinden büyük, sapma negatif ise gerçek değerlerinden küçük olarak tahmin edilmektedirler.
- Parametre tahminleyenlerinin etkin olmayışları, yapılacak aralık tahminlerini, t ve F testlerinin güvenilirliğini etkileyecektir. Varyansların küçük tahmin edilmeleri, otokorelasyonda olduğu gibi, t ve F test istatistiklerinin gerçek değerlerinden daha büyük hesaplanmasına ve dolayısıyla kararların yanlış alınmasına neden olacaktır. Aynı zamanda tahminleyenler için hesaplanan tahmin aralıkları da daralacaktır. Varyanslar gerçek değerlerinden büyük tahmin edildiklerinde ise yapılan açıklamaların tersi durumlar geçerli olacaktır.

---

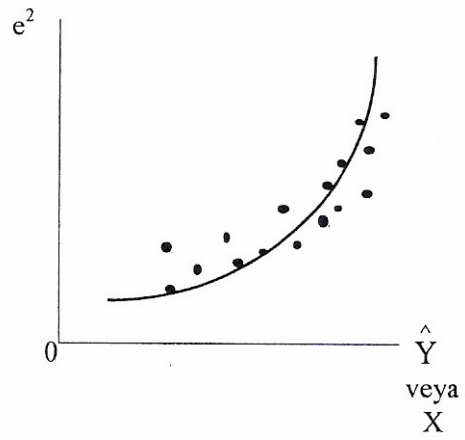
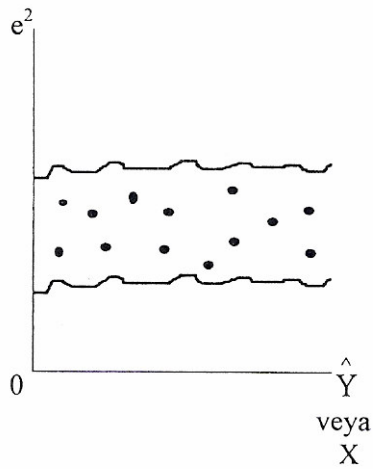
<sup>96</sup> İspat için bkz.: William H. Greene, **Econometric Analysis**, New York, Macmillan Publishing Company, 2000, s. 501-502.

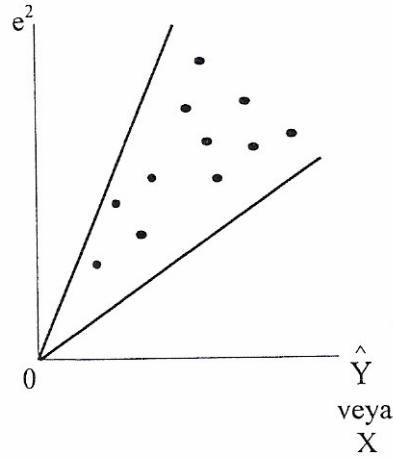
### 2.2.3. Farklı Varyanslılığın Belirlenmesi

Bu bölümde, kurulan regresyon modelinde farklı varyanslılık varsayımının geçerliliğini araştırmak için kullanılan testler ele alınacaktır.

#### 2.2.3.1. Grafik Yöntem

Kurallı bir yöntem olmadığı için, kesin sonuçlar vermemekte, ancak uygulamalarda hata terimleri ile bağımlı veya bağımsız değişkenler arasındaki ilişkiyi görsel olarak ortaya koyduğu için dikkate alınmaktadır. Bu yöntemde, eşit varyanslılık varsayımı altında, örnek verileriyle oluşturulan modelin parametreleri tahmin edilmekte ve tahmini hata terimleri  $e_i$ 'ler hesaplanmaktadır. Düşey eksene  $e^2$  veya standardize hatalar ( $e/s$ ), yatay eksene basit doğrusal regresyon analizinde  $y'$  veya  $x$  değerleri yerleştirilerek hatanın düzenli bir yapı sergileyip sergilemediğine bakılır. Çoklu regresyon analizinde ise  $e^2$  veya  $e/s$  değerlerinin grafiği herbir  $x$  değeri için ayrı ayrı çizilir. Eşit varyanslılık varsayımının geçerli olabilmesi için, grafiklerde hata terimiyle  $y'$  veya  $x$ 'ler arasında belirli bir ilişki gözlenmemesi gerekmektedir.





**Şekil (2.5):  $e^2$  ile  $x$  Arasındaki İlişkiyi Gösteren Grafikler**

Yukarıdaki şekilde verilen grafiklerden birincisinde,  $e^2$  ile  $x$  arasında düzenli bir örüntü gözlenememekte, ikincisinde ikinci dereceden bir ilişki gözlemlenmekte, üçüncüsünde ise  $x$ 'ler arttıkça hata terimlerinin varyansının da arttığı saptanmaktadır. Ancak, bu sonuçların hiçbiri belirli bir test istatistiğine dayanmadığı için kesin değildir.

### 2.2.3.2. White Testi

Yaygın olarak kullanılan bu test, farklı varyanslılığın yapısı hakkında herhangi bir bilgiye gereksinim duyulmadan kullanılmaktadır. Bu yüzden White'in Genel Testi<sup>97</sup> adı da verilmektedir. Farklı varyanslılığı test etmek için hipotezler

$$H_0: \sigma_i^2 = \sigma^2$$

$$H_1: \sigma_i^2 \neq \sigma^2$$

olarak oluşturulurlar.

Bilindiği gibi tahminleyenlerin varyans matrisi,

<sup>97</sup> Hal White, "A Heteroscedasticity-Consistency Covariance Matrix estimator and a Direct Test of Heteroscedasticity", *Econometrica*, Vol.XLVIII, 1980, s. 817-838.

$$V(b) = s^2 (X'X)^{-1}$$

eşitliğiyle elde edilmektedir. Bu yöntemde, eşit varyanslılık varsayımının geçerliliği durumunda bu tahminleyenin tutarlı olduğu, aksi takdirde tutarlı olmadığı gerçeği dikkate alınmıştır. Tutarlılığı test etmek için  $e^2$ 'nin bağımlı, modelin tüm bağımsız değişkenleri, bu değişkenlerin kareleri ve ikili çarpımlarının bağımsız değişken oldukları yardımcı regresyon denklemleri ele alınmaktadır.<sup>98</sup>

Genel olarak, çoklu regresyon modeli örnek verileriyle ele alındığında,

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots + b_{p-1} x_{i,p-1} + e_i \quad (2.5)$$

White testi aşağıdaki adımlarla uygulanmaktadır:

1. En Küçük Kareler yöntemiyle (2.5) modelinin hata terimleri elde edilir.
2. Bu değerlerle yardımcı regresyon modeli,

$$e_i^2 = \alpha_1 + \alpha_2 x_{i1} + \dots + \alpha_{p-1} x_{i,p-1} + \gamma_1 x_{i1}^2 + \dots + \gamma_{p-1} x_{i,p-1}^2 + \theta_1 x_{i1} x_{i2} + \dots + \theta_{p-1} x_{i1} x_{i,p-1} + \delta_1 x_{i2} x_{i3} + \dots + v_i$$

eşitliğiyle kurulur. Bu modelin belirlilik katsayısı, En Küçük Kareler ile tahmin edilir. Eşit varyanslılık varsayımı altında,  $nr^2$  değerinin,  $k$  (yardımcı regresyon modelindeki toplam bağımsız değişken sayısı) serbestlik dereceli ki-kare dağılımına uyduğu gösterilebilmektedir.

$$nr^2 \approx \chi_k^2$$

3. Seçilmiş bir  $\alpha$  anlamlılık seviyesinde ve  $k$  serbestlik derecesiyle ki-kare tablosundan bulunan değer ile  $nr^2$  değeri karşılaştırılmakta ve karar

---

<sup>98</sup> Greene, a.g.e., s. 392-393.



$nr^2 \leq \chi_{\alpha,k}^2$  ise  $H_0$  kabul, eşit varyanslılık varsayımı geçerlidir,

$nr^2 > \chi_{\alpha,k}^2$  ise  $H_0$  red, eşit varyanslılık varsayımı geçerli değildir

şeklinde alınmaktadır.

Yaygın olarak kullanılmasına karşın, bu testin literatürde sözü edilen iki zayıf yönü bulunmaktadır. Birincisi, test istatistiğinin anlamlı çıktığı bazı durumlarda, neden, farklı varyanslılık değil, model kurma hataları olabilmektedir. Thursby<sup>99</sup> çalışmasında, hata terimlerinin sabit varyansa sahip oldukları bir basit regresyon modelinde, White testi uygulandığında temel hipotezin reddedildiğini göstermiştir. Nedeni ise modelde  $x^2$  teriminin eksik alınmış olmasıdır. Bir diğer zayıf yönü ise farklı varyanslılığın hangi bağımsız değişkenler için geçerli olduğu hakkında bir bilgi vermemesi, yani daha sonraki çalışmalara ışık tutmamasıdır.

Örnek sayısı küçük olduğunda, bu testteki test istatistiğinin ki-kare dağılımı gösterdiği varsayımı geçerli olamayabilmekte ve bu durumda White testinin sonuçları tatmin edici olmaktan çıkmaktadır. Yapılan son araştırmalardan birinde, Jeong ve Lee<sup>100</sup> White testinin performansını küçük örnekler için incelemek üzere yeniden örnekleme yöntemlerinden biri olan Bootstrap<sup>101</sup> yöntemini kullanmışlardır. Örnek verileri Monte Carlo simülasyonu kullanılarak oluşturulmuş ve Bootstrap yönteminin, özellikle örnek büyüklüğü küçük olduğunda, White testinin performansını ve güvenilirliğini arttırdığı gösterilmiştir.

### 2.2.3.3. Goldfeld-Quandt Testi

---

<sup>99</sup> Jerry Thursby, "Misspecification, Heteroscedasticity, and the Chow and Goldfeld-Quandt Tests", **Review of Economics and Statistics**, Vol. LXIV, 1982, s. 314-321.

<sup>100</sup> Jinook Jeong, Kyoungwoo Lee, "Bootstrapped White's Test for Heteroscedasticity in Regression Models", **Economics Letters**, Vol. LXIII, 1999, s. 261-267.

<sup>101</sup> Parametrik olmayan bu yöntemde, model için seçilen örnek verilerinin anakütleyi iyi temsil ettiği varsayımıyla bu örnek küleden seçilen verilerle çeşitli örnek kütleleri oluşturulmakta (resampling) ve anakütle dağılımı incelenmektedir.

White testine oranla daha güçlü bulunan bu testte<sup>102</sup>, değişen varyans  $\sigma_i^2$ 'nin modeldeki bağımsız değişkenlerden biri  $x_k$  ile aynı yönlü bir ilişki içinde olduğu varsayılmaktadır. Bu ilişki,

$$\sigma_i^2 = \sigma^2 x_k^2$$

şeklinde ifade edilmektedir. Bu varsayıma göre  $x_k$ 'nın değerleri büyüdükçe  $\sigma_i^2$  de büyüyecektir ve bu durumda modelde farklı varyans bulunması olasılığı yüksektir.

Hipotezler White testinde olduğu gibi, temel hipotezde eşit varyanslılığın geçerli olduğu, alternatif hipotezde de geçerli olmadığı savlarıyla oluşturulmaktadır. Bu test, aşağıda verilen adımlar izlenerek uygulanmaktadır.

1.  $x_k$  bağımsız değişkenin değerleri, küçükten büyüğe doğru sıralanmaktadır. Bu sıralamaya bağlı olarak,  $y$  ve modeldeki tüm diğer bağımsız değişkenlerin gözlem değerlerinin sırası da değişecektir.
2. Tüm değişkenlerin gözlem değerlerinden, önceden belirlenen  $c$  kadar gözlem tam ortadan çıkarılarak atılmakta, böylece tüm değerleri her yarıda  $(n-c)/2$  adet olacak şekilde iki guruba ayrılmış olmaktadır.
3. İki grup için,  $y$ 'ler ile  $x$  değerleri arasında çoklu regresyon modelleri oluşturulur ve En Küçük Kareler yöntemi uygulanarak hata kareleri toplamı iki model için hesaplanır.  $HKT_1$   $x_k$ 'nin küçük değerlerine ait gurubun,  $HKT_2$  ise büyük değerlerine ait gurubun hata kareleri toplamını ifade ederse,  $\frac{HKT_2}{HKT_1}$  oranının eşit varyans varsayımı altında, pay ve payda için  $(n-c-p)/2$ <sup>103</sup> serbestlik dereceli F dağılımına uyduğu gösterilmektedir. Bu durumda, test istatistiği

---

<sup>102</sup> Stephen Goldfeld, Richard Quandt, "Some Tests for Heteroscedasticity", **Journal of the American Statistical Association**, Vol.LX, 1965, s. 539-547.

<sup>103</sup>  $p$  bağımlı değişken dahil, modelin tüm değişkenlerinin sayısıdır.

$$F = \frac{HKT_2}{HKT_1}$$

formülüyle hesaplanır.

4. Belirlenen  $\alpha$  hata payı ve ilgili serbestlik dereceleri ile F tablosundan bulunacak  $F^*$  değeri ile test istatistiği karşılaştırılır ve karar

$F \leq F^*$  ise  $H_0$  kabul, eşit varyanslılık varsayımı geçerlidir

$F > F^*$  ise  $H_0$  red, eşit varyanslılık varsayımı geçerli değildir

şeklinde alınır.

Goldfeld-Quandt testi basit regresyon modellerine uygulandığında,  $x_k$  değişkeni doğal olarak modelin tek bağımsız değişkeni  $x$  olmakta ve bağımsız değişkenle hata terimleri arasındaki ilişki incelenmektedir. Aynı test çoklu regresyon modellerine uygulanırken, farklı varyanslılığa neden olabilecek değişken tahmin edilemiyorsa, test tüm bağımsız değişkenler için ayrı ayrı uygulanmaktadır. Böylece hem eşit varyanslılık varsayımının geçerli olup olmadığı hem de geçerli değilse neden olan bağımsız değişken veya değişkenler belirlenmiş olacaktır. Bu özellik testin güçlü yönlerinden biridir.

Bu testin zayıf sayılabilecek bir özelliği ise ikinci aşamasında belirlenmesi gereken  $c$  sayısı hakkında kesin bir bilgi bulunmayışıdır. Yapılan araştırmalarda, bu değer çok küçük veya büyük olmaması önerilmektedir. Büyük seçilmesi durumunda, F testindeki serbestlik dereceleri küçülecek bu da testin başarısının (kuvvetinin) azalmasına neden olacaktır. Harvey ve Phillips<sup>104</sup>, araştırmalarında  $c < (n/3)$  olmasını savunmaktadırlar. Goldfeld ve Quandt basit regresyon modeli için

---

<sup>104</sup> Andrew Harvey, Garry Phillips, "A Comparison of the Power of Some Tests for Heteroscedasticity in the General Linear Model", **Journal of Econometrics**, Vol.II, 1974, s. 307-316.

uyguladıkları Monte Carlo denemelerinde<sup>105</sup>, örnek büyüklüğü 30 civarındaysa c'nin 8, 60 civarındaysa 16 alınmasını önermektedirler. Bir başka çalışmada ise, Judge v.d.<sup>106</sup>, n ≅ 30 iken c=4 ve n ≅ 60 iken c=8 alınmasını yeterli bulmaktadırlar.

#### 2.2.3.4. Breusch-Pagan/Godfrey Testi

Goldfeld-Quandt testinin başarısı sadece atılacak gözlem sayısı c'nin seçimine değil, aynı zamanda gözlemlerin sıralanacağı x değişkeninin seçimine de bağlıdır. Breusch-Pagan<sup>107</sup>/Godfrey<sup>108</sup> testinde ise modelde yeralan bağımsız değişkenlerden farklı varyanslılığa neden olduğu düşünülen değişkenler z ile ifade edilmişlerdir. Bu z değişkenleri bağımsız değişkenlerin kendileri, kareleri veya e<sup>x</sup> değerleri de olabilmektedir.

Çoklu regresyon modeli,

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots + b_{p-1} x_{i,p-1} + e_i \quad (2.6)$$

eşitliği ile ele alındığında, farklı varyanslılık durumunda hata varyansının z değişkenlerinin bir fonksiyonu olarak,

$$\sigma_i^2 = f(\alpha_0 + \alpha_1 z_{i1} + \alpha_2 z_{i2} + \dots + \alpha_m z_{im})$$

formülüyle tanımlanmaktadır. Bu ilişkinin doğrusal olduğu varsayımıyla, yardımcı regresyon modeli

$$\sigma_i^2 = \alpha_0 + \alpha_1 z_{i1} + \alpha_2 z_{i2} + \dots + \alpha_m z_{im} \quad (2.7)$$

---

<sup>105</sup> Monte Carlo denemeleri, teorik olarak elde edilmeleri mümkün olmayan tahminleyenelerin küçük örnek özellikleri için yapılan simülasyon denemeleridir.

<sup>106</sup> George Judge, Carter Hill, William Griffiths, Tsoung Lee, Helmut Lutkepöl, **Introduction to the Theory and Practice of Econometrics**, New York, John Wiley & Sons, 1982, s. 422.

<sup>107</sup> Trevor Breusch, Adrian Pagan, "A Simple Test for Heteroscedasticity and Random Coefficient Variation", **Econometrica**, Vol.XLVII, 1979, s. 1287-1294.

<sup>108</sup> Leslie Godfrey, "Testing for Multiplicative Heteroscedasticity", **Journal of Econometrics**, Vol.VIII, 1987, s. 227-236.

şeklinde yazılmaktadır.  $\alpha_0$  katsayısı hariç diğer tüm  $\alpha$  katsayılarının 0'a eşit olması durumunda varyans sabit  $\alpha_0$  değerine eşit olacaktır. Buradan hareketle, farklı varyanslılığı test etmek için hipotezler

$$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_m = 0$$

$$H_1: \alpha_1 \neq \alpha_2 \neq \dots \neq \alpha_m \neq 0$$

olarak oluşturulur.

Testin aşamaları şu şekilde geliştirilmiştir:<sup>109</sup>

1. (2.6) modeline En Küçük Kareler yöntemiyle hata terimleri (matris notasyonu ile  $e$  vektörü) elde edilir.
2. Regresyon denkleminin varyansı  $s^2 = (e'e)/n$  formülüyle<sup>110</sup> hesaplanır.
3. Hata terimlerinin karelerinin  $s^2$ 'ye bölünmesiyle  $g_i$  değerleri hesaplanır.

$$g_i = \frac{e_i^2}{s^2}$$

4.  $g_i$  terimlerinin  $z$ 'lere göre regresyonu

$$g_i = \alpha_0 + \alpha_1 z_{i1} + \alpha_2 z_{i2} + \dots + \alpha_m z_{im} + v_i$$

modeliyle kurulur ve parametreleri tahmin edilir. Bu modelin regresyon ile açıklanan değişkenliği bulunup, şu tanım<sup>111</sup> yapılır.

---

<sup>109</sup> Gujarati, **a.g.e.**, s. 377.

<sup>110</sup> Bu tahminleyen en çok benzerlik yöntemiyle elde edilmiştir.

<sup>111</sup> Bu terim literatürde Lagranj Çarpanı adıyla geçmektedir.

$$\Theta = \frac{1}{2} \sum (g' - \bar{g})^2$$

Matris notasyonu kullanıldığında, aynı tanım

$$\Theta = \frac{1}{2} ([G'Z(Z'Z)^{-1}Z'G] - n)$$

eşitliğine dönüşmektedir. Burada  $Z$ ,  $(n \times (m+1))$  boyutunda  $z$  değişkenlerini içeren matrisi,  $G$  ise  $(n \times 1)$  boyutunda  $g_i$  değerlerini içeren vektörü ifade etmektedirler.

(2.6) modelindeki hata terimleri  $e_i$ 'lerin normal dağıldıkları varsayımı ile, eşit varyans varsayımı geçerli iken ve örnek büyüklüğü sonsuza doğru artarken  $\Theta$ 'nin  $(m-1)$  serbestlik dereceli ki-kare dağılımına uyduğu gösterilebilir.

$$\Theta \approx \chi_{m-1}^2$$

5. Seçilmiş bir  $\alpha$  anlamlılık seviyesinde ve  $(m-1)$  serbestlik derecesiyle ki-kare tablosundan bulunan değer ile  $\Theta$  değeri karşılaştırılmakta ve karar

$$\Theta \leq \chi_{\alpha, m-1}^2 \text{ ise } H_0 \text{ kabul, eşit varyanslılık varsayımı geçerlidir,}$$

$$\Theta > \chi_{\alpha, m-1}^2 \text{ ise } H_0 \text{ red, eşit varyanslılık varsayımı geçerli değildir}$$

şeklinde alınmaktadır.

Yapılan çalışmalarda, Breusch-Pagan/Godfrey testinin normallik varsayımına çok hassas olduğu bulunmuş, bu zayıf noktayı ortadan kaldırmak için,  $\Theta$  teriminin, hata terimlerinin daha güçlü bir tahminleyenine bağlı olarak bulunabileceği

savunulmuştur. Koenker<sup>112</sup>, Koenker ve Basset<sup>113</sup> test istatistiğinin, regresyon denkleminin varyansının daha güçlü bir tahminleyeni ile

$$V = \frac{1}{n} \sum_{i=1}^n \left( e_i^2 - \frac{e'e}{n} \right)^2$$

hesaplanabileceğinin savunmaktadırlar. Bu değer ile test istatistiği,

$$\Theta = \left[ \frac{1}{V} \right] (u - \bar{u}i)' Z(Z'Z)^{-1} Z'(u - \bar{u}i)$$

formülüyle tanımlanmaktadır. Bu formülde, u hata terimlerini içeren  $(e_1^2, e_2^2, \dots, e_n^2)$  vektörünü, i (n×1) boyutunda tüm elemanları 1 olan vektörü ve  $\bar{u} = e'e/n$  değerini ifade etmektedir.

Normallik varsayımı altında, bu test istatistiği, Breusch-Pagan/Godfrey test istatistiği ile aynı asimptotik dağılımı göstermekte, hata terimlerinin normal dağılmadığı durumda ise, daha güçlü bir test istatistiği sağladığı yönündeki iddalar, yapılan çalışmalarda gösterilmektedir.

Waldman<sup>114</sup> ise, Breusch-Pagan/Godfrey testindeki z değişkenlerinin, White testindeki yardımcı regresyon modelindeki bağımsız değişkenlerle aynı seçildiklerinde iki testin de aynı cebirsel sonuçları verdiğini göstermiştir.

### 2.2.3.5. Spearman Sıra Korelasyon Testi

---

<sup>112</sup> Roger Koenker, "A Note on Studentizing a Test for Heteroscedasticity", **Journal of Econometrics**, Vol.XVII, 1981, s. 107-112.

<sup>113</sup> Roger Koenker, Gilbert Basset, "Robust Tests for Heteroscedasticity Based on Regression Quantiles", **Econometrica**, Vol.L, 1982, s. 43-61.

<sup>114</sup> Donald Waldman, "A Note on the Algebraic Equivalence of White's Test and a Variant of the Godfrey/Breusch-Pagan Test for Heteroscedasticity", **Econometric Letters**, Vol.XIII, 1983, s. 197-200.

Parametrik olmayan bir test olan Spearman Sıra Korelasyon testi, eşit varyans varsayımının geçerli olup olmadığını belirlemede de kullanılabilmekte ve aşağıdaki gibi uygulanmaktadır.<sup>115</sup>

Bu testte değişkenler arasındaki ilişki sıra korelasyon katsayısı ile belirlenebilir. Bunun için önce model (2.6) eşitliği ile belirlenir ve En Küçük Kareler yöntemiyle hata terimleri elde edilir. Bu terimlerin mutlak değerleri alınarak bu değerler ile  $x_k$  (farklı varyanslılığa neden olan bağımsız değişken) değerleri, ayrı ayrı sıra numaraları (en küçüğe 1 veya en büyüğe 1) verilerek sıralanırlar. Sıra numarası verilirken aynı sayıda birden fazla değer varsa bu değerlere aynı sıra numarası (verilmesi gerekli sıra numaralarının ortalaması) verilir. İki değişkenin sıra numaraları arasındaki ilişki ile sıra korelasyon katsayısı  $r_s$ ,

$$r_s = 1 - 6 \frac{\sum D^2}{n(n^2 - 1)}$$

formülüyle hesaplanır. Burada D, sıra numaraları arasındaki farktır. Bu testte test istatistiği,

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

formülüyle hesaplanmaktadır.

Örnek verileriyle sıra korelasyon katsayısı  $r_s$ 'in anlamlılığını anakütlede sınamak için hipotezler

$$H_0: \rho_s=0$$

$$H_1: \rho_s \neq 0$$

şeklinde kurulur. Belirlenen  $\alpha$  hata payı ile  $n>30$  ise Z tablosundan  $Z_{\alpha/2}$ ,  $n \leq 30$  ise tablosundan  $(n-2)$  serbestlik dereceli  $t_{\alpha/2, n-2}$  değeri belirlenir ve karar

---

<sup>115</sup> Güriş, Çağlayan, **a.g.e.**, s. 512-516.



$t^* \leq Z_{\alpha/2}$  veya  $t_{\alpha/2, n-2}$  ise  $H_0$  kabul eşit varyans varsayımı geçerli  
 $t^* > Z_{\alpha/2}$  veya  $t_{\alpha/2, n-2}$  ise  $H_0$  red eşit varyans varsayımı geçerli değil

şeklinde alınır.

Çoklu regresyon modellerinde bu test uygulanırken sıra korelasyon katsayısı  $r_s$ ,  $|e_j|$  ile her bir  $x$  değeri arasında hesaplanır. Modelede, eşit varyans varsayımının geçerli olabilmesi için, herbir testte  $H_0$  hipotezinin kabul edilmesi gerekmektedir. Herhangi biri reddedildiğinde, farklı varyanslılığa neden olan değişken veya değişkenler de belirlenebilmektedir.

### 2.2.3.6. Park Testi

Park testinde<sup>116</sup>, hata terimleri ile farklı varyanslılığa neden olan değişken arasındaki ilişkinin,

$$\ln(e_i^2) = \beta + \delta \ln x_i + v_i \quad (2.8)$$

şeklinde olduğu varsayılmaktadır. Burada,  $\beta = \ln s^2$  yani regresyon denkleminin tahmini varyansının doğal logaritmasıdır. Eşit varyans varsayımının geçerliliği, hata kareleri ile  $\ln x_i$  arasındaki ilişkinin anlamlılığına yani  $\delta$  parametresinin anlamlılığına bağlıdır. Bu yüzden hipotezler,

$$H_0: \delta=0$$

$$H_1: \delta \neq 0$$

şeklinde ifade edilirler. Test için (2.8)'deki basit regresyon modeli tahmin edilerek  $\delta$  parametresinin anlamlılığı Z veya t testi ile incelenir ve eşit varyans varsayımının geçerli olup olmadığına karar verilir.

---

<sup>116</sup> R. E. Park, "Estimation with Heteroscedastic Error Terms", *Econometrica*, Vol. XXXIV, 1966, s. 888.

Diğer testlerde olduğu gibi, bu testte de çoklu regresyon modellerinde eşit varyanslılık varsayımını geçerliliği araştırılırken, hangi bağımsız değişkenin bu varsayımdan sapmalara neden olduğu bilinmiyorsa, test ayrı ayrı tüm değişkenler için uygulanmaktadır. Önemli olan sadece model için genel bir test ise, (2.8) denkleminde x yerine bağımlı değişkenin tahmini y' değerleri konarak modelin hata terimlerinin varyanslarının sabit olup olmadıkları araştırılmaktadır.

### 2.2.3.7. Glejser Testi

Bu test<sup>117</sup> de Park testine benzer olarak hata terimleri ile farklı varyanslılığa neden olan değişken arasındaki ilişki, basit regresyon modelleriyle açıklanmaktadır. Bu testte, bağımlı değişken olarak hata terimlerinin mutlak değeri alınmakta, bağımsız değişken için ise değişik dönüşümler uygulanabilmektedir. Aşağıda verilen modellerden herhangi biri seçilerek  $\beta_1$  parametresinin anlamlı olup olmadığı Z, t veya F testleriyle incelenmektedir.

$$|e_i| = \beta_0 + \beta_1 x_i + v_i$$

$$|e_i| = \beta_0 + \beta_1 \sqrt{x_i} + v_i$$

$$|e_i| = \beta_0 + \beta_1 \frac{1}{x_i} + v_i$$

$$|e_i| = \beta_0 + \beta_1 \frac{1}{\sqrt{x_i}} + v_i$$

$$|e_i| = \sqrt{\beta_0 + \beta_1 x_i} + v_i$$

Bu test için de, Park testinin çoklu regresyon modelleri için yapılan açıklamaları geçerlidir.

---

<sup>117</sup> Herbert Glejser, "A New Test for Heteroscedasticity", **Journal of the American Statistical Association**, Vol.LXIV, 1969, s. 316-323.

Pesaran ve Taylor<sup>118</sup> ise önerdikleri testte, hata terimleri ile bağımsız değişken arasındaki ilişkiyi,

$$e_i^2 = \gamma_1 + \gamma_2 \hat{x}_i + \eta_i$$

denklemini ile tanımlamaktadırlar. Bu formüldeki  $\hat{x}$  değişkeni farklı varyanslılığa neden olan bağımsız değişkenin kendi veya bu değişkenin *yardımcı değişken regresyonu*<sup>119</sup> adı verilen yöntemle elde edilen değerini göstermektedir. Diğer testlerde olduğu gibi bu testte de, farklı varyanslılığın geçerliliği parametrelerin anlamlılığı test edilerek araştırılmaktadır.

Yukarıda verilen testlerin hepsinde de bağımsız değişken veya değişkenlerde ölçme hatalarının yapılmadığı varsayılmaktadır. Farklı varyanslılık konusunda yapılan son araştırmalardan bazılarında, bağımsız değişkenlerin değerlerini belirlerken yapılabilecek ölçme hataları da dikkate alınarak bu durumda uygulanan testlerin performansları ve doğrulukları ele alınmıştır. Bu araştırmalardan birinde Wallentin ve Agren<sup>120</sup> basit regresyon modellerinde bağımsız değişkende yapılan ölçüm hatalarının farklı varyanslılık testlerini ne şekilde etkilediği araştırılmıştır. Simülasyon yöntemiyle oluşturulan örnek verileriyle hata terimleri hem basit regresyon yöntemiyle, hem de yardımcı değişken regresyon yöntemiyle elde edilmiştir. Bu değerlerle Glejser, Goldfeld-Quandt, White ve Pesaran-Taylor testleri uygulanmış ve White testi dışında diğerlerinin performanslarında önemli bir farka rastlanmamıştır. White testi ise, ölçme hataları bulunması durumunda, farklı varyanslılığı test etmede etkin bir test olmaktan çıkmaktadır.

#### 2.2.4. Farklı Varyanslılık Durumunu Çözümleyen Yöntemler

Basit veya çoklu regresyon modellerinde, bir önceki bölümde sözü edilen yöntemlerden biri ile farklı varyanslılık durumu saptanmışsa, bu durumun etkilerini

---

<sup>118</sup> Hashem Pesaran, L.W. Taylor, "Diagnostics for IV Regressions", **Oxford Bulletin of Economics and Statistics**, Vol.LXI, 1999, s. 255-281.

<sup>119</sup> "Instrumental Variable Regression" teriminin Türkçe karşılığıdır.

<sup>120</sup> Bo Wallentin, Anders Agren, "Test of Heteroscedasticity in a Regression Model in the Presence of Measurement Errors", **Economic Letters**, Vol.LXXVI, 2002, s. 205-211.

ortadan kaldıran çözüm yöntemlerinden uygun olanın seçilerek modelin yeniden oluşturulması gerekmektedir.

### 2.2.4.1. Genelleştirilmiş En Küçük Kareler Yöntemi

Bu yöntem, En Küçük Kareler yönteminin ilk iki temel varsayımının (otokorelasyon olmaması ve eşit varyanslılık) geçerli olmamaları durumunda uygulanabilecek bir yöntemdir. Bilindiği üzere, basit veya çoklu regresyon modelinde otokorelasyon veya farklı varyanslılık durumları saptanmışsa, En Küçük Kareler tahminleyicileri etkin olma özelliklerini kaybetmektedirler. Genelleştirilmiş En Küçük Kareler (Generalized Least Squares) yöntemi bu varsayımlardan sapma olması durumunda uygulanabilen ve etkin tahminleyenlere sahip bir yöntemdir.

Genelleştirilmiş doğrusal regresyon modeli,

$$Y=X\beta+\varepsilon \quad (2.9)$$

$$E(\varepsilon|X)=0$$

$$E(\varepsilon\varepsilon'|X)=\sigma^2\Omega$$

ifadelerle tanımlanmaktadır. Buradaki  $\Omega$  matrisi bir kesin artı(positive definite) matrisidir. Hata terimlerinin birbirleriyle ilişkisiz fakat farklı varyansa eşit olmaları durumu, yani farklı varyanslılık durumunda,  $\sigma^2\Omega$  matrisi aşağıdaki şekilde gösterilmektedir.

$$\sigma^2\Omega = \begin{bmatrix} \sigma_1^2 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \sigma_2^2 & 0 & \cdot & \cdot & 0 \\ \cdot & 0 & \sigma_3^2 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & \sigma_n^2 \end{bmatrix},$$

Aynı matris otokorelasyon bulunması durumunda ise, sabit varyans varsayımı ile

$$\sigma^2 \Omega = \sigma^2 \begin{bmatrix} 1 & \rho_1 & \cdot & \cdot & \rho_{n-1} \\ \rho_1 & 1 & \cdot & \cdot & \rho_{n-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_{n-1} & \rho_{n-2} & \cdot & \cdot & 1 \end{bmatrix},$$

şeklinde ifade edilmektedir. Klasik regresyon modelinde ise, otokorelasyon bulunmaması ve eşit varyanslılık varsayımları ile bu matris,

$$\sigma^2 \Omega = \sigma^2 I$$

ile yani sabit  $\sigma^2$  ile birim matrisinin çarpımı ile gösterilmektedir.

Genelleştirilmiş En Küçük Kareler (GEKK) yönteminde,  $\beta$ 'nin etkin tahminleyeni  $b$ 'nin hesaplanması  $\Omega$  matrisinin bilinmesine bağlıdır.  $\Omega$  bilinen bir kesin artı matrisi ise,

$$\Omega^{-1} = P'P$$

ile belirlenen  $P$  matrisinin (2.9) eşitliğine soldan çarpımı ile,

$$PY = PX\beta + P\varepsilon$$

elde edilmekte ve  $Y_* = PY$ ,  $X_* = PX$ ,  $\varepsilon_* = P\varepsilon$  tanımları kullanıldığında aynı eşitlik

$$Y_* = X_* \beta + \varepsilon_*$$

şekline dönüşmektedir. Bu dönüştürülmüş değişkenler ile elde edilen modelde, minimize edilecek hata kareleri toplamı matris notasyonu ile

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

şeklinde gösterilir. Bu terimlerin minimize edilmesiyle, hata terimlerinin varyansı,

$$E(\varepsilon_* \varepsilon_*') = P\sigma^2\Omega P' = \sigma^2 I$$

ve parametre tahminleyen vektörü de,

$$\begin{aligned} \hat{b} &= (X_*' X_*)^{-1} X_*' Y_* \\ &= (X' P' P X)^{-1} X' P' P Y \\ &= (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} Y \end{aligned}$$

formülleriyle elde edilmektedirler. GEKK yönteminin tahminleyeni şu özellikleri taşımaktadır:<sup>121</sup>

- Eğer  $E(\varepsilon_* | X) = 0$  ise

$E(\hat{b}) = E((X_*' X_*)^{-1} X_*' Y_*) = \beta + E((X_*' X_*)^{-1} X_*' \varepsilon_*) = \beta$ 'dir. Yani  $\beta$ 'nin yansız tahminleyendir.

- $\hat{b}$ , tutarlı bir tahminleyendir.

- $\hat{b}$ , ortalaması  $\beta$  ve varyansı

$$V(\hat{b}) = \sigma^2 (X_*' X_*)^{-1} = \sigma^2 (X' \Omega^{-1} X)^{-1}$$

ile asimptotik normal bir dağılıma uymaktadır.

- $\hat{b}$ , GEKK yönteminin en küçük varyanslı tahminleyendir. Başka bir ifade ile etkin bir tahminleyendir.

---

<sup>121</sup> Greene, **a.g.e.**, s. 311-312.

## 2.2.4.2. Ağırlıklı En Küçük Kareler Yöntemi

Ağırlıklı En Küçük Kareler<sup>122</sup> (AEKK) yöntemi (Weighted Least Squares Method), GEKK yönteminin özel bir hali olup<sup>123</sup>, farklı varyanslılık durumunda çözümlene yöntemlerinden biri olarak uygulanabilmektedir. Farklı varyanslılık durumunda, hata terimlerinin sabit olmayan varyansları,

$$V(\varepsilon_i) = \sigma_i^2 = \sigma^2 \omega_i$$

formülüyle gösterilirse,  $\Omega$  matrisi

$$\Omega = \begin{bmatrix} \omega_1 & 0 & . & . & . & 0 \\ 0 & \omega_2 & 0 & . & . & 0 \\ . & 0 & \omega_3 & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & \omega_n \end{bmatrix},$$

şeklinde diagonal bir matris olmakta ve tersi  $\Omega^{-1}$ 'de köşegen değerleri  $1/\omega_i$ 'ye eşit olan diagonal bir matrisi göstermektedir. Buradaki  $1/\omega_i$  terimleri hata terimlerine verilen *ağırlık*'ları ifade etmektedirler. Bu durumda, PY ve PX matrisleri,

$$PY = \begin{bmatrix} y_1 / \sqrt{\omega_1} \\ y_2 / \sqrt{\omega_2} \\ . \\ . \\ y_n / \sqrt{\omega_n} \end{bmatrix}, \quad PX = \begin{bmatrix} x_{11} / \sqrt{\omega_1} & . & . & . & x_{(p-1)1} / \sqrt{\omega_1} \\ x_{12} / \sqrt{\omega_2} & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{1n} / \sqrt{\omega_n} & . & . & . & x_{(p-1)n} / \sqrt{\omega_n} \end{bmatrix}$$

<sup>122</sup> Literatürde Tartılı En Küçük Kareler yöntemi adı da verilmektedir.

<sup>123</sup> Farklı varyanslılık durumunda iki yöntem birbirine eşit olmaktadır.

değerlerini almaktadırlar. Ağırlıklı En Küçük Kareler yöntemi hataların ağırlıklı karelerinin toplamını,

$$\sum_{i=1}^n \frac{1}{\omega_i} e_i^2$$

minimum yapacak şekilde parametreleri, daha önce de gösterildiği gibi

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

matrisi ile tahmin etmektedir. Dikkat edilirse, En Küçük Kareler yöntemi, Ağırlıklı En Küçük Kareler yönteminin özel bir hali olup, tüm ağırlıkların eşit seçilip 1'e eşitlenmesi durumunu göstermektedir.

Farklı varyanslılık düzeltme yöntemleri, anakütlede hata terimleri varyansı  $\sigma_i^2$ 'nin bilinip bilinmemesi durumuna göre değişmektedir.

### 2.2.4.3. $\sigma_i^2$ 'nin Bilinmesi Durumu

Hata terimlerinin varyansı biliniyorsa, genellikle AEKK yöntemi, varyansla ters orantılı ağırlık verilerek uygulanmaktadır. Carrol ve Ruppert'e<sup>124</sup> göre bağımlı değişkenin veya hata karelerinin standart hatalarının en büyük değeri ile en küçüğünün oranının 1,5'i aşmaması durumunda klasik EKK yönteminin uygulanıp farklı varyanslılığın test edilmesi gerektiği, bu oranın 1,5'dan büyük olması durumunda ancak AEKK yönteminin kullanılmasının, modelin tahmininde dikkate değer bir iyileştirme getireceği vurgulanmaktadır. Böylece, hem modelin hem de parametre tahminleyenlerinin standart hataları, En Küçük Kareler yöntemiyle

---

<sup>124</sup> Raymond J. Carrol, David Ruppert, **Transformation and Weighting in Regression**, New York, Chapman and Hall, 1988, s. 16.



bulunanlardan daha düşük olarak elde edilmektedirler. Bu sonuç, özellikle büyük örnek verileri ile çalışılması durumunda her zaman geçerlidir.<sup>125</sup>

$\sigma_i^2$ 'nin bilinmesi durumunda, ağırlıkların  $1/w_i=1/\sigma_i^2$  olarak seçilerek, AEKK yöntemi uygulanmaktadır. Böylece büyük varyanslı gözlemlerin ağırlıkları göreceli olarak küçük, küçük varyanslı gözlemlerin ağırlıkları göreceli olarak büyük alınmaktadır. Daha önce de gösterildiği gibi bu yöntemle elde edilen modelin hata terimlerinin varyansı sabit ve parametre tahminleyenleri ise etkin olacaktır. Matris notasyonu kullanıldığında, modeldeki  $\sigma^2\Omega$  matrisi ile PY ve PX vektörleri aşağıdaki değerleri alacaklardır.

$$\sigma^2\Omega = \begin{bmatrix} \sigma_1^2 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \sigma_2^2 & 0 & \cdot & \cdot & 0 \\ \cdot & 0 & \sigma_3^2 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & \sigma_n^2 \end{bmatrix},$$

$$PY = \begin{bmatrix} y_1/\sigma_1 \\ y_2/\sigma_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n/\sigma_n \end{bmatrix}, \quad PX = \begin{bmatrix} x_{11}/\sigma_1 & \cdot & \cdot & \cdot & x_{(p-1)1}/\sigma_1 \\ x_{12}/\sigma_2 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{1n}/\sigma_n & \cdot & \cdot & \cdot & x_{(p-1)n}/\sigma_n \end{bmatrix}$$

Görüldüğü gibi, modeldeki tüm değişkenlerin  $\sigma_i$ 'ye bölünmesiyle değişken dönüşümü uygulanmakta, ve aynı dönüşümün hata terimlerine uygulanması ile hata teriminin varyansı

<sup>125</sup> Raymond J. Carrol, David Ruppert, "Robust Estimation in Heteroscedastic Linear Models", *Annals of Statistics*, Vol.X, 1982, s. 429-441.

$$V\left(\frac{\varepsilon_i}{\sigma_i}\right) = E\left[\frac{\varepsilon_i}{\sigma_i} - E\left[\frac{\varepsilon_i}{\sigma_i}\right]\right]^2 = E\left[\frac{\varepsilon_i}{\sigma_i}\right]^2 = \frac{1}{\sigma_i^2} E(\varepsilon_i^2) = \frac{\sigma_i^2}{\sigma_i^2} = 1$$

olarak hesaplanmaktadır. Dolayısıyla,  $V\left(\frac{\varepsilon_i}{\sigma_i}\right)$  sabit ve homoscedastiktir.

Tahminleyen vektörü,  $\Omega$  matrisi ile AEKK yöntemiyle

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

formülünden, veya değişken dönüşümleri kullanılarak En Küçük Kareler yöntemiyle elde edilebilmektedir. Her iki yöntem de aynı etkin ve yansız tahminleyenleri vereceklerdir. Tahminleyenlerin varyansı ise,

$$V(\hat{b}) = \sigma^2 (X' \Omega^{-1} X)^{-1}$$

formülünden hesaplanmaktadır.

#### 2.2.4.4. $\sigma_i^2$ 'nin Bilinmemesi Durumu

Hata terimlerinin varyansının bilinmesi durumuna genellikle ender olarak rastlanmaktadır. Genel olarak, bilinmeyen  $\sigma_i^2$  tahmini için bazı varsayımlar yapılmakta ve regresyon modeli farklı varyanslılığı olmayan başka bir modele dönüştürülmektedir.

**Varsayım 1:** Bazı analizlerde, bağımsız değişkenlerin büyük değerleri büyük hata terimlerine neden olmakta ve farklı varyanslılık ortaya çıkmaktadır. Bu durumda daha önceki bölümde sözü edilen testler uyguladığında, modelde farklı varyanslılık durumu saptanmakta, hatta bazı testler hangi değişkenin farklı varyanslılığa neden olduğunu da ortaya çıkarmaktadır.  $\sigma_i^2$ 'nin bu bağımsız değişkenle ( $X_k$ )

ilişkilendirilmesi ile hata terimlerinin eşit varyanslı oldukları bir model tahmin edilebilmektedir. Genel olarak bu ilişki,<sup>126</sup>

$$\sigma_i^2 = \sigma^2 x_{ik}^\gamma$$

eşitliği ile tanımlanmaktadır.<sup>127</sup>  $\gamma=0$  ise sabit varyans varsayımı geçerlidir. Uygulamalarda genellikle  $\gamma=2$  veya  $\gamma=1$  alınmaktadır.

1)  $\gamma=2$  durumunda,

$$\sigma_i^2 = \sigma^2 x_{ik}^2$$

olarak ifade edilebilir. Ağırlıklar,  $\frac{1}{\omega_i} = \frac{1}{x_{ik}^2}$  ile belirlenip AEKK yöntemi uygulanarak elde edilen regresyon denklemi farklı varyanslıktan arındırılmış olacaktır.

En Küçük Kareler yönteminin kullanılması durumunda ise değişkenlere dönüşüm uygulanması gerekmektedir. Daha önce belirtildiği gibi, değişkenlerin  $\sqrt{\omega_i}$  değerlerine bölünmesi ile bu dönüşümler uygulanmaktadır. Bu durumda çoklu regresyon denklemi

$$\frac{y_i}{x_{ik}} = b_0 \frac{1}{x_{ik}} + b_1 \frac{x_{i1}}{x_{ik}} + \dots + b_{p-1} \frac{x_{i,p-1}}{x_{ik}} + \frac{e_i}{x_{ik}}$$

eşitliği ile ele alınmakta ve bu modelin hata teriminin varyansı,

$$V\left(\frac{e_i}{x_{ik}}\right) = E\left[\left(\frac{e_i}{x_{ik}}\right)^2\right] = \frac{1}{x_{ik}^2} E(e_i^2) = \frac{1}{x_{ik}^2} \cdot \sigma^2 x_{ik}^2 = \sigma^2$$

<sup>126</sup> Güriş, Çağlayan, **a.g.e.**, s. 537.

<sup>127</sup> Formüldeki  $x_i$  genel olarak bağımsız değişkenlerden biri olarak alındığı halde, bazen de modelde yeralmayan bir bağımsız değişken de olabilmektedir.

sabit değerine eşit olmaktadır. Dolayısıyla, dönüştürülmüş değişkenler ile elde edilen bu modelde eşit varyanslılık durumu geçerli olacak ve etkin tahminleyenler elde edilecektir.  $x_{ik}$  bağımsız değişkenlerden biri olduğundan, modelde bu değişkene ait katsayı modelin sabit katsayısını oluşturacaktır.

- 2)  $\gamma=1$  durumunda varyans, farklı varyanslılığa neden olan bağımsız değişken ile ilişkilendirilmektedir.<sup>128</sup>

$$\sigma_i^2 = \sigma^2 x_{ik}$$

Ağırlıklar,  $1/\sqrt{x_{ik}}$  olarak seçilerek AEKK yöntemi uygulanabilmekte veya tüm değişkenler  $\sqrt{x_{ik}}$  değerine bölünerek oluşturulan model En Küçük Kareler yöntemi ile tahmin edilmektedir. Bir önceki durumda olduğu gibi, burada da hata terimleri varyansı sabit ve  $\sigma^2$  değerine eşittir.

İstatistik programlarında AEKK yöntemi uygulanırken yaygın olarak kullanılan yaklaşım,  $\gamma$ 'yı belirli bir başlangıç değerinden başlayarak belirli bir son değere ulaşınca kadar her adımda sabit değerlerle arttırarak yöntemi uygulamak ve elde edilen modeller arasında en iyisine karar verebilmek için En Büyük Olabilirlik fonksiyonun hesaplamaktır. Bilindiği gibi, bu fonksiyonu enbüyükleyen ağırlık değeri ile elde edilen model en iyi model olmaktadır. AEKK modeli için En Büyük Olabilirlik fonksiyonu

$$L = \frac{-n}{2} (\ln(2\pi) + \ln(\sigma^2)) + \frac{1}{2} \sum_{i=1}^n \ln w_i - \frac{1}{2} (n - p + 1) \quad (2.10)$$

formülüyle hesaplanmaktadır.

---

<sup>128</sup> A.e., s. 538.

**Varsayım 2:** Bazı analizlerde de, hata terimlerinin varyansının bağımlı değişkenin beklenen değerinin karesiyle doğru orantılı olarak değiştiği varsayılmaktadır. Bu durumda, ilişki

$$\sigma_i^2 = \sigma^2 [E(Y_i)]^2$$

şeklinde tanımlanmaktadır. Değişkenlerin ve hata teriminin  $E(Y_i)$ 'ye bölünmesiyle elde edilen denklemde hata terimleri varyansı sabit  $\sigma^2$  değerine eşittir. Ancak  $E(Y_i)$ 'nin gerçek değeri bilinmediği için, onun tutarlı bir tahminleyeni olan  $y_i$ '<sup>129</sup> En Küçük Kareler yöntemiyle elde edilmekte ve bu değerlerin karesi ağırlık olarak alınarak Ağırlıklı EKK yöntemi uygulanmaktadır.

**Varsayım 3:** Değişkenlere logaritmik dönüşümler uygulandığında farklı varyanslılık sorunu çözülebilmektedir. Bunun nedeni, logaritmanın, değişkenlerin ölçüldüğü ölçeği daraltarak, değişkenler arasındaki farkları azaltmasıdır. Değişkenlerin tümü logaritmik olarak modele girdiğinde parametreler,

$$\ln Y_i = \beta_0 + \beta_1 \ln X_{i1} + \beta_2 \ln X_{i2} + \dots + \beta_{p-1} \ln X_{i,p-1} + \varepsilon_i$$

modeliyle tahmin edilmektedir. Bazı durumlarda ise sadece bağımlı değişkene uygulanan logaritmik dönüşüm ile farklı varyanslılık ortadan kaldırılmaktadır.<sup>130</sup>

**Varsayım 4:** Bağımlı değişkenin beklenen değeri ile varyans arasında belirli bir ilişki gözlemlenebiliyorsa, bağımlı değişkene uygulanan dönüşümler ile farklı varyanslılık durumu ortadan kaldırılmaktadır. Bu ilişkinin şekline göre dönüşümler aşağıdaki tabloda özetlendiği gibi uygulanabilmektedir.<sup>131</sup>

### **Tablo (2.2): $\sigma^2$ ile $E(y)$ Arasındaki İlişkiye Göre Önerilen Dönüşümler**

<sup>129</sup> Örnek büyüklüğü sonsuza doğru arttıkça, gerçek  $E(Y_i)$  değerine yakınsarlar.

<sup>130</sup> Chatterjee, Price, **a.g.e.**, s. 54-57.

<sup>131</sup> David G. Kleinbaum, Lawrence L. Kupper, Keith E. Muller, Azhar Nizam, **Applied Regression Analysis and Other Multivariable Methods**, 3rd edition, Washington, Duxbury Press, 1998, s. 252.

$\sigma^2$ ile E(y) arasındaki ilişki	Önerilen Dönüşüm
$\sigma^2 \cong E(y)$	$y' = \sqrt{y}$
$\sigma^2 \cong E(y)(1-E(y))$	$y' = \sin^{-1} \sqrt{y}$
$\sigma^2 \cong (E(y))^3$	$y' = y^{-1/2}$
$\sigma^2 \cong (E(y))^4$	$y' = y^{-1}$

**Varsayım 5:** Farklı varyanslılığa neden olan değişken tekrarlı değerlere sahipse veya birbirine yakın değerlerin gruplanmasından oluşuyorsa, literatürde bilinmeyen ağırlıkların tahmini için önerilen bazı yöntemler bulunmaktadır.

1. Bağımsız değişken değerleri tekrarlı sayılardan oluşuyorsa, bilinmeyen  $\sigma_i^2$  yerine örnek verileriyle her farklı  $x_i$  değerine karşı gelen  $y$  değerlerinin varyansı  $s_y^2$  hesaplanabilir ve  $\sigma_i^2$ 'nin tahmini değeri olarak kullanılabilir. Fakat örnek sayısı küçük olduğunda,  $x_i$ 'nin tekrarlı sayıları da az olacağından  $s_y^2$  iyi bir tahminleyen olmaktan uzaktır. Gerçekçi bir tahminleyen için Carroll ve Cline'a<sup>132</sup> göre  $x_i$ 'nin tekrarlanan değerleri en az 10 adet, Deaton, Reynolds ve Myers'e<sup>133</sup> göre ise en az 9 adet olmalıdır. Fakat gerçek uygulamalarda, verilerde birbirine eşit 9 veya 10 adet bağımsız gözlem değeri bulunmasına nadir olarak rastlanmaktadır. Bu yüzden  $s_y^2$  değeri yerine hata kareleri ortalaması,

$$\sum (y - y')^2 / m_i$$

<sup>132</sup> Raymond J. Carroll, Daren B. H. Cline, "An Asymptotic Theory for Weighted Least Squares with Weights Estimated by Replication", **Biometrika**, Vol.LXXV, 1988, s. 35-43.

<sup>133</sup> M.L. Deaton, Marion R. Reynolds, Raymond H. Myers, "Estimation and Hypothesis Testing in Regression in the Presence of Nonhomogeneous Error Variances", **Communications in Statistics, Simulation and Computation**, Vol.XII, 1983, s. 45-66.

$\sigma_i^2$ 'nin tahmini değeri olarak kullanılabilir. <sup>134</sup> Buradaki  $m_i$ ,  $x_i$ 'nin tekrar sayısıdır. Carroll ve Cline bu sayının tüm  $x_i$ 'ler için aynı olduğu durumu incelerken, Davidian <sup>135</sup> ise farklı olma durumunu ele almıştır.

2. Draper ve Smith'in önerdikleri yöntemde <sup>136</sup>, farklı varyanslılığa neden olan bağımsız değişken ( $x_k$ ) değerleri, birbirlerine eşit veya yakın değerlerden oluşan gurupların birleşimiyle gösterilebiliyorsa, öncelikle veriler bu değişkenin değerlerinin küçükten büyüğe doğru sıralanması ile yeniden düzenlenir. Verilerde t adet grup olduğu varsayımıyla aşağıdaki adımların uygulanması ile tahmini ağırlıklar hesaplanır.

1) Her gruba ait ortalamalar ( $\bar{x}_{kj}$ ) ve her gruba ait y değerlerinden hata terimlerinin varyansının tahmini değeri  $s_{yj}^2$  hesaplanır.

2) Hesaplanan bu değerler veri <sup>137</sup> olarak kullanılarak, varyans ile bağımsız değişkenin ortalaması arasındaki ilişki,

$$s_y^2 = \beta_0 + \beta_1 \bar{x}_k + \beta_1 \bar{x}_k^2 \quad (2.11)$$

regresyon denklemi ile ifade edilip, En Küçük Kareler yöntemiyle bilinmeyen parametrelerin tahmini değerleri elde edilir.

3) Bağımsız değişken  $x_k$ 'nın her gözlem değeri  $x_{ki}$  (2.11) denkleminde  $\bar{x}_k$  yerine yerleştirilerek, gözlem sayısı kadar  $s_{yi}^2$  değeri hesaplanarak her bir gözlem değeri için tahmini  $\sigma_i^2$  değeri hesaplanır.

---

<sup>134</sup> Carroll, Cline, a. y.

<sup>135</sup> Marie Davidian, "Estimation of Variance Components in Assays with Possibly Unequal Replication and Nonnormal Data", **Biometrika**, Vol.LXXVII, 1990, s. 43-54.

<sup>136</sup> Draper, Smith, a.g.e., s.112-116.

<sup>137</sup> t adet gurup bulunduğunda, t adet ortalama ve varyans değeri elde edilecektir.

4) Son olarak ağırlıklar  $1/s_{yi}^2$  şeklinde kullanılmaktadır.

Bu tahmini varyanslar, örnek birim sayısının yeteri kadar büyük olması durumunda kullanılarak AEKK yöntemi uygulanabilmekte ve regresyon denkleminin yansız ve etkin tahminleyenleri elde edilmektedir. Ancak örnek verilerinde gözlem sayısı az ise, tahmini ağırlık değerleri gerçekçi olmaktan uzaktırlar. Bu durumda, daha iyi tahminleyenler elde etmek için iterasyon yöntemi kullanılmaktadır. *Yeniden Ağırlıklandırılmış En Küçük Kareler Yöntemi (Iteratively Reweighted Least Squares Method)* adı verilen bu yaklaşımda, ağırlıkların başlangıç değeri daha önce verilen yöntemlerden biriyle tahmin edilerek AEKK yöntemiyle regresyon denklemi oluşturulur. Elde edilen bu modelden  $y'$  değerleri ve yeni hata terimleri elde edilir. Bu değerlerden hareketle yeni tahmini ağırlıklar hesaplanır. Bu yeni değerlerle yeniden AEKK yöntemi uygulanır ve iterasyon, regresyon denkleminin parametrelerinin gerçek değerlerine yakınsaması, başka bir ifadeyle elde edilen parametre değerleri arasındaki farkın 0'a çok yakın olması durumunda sonlandırılır.<sup>138</sup>

### 2.3. Normallik Varsayımından Sapmaların Belirlenmesi ve Çözümleme Yöntemleri

Regresyon analizinin temel varsayımlarından bir diğeri de, anakütle hata terimlerinin sıfır ortalama ve sabit varyansla normal dağılım göstermeleridir. Normallik varsayımı sembollerle,

$$E(\varepsilon_i) = 0$$

$$E(\varepsilon_i^2) = \sigma^2$$

$$E(\varepsilon_i \cdot \varepsilon_j) = 0 \quad i \neq j$$

---

<sup>138</sup> Raymond H. Myers, **Classical and Modern Regression with Applications**, second edition, Washington, Duxbury, 1990, s. 351-352.



$$\varepsilon_i \approx N(0, \sigma^2)$$

olarak ifade edilebilir. Bu varsayım çoğu durumlarda gerçekleşmesi beklenen bir varsayımdır. Hata terimleri genellikle modele dahil edilmeyen bağımsız değişkenlerin etkisini temsil etmektedir. Bu yüzden unutulmuş değişkenlerin modele dahil edilmesiyle hata terimleri dağılımı normale yaklaşmaktadır.

### **2.3.1. Normallik Varsayımından Sapmaların En Küçük Kareler Yöntemine Etkileri**

Normal dağılım varsayımı özellikle parametrik testlerin yapılabilmesi için gereklidir. Hata terimleri anakütlede normal dağılım gösterse de göstermese de, En Küçük Kareler yöntemiyle saptanan tahminleyenler en iyi doğrusal yansız (BLUE) tahminleyendirler. Bilindiği üzere normal dağılan değişkenlerin fonksiyonları da normal dağılır. Bu yüzden ancak normallik varsayımı altında, bu tahminleyenlerin de normal dağıldıkları ve dolayısıyla çeşitli test istatistiklerinde t, F ve  $\chi^2$  dağılımlarının kullanılabilmesi saptanmaktadır. Başka bir ifade ile, normallik varsayımı geçerli değil fakat diğer varsayımlar geçerli iseler, parametrelerin nokta tahminleri elde edilir fakat bu parametreler için anlamlılık testleri ve aralık tahminlerinin güvenilirliğinden söz edilemez. Ancak, örnek birim sayısı arttıkça merkezi limit teoremi gereği dağılım, normal dağılıma yaklaşır.<sup>139</sup>

### **2.3.2. Normallik Varsayımından Sapmaların Belirlenmesi**

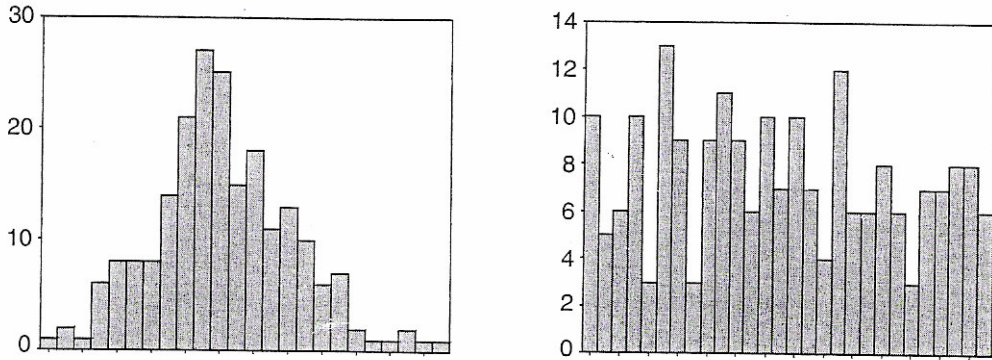
Normallik varsayımının sınanması için literatürde önerilen testler aşağıdaki sırada ele alınabilirler.

---

<sup>139</sup> Güriş, Çağlayan, **a.g.e.**, s. 576-577.

### 2.3.2.1. Grafik Yöntem

En pratik yöntemlerden biri olan grafik yöntemde, hata terimlerinin histogramları çizilerek normal dağılıma uyup uymadıkları saptanabileceği gibi, normal olasılık grafiklerinden de yararlanılabilmektedir.<sup>140</sup> Bilindiği gibi, gruplanmış seriler için çizilen histogramlarda, koordinat sisteminin yatay eksenine ilgili değişkenin sınıf sınırları, dikey eksene ise frekanslar veya nispi frekansların yerleştirilmesiyle oluşan dikdörtgenlerle frekans dağılışı grafiksel olarak incelenmektedir<sup>141</sup>. Hata terimleri için çizilen histogramlarda, oluşan şeklin simetrik ve yaklaşık çan eğrisi biçiminde olması, büyük bir olasılıkla normal veya normale çok yakın sayılabilecek bir dağılıma uygunluğunu göstermektedir. Aşağıdaki grafiklerin ilkinde normale yakın, ikincisinde ise normale uymayan bir dağılım gözlemlenmektedir.



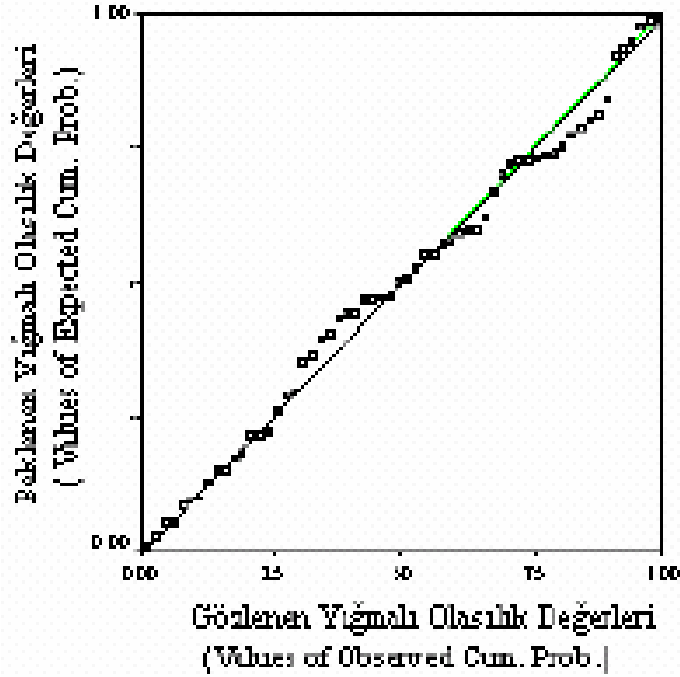
Şekil (2.6): Histogramlar

Normal olasılık grafiklerinde ise hata terimleri beklenen değerlerine karşı çizilmekte, ve normal dağılım varsayımı altında elde edilen grafik, bir doğruyu

<sup>140</sup> Jeremy Miles, Mark Shevlin, **Applying Regression & Correlation**, London, SAGE Publications, 2001, s. 67-73.

<sup>141</sup> David R. Anderson, Dennis J. Sweeney, Thomas A. Williams, **Statistics for Business and Economics**, eighth edition, Ohio, South-Western Thomson Learning, 2002, s. 33.

şekillendirmekte, yani doğrusal bir ilişkiyi temsil etmektedir. Doğrudan olan sapmalar, ilgili gözlemlerin normal dağılımdan sapmalarını ifade etmektedir. Normalden sapan gözlemlerin bulunması uygun olmayan bir regresyon modelinin kullanıldığını veya varyansın homojen olmadığını ifade eder. Aşağıdaki şekilde hata terimlerinin gözlenen değerleri, beklenen değerlerine karşı çizilmiş, ve doğrusal bir ilişki gözlemlenmiştir. Bu yüzden çalışılan örnekte normallik varsayımının bulunduğu saptanmıştır.<sup>142</sup>



Şekil (2.7): Normal Olasılık Grafikleri

### 2.3.2.2. Momentlere Dayanan Asimetri ve Basıklık Ölçülerine göre Yapılan Test

Normal dağılımda standart aralıklarda seri terimleri belirli oranlarda bulunacağından, bu aralıklarda  $(\bar{x} \pm z\sigma)$  yer alan terimlerin oransal payların hesaplanması ve standart aralıklarla karşılaştırılması ile dağılımın normalliği

<sup>142</sup> Suat Şahinler, “En Küçük Kareler Yöntemi ile Doğrusal Regresyon modeli Oluşturmanın Temel Prensipleri”, MKÜ Ziraat Fakültesi Dergisi, C.V, 2000, s. 57-73.

incelenebilirken, momentlere dayanan asimetri ve basıklık ölçülerinden de yararlanılabilir.

Bilindiği gibi, aritmetik ortalamadan farklarla üçüncü dereceden moment, standart sapmanın kübüne bölüldüğünde asimetri ölçüsü, dördüncü dereceden moment standart sapmanın dördüncü kuvvetine bölüldüğünde ise basıklık ölçüsü elde edilmektedir.<sup>143</sup> Hata terimleri için asimetri (S) ve basıklık (K) ölçüleri,

$$S = \frac{\sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})^3}{\sigma^3}, \quad (2.12)$$

ve

$$K = \frac{\sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})^4}{\sigma^4} \quad (2.13)$$

formülleriyle ifade edilmektedirler. Bu ölçülerin standart sapma değerleri ise, sırasıyla

$$s(S) = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}}$$

ve

$$s(K) = \sqrt{\frac{24n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)}}$$

---

<sup>143</sup> Neyran Orhunbilge, **Tanımsal İstatistik Olasılık ve Olasılık Dağılımları**, İstanbul, Avcıol Basım Yayın, 2000, s. 136-137.

formülleriyle hesaplanmaktadır.<sup>144</sup>

Normallik varsayımı geçerli ise hata terimlerine ait asimetri ölçüsü 0 veya 0'a çok yakın, basıklık ölçüsü ise 3 veya 3'e çok yakın olmalıdır. Bu değerlerden sapmalar arttıkça dağılımın normal dağılıma uymadığı sonucu çıkarılmaktadır. Bu sonuçların anakütlede anlamlılığını test etmek için hipotezler asimetri ölçüsü için:

$$H_0: S=0$$

$$H_1: S \neq 0$$

ve basıklık ölçüsü için:

$$H_0: K=3$$

$$H_1: K \neq 3$$

şeklinde kurulurlar. Hipotezler, t veya Z testleriyle sınanmaktadır. Bu amaçla, test istatistiği sırasıyla,

$$t^*(\hat{S}) = \frac{\hat{S}}{s(\hat{S})}$$

ve

$$t^*(\hat{K}) = \frac{\hat{K} - 3}{s(\hat{K})}$$

formülleriyle hesaplanmaktadır.

---

<sup>144</sup> Büyük n değerleri için  $s(S) = \sqrt{6/n}$  ve  $s(K) = \sqrt{24/n}$  değerlerine yaklaşık olarak eşit olmaktadır.

Belirlenen  $\alpha$  hata payı ile  $n > 30$  ise Z tablosundan  $Z_{\alpha/2}$ ,  $n \leq 30$  ise tablosundan  $(n-1)$  serbestlik dereceli  $t_{\alpha/2, n-1}$  değeri belirlenir ve karar asimetri ölçüsü için:

$|t^*(\hat{S})| \leq Z_{\alpha/2}$  veya  $t_{\alpha/2, n-1}$  ise  $H_0$  kabul asimetri ölçüsü anakütlede anlamlı

$|t^*(\hat{S})| > Z_{\alpha/2}$  veya  $t_{\alpha/2, n-1}$  ise  $H_0$  red eşit asimetri ölçüsü anakütlede anlamlı değil

şeklinde alınır. Benzer şekilde basıklık ölçüsü için de kararlar aşağıdaki şekilde alınır:

$|t^*(\hat{K})| \leq Z_{\alpha/2}$  veya  $t_{\alpha/2, n-1}$  ise  $H_0$  kabul basıklık ölçüsü anakütlede anlamlı,

$|t^*(\hat{K})| > Z_{\alpha/2}$  veya  $t_{\alpha/2, n-1}$  ise  $H_0$  red eşit basıklık ölçüsü anakütlede anlamlı değil.

### 2.3.2.3. Ki-Kare Uygunluk Testi

Hata terimlerinin gözlenen değerleriyle, normal dağılıma uygun olması durumunda alması gereken beklenen değerleri arasındaki farkın anlamlı olup olmadığını sınavan bu test, aşağıdaki adımlar izlenerek uygulanmaktadır.<sup>145</sup>

1. Basit veya çoklu regresyon modeli En Küçük Kareler yöntemiyle tahmin edilerek hata terimleri bulunur ve bu terimlerin standart sapması hesaplanır. Hata terimleri ortalama değerlerinden sapmalara göre gruplandırılır yani ortalamaları olan 0'dan kaç standart sapma uzakta olmalarına göre çeşitli kümelere ayrılır. Gruplarda yer alan hata terimi sayısı gözlenen frekansı verir. Gözlenen frekanslar, sıfırın altında ve üstünde belli standart sapmalar arasındaki kalıntıların sıklık dağılımını gösterir.

---

<sup>145</sup> Güriş, Çağlayan, **a.g.e.**, s. 577-578.

2. Normal dağılımdan yararlanılarak, hata terimlerinin aynı aralıklar için beklenen frekansları belirlenir.
3. Test istatistiği, gözlenen ve beklenen frekanslar arasındaki farkların karesinin beklenen frekanslara bölünerek toplamların alınması ile oluşturulur.

$$\chi^2 = \sum_{i=1}^k \frac{(G_i - B_i)^2}{B_i}$$

Burada, k aralık sayısını göstermektedir.

4. Belirlenen  $\alpha$  anlamlılık seviyesinde, (k-1) serbestlik dereceli ki-kare tablo değeri belirlenir ve karar aşağıdaki şekilde alınır.

$\chi^2 \leq \chi_{\alpha, k-1}^2$  ise  $H_0$  kabul, normallik varsayımı geçerlidir,

$\chi^2 > \chi_{\alpha, k-1}^2$  ise  $H_0$  red, normallik varsayımı geçerli değildir.

#### 2.3.2.4. Jarque-Bera Testi

Jarque-Bera testinde, En Küçük Kareler yöntemiyle elde edilen hata terimlerinin asimetri ve basıklık ölçülerine dayanan test istatistiğinin,  $\chi^2$  ile karşılaştırılması ile normallik varsayımının geçerli olup olmadığı saptanmaktadır.

Test istatistiği,

$$JB = n \left[ \frac{S^2}{6} + \frac{(K-3)^2}{24} \right] \quad (2.14)$$

formülüyle hesaplanmaktadır. Buradaki S ve K sırasıyla (2.12) ve (2.13) formülleriyle hesaplanan hata terimlerinin asimetri ve basıklık ölçüleridir. Jarque ve Bera, normallik varsayımın geçerli olması durumunda JB test istatistiğinin, 2 serbestlik dereceli  $\chi^2$  dağılımına uyduğunu göstermişlerdir.<sup>146</sup>

Uygulamalarda JB test istatistiği hesaplanarak,  $\alpha$  hata payı ile 2 serbestlik dereceli ki-kare tablo değeri belirlenir ve karar aşağıdaki şekilde alınır.

$JB \leq \chi_{\alpha,2}^2$  ise  $H_0$  kabul, normallik varsayımı geçerlidir,

$JB > \chi_{\alpha,2}^2$  ise  $H_0$  red, normallik varsayımı geçerli değildir

Bowman ve Shelton'un önerdikleri testte, test istatistiği, (2.14) ile hesaplanan JB istatistiğinin aynısıdır, fakat kritik değerleri için  $\chi^2$  tablosu yerine %1 ve %5 anlamlılık seviyeleri ve çeşitli n değerleri için düzenlenen bir tablo kullanılmaktadır. Jarque-Bera testine benzer olarak JB değeri tablodan bulunan kritik değerinden büyük olduğunda hipotez reddedilecek, aksi durumda ise kabul edilecektir.

### 2.3.2.5. Kolmogorov-Smirnov Testi

Dağılıma uygunluk testlerinden biri olan Kolmogorov-Smirnov sınaması, regresyon analizinde, elde edilen hata terimlerinin dağılımının anakütlede normal dağılıma uygunluk gösterip göstermediğini araştırmak üzere kullanılabilir. Bu testte de, Ki-Kare testinde olduğu gibi hata terimleri ortalamadan sapmalara göre gruplandırılır. Her grup için, gözlenen frekans değerlerinden hareketle hesaplanan gözlenen kümülatif nispi frekans dağılımı (f) ile normal dağılımın teorik frekans dağılımı (f') arasındaki fark hesaplanır. Test istatistiği D, bu farkların mutlak değerlerinin maksimumuna eşittir.

---

<sup>146</sup> Carlos M. Jarque, Anil K. Bera, "A Test for Normality of Observations and Regression Residuals", **International Statistical Review**, Vol.LV, 1987, s. 163-172.



$$D = \max |f - f'|$$

Kritik D deęerleri, Kolmogorov-Smirnov tablolarında çeşitli  $\alpha$  ve n deęerleri için verilmektedir. Dięer testlerde olduęu gibi hesaplanan D deęeri, tablodan bulunan kritik deęerinden büyük olduęunda normallik varsayımı reddedilmekte, aksi durumda ise kabul edilmektedir.

### 2.3.2.6. Shapiro-Wilk Testi

Shapiro-Wilk testi<sup>147</sup>, tesadüfi olarak seçilen örnek verilerinin, normal dağılım gösteren bir anakütleden gelip gelmediklerini arařtırmak için kullanılan uygunluk testlerinden biridir. Küçük örnek verileri (n<50) için geliştirilmiş olan bu test, regresyon analizinde de hata terimlerinin normallik varsayımına uygun olarak dağılıp dağılmadığını arařtırmak için kullanılmaktadır.

Regresyon analizinde, En Küçük Kareler yöntemiyle elde edilen hata terimleri öncelikle en küçükten büyüęe doęru ( $e_1 < e_2 < e_3 \dots$ ) sıralanmakta ve standart sapma,

$$s = \sqrt{\frac{\sum (e_i - \bar{e})^2}{n-1}}$$

formülüyle hesaplanmaktadır. Bu deęerlerden hareketle, Shapiro-Wilk test istatistięi

$$W = \frac{\sum a_i e_i}{s}$$

eşitlięi ile belirlenmektedir. Burada  $a_1, a_2, \dots, a_n$  deęerleri, standart normal dağılımdan hesaplanan sıra istatistik<sup>148</sup> deęerlerinin beklenen deęerlerini göstermektedir. Test istatistięi W, 0 ile 1 arasında deęerler almakta, 0'a yakın deęerler dağılımın

---

<sup>147</sup> S. S. Shapiro, M. B. Wilk, "An Analysis of Variance Test for Normality (Complete Samples)", **Biometrika**, Vol. LII, 1965, s. 591-611.

<sup>148</sup> Order statistics teriminin Türkçe karşılıęı olarak kullanılmıştır.

normalden uzaklaştığını, doğal olarak 1'e yakın değerler ise normal sayılacak bir dağılımı işaret etmektedir. Normallik varsayımını ifade eden temel hipotez,

$$W \leq W_{\alpha}$$

olması durumunda reddedilmektedir.<sup>149</sup> Buradaki  $W_{\alpha}$  kritik değeri göstermektedir. Shapiro ve Wilk<sup>150</sup>, Madansky ve diğer araştırmacılar testteki sabit  $a_1, a_2, \dots, a_n$  değerleri ile çeşitli anlamlılık seviyeleri için  $W_{\alpha}$  kritik değerlerini hesaplayarak çalışmalarında tablolarla sunmaktadırlar.

Royston<sup>151</sup>, küçük örnekler için geliştirilen Shapiro-Wilk testini genişleterek  $W_{\alpha}$  kritik değerlerini  $n < 2000$  için hesaplamıştır. Leslie, Stephens ve Fotopoulos<sup>152</sup> yaptıkları araştırmada W test istatistiğinin, asimptotik dağılımı incelemiş ve tutarlılık özelliğinin bulunduğunu göstermişlerdir. Bir diğer araştırmada<sup>153</sup> ise Shapiro-Wilk test istatistiği maksimum benzerlik yöntemine dayalı olarak hesaplanmış ve analiz *nuisance regression* bulunması durumu için genişletilmiştir.

### 2.3.3. Normallik Varsayımından Sapmaların Bulunması Durumunda Önerilen Çözüm Yöntemleri

Basit veya çoklu regresyon analizinde En Küçük Kareler yöntemiyle elde edilen hata terimlerinin dağılımının normal dağılıma uymadığı bir önceki bölümde sözü edilen testlerden biri ile saptanmışsa, öncelikle aşağıda sıralanan çözüm önerileri dikkate alınmalıdır.

---

<sup>149</sup> Ashish Sen, Muni Srivastava, **Regression Analysis Theory, Methods, and Applications**, New York, Springer-Verlag, 1990, s. 105.

<sup>150</sup> Shapiro, Wilk, **a.y.**

<sup>151</sup> J. P. Royston, "An extension of Shapiro and Wilk's W Test for Normality to Large Samples", **Applied Statistics**, Vol. XXXI, 1982, s. 115-124.

<sup>152</sup> J. R. Leslie, M.A. Stephens, S. Fotopoulos, "Asymptotic Distribution of the Shapiro-Wilk W For Testing for Normality", **The Annals of Statistics**, Vol. XIV, 1986, s. 1497-1506.

<sup>153</sup> Pranab Kumar Sen, Jana Jureckova, Jan Picek, "Goodness-of-Fit Test of Shapiro-Wilk Type with Nuisance Regression and Scale", **Austrian Journal of Statistics**, Vol. XXXII, 2003, s. 163-177.

- Modelde unutulmuş bağımsız değişken veya değişkenlerin modele dahil edilmesi.
- Bölüm (2.2.4.4.)’de bahsedilen varsayım 3 ve 4’deki dönüşümlerin uygulanması.
- Örnek büyüklüğü  $n$ ’yi arttırmak.
- Ağırlıklı En Küçük Kareler yönteminin uygulanması.
- Yeniden Ağırlıklandırılmış En Küçük Kareler yönteminin uygulanması.

Bu çözüm yöntemlerinden bir veya birkaçını uygulayarak model tekrar tahmin edilmeli ve elde edilen hata terimlerinin dağılımının normale yaklaşım yaklaşmadığı saptanmalıdır. Normal dağılımın saptanmadığı durumda, literatürde En Küçük Kareler yöntemine alternatif yöntemlerden biri olarak önerilen *Güçlü Regresyon (Robust Regression)* yöntemi kullanılarak modelin parametreleri tahmin edilmektedir.

### **2.3.3.1. Güçlü Regresyon (Robust Regression)**

Güçlü regresyon özellikle modelde aşırı (uç) değerlerin bulunması durumunda önerilen bir yöntem olup hata terimleri dağılımının normal olmaması durumunda da kullanılabilen bir çözüm yöntemidir. Aşırı değerlere, bağımlı değişken ve/veya bağımsız değişkenlerde rastlanabileceği gibi standardize edilmiş hata terimlerinde ve/veya regresyon ile tahmin edilen  $y'$  değerlerinde de karşılaşılabilmektedir. Modelde aşırı değerler bulunması durumunda, genel olarak uygulanan yöntem öncelikle çeşitli testlerden biriyle bu değerlerin saptanması ve daha sonra bu değerlerin silinmesi başka bir ifade ile model dışı tutulması ve En Küçük Kareler yönteminin uygulanmasıdır. Bu yöntemin tersine, Güçlü Regresyon analizinde amaç, öncelikle verilerin çoğunluğuna uyan bir model tahmin edilmesi ve daha sonra bu modelden elde edilen hata terimindeki büyük değerlerden hareketle

aşırı değerlerin belirlenmesidir.<sup>154</sup> Huber<sup>155</sup> ise güçlülüğü (robustness), varsayımlardan küçük sapmalara duyarlı olmama olarak tanımlamaktadır.

Güçlü regresyon yöntemiyle elde edilen tahminleyenler, hatanın normal dağılması durumunda En Küçük Kareler tahminleyenleri kadar etkin, hataların normalden farklı bir dağılım (asimetrik veya simetrik olduğu halde normal olmayan bir dağılım) göstermeleri durumunda ise En Küçük Kareler tahminleyenlerinden daha etkin olmaktadır. Bunun nedeni ise Güçlü regresyon tahminleyenlerinin aşırı değerlere veya regresyon analizinin temel varsayımlarına (eşit varyanslılık ve normallik) En Küçük Kareler tahminleyenleri kadar hassas olmamasıdır.<sup>156</sup> Bu yüzden, bu yöntem sadece normalden sapmalar olması durumunda değil farklı varyanslılık durumunda da önerilen bir çözüm yöntemidir.

Güçlü regresyonda tahminleyenler, hata kareleri toplamının minimizasyonu yerine aşırı değerlere daha az hassas olan bir terimin minimize edilmesiyle elde edilmektedir. Bu terim, hata terimlerinin mutlak değerlerinin, karelerinin medyanı veya toplamı olabileceği gibi hata terimlerinin mutlak değerlerinin toplamı da olabilmektedir.

Güçlü Regresyon yönteminde, tahminleyenlerin elde edilmesinde aşağıdaki yaklaşımlar önerilmektedir.

### 1) *M-Tahminleyenleri*

M-tahminleyenleri aşağıda verilen fonksiyonun minimizasyonu ile belirlenmektedirler.

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

---

<sup>154</sup> Peter J. Rousseeuw, Annick M. Leroy, **Robust Regression and Outlier Detection**, New York, Wiley-Interscience, 2003, s. 8.

<sup>155</sup> Peter J. Huber, **Robust Statistics**, New York, Wiley-Interscience, 2004, s.1.

<sup>156</sup> Kleinbaum, Kupper, Muller, Nizam, **a.g.e.**, s. 250.

Bu denklemin regresyon parametrelerine göre türevinin alınıp 0'a eşitlenmesiyle,

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

p bilinmeyenli, denklem sistemi elde edilmektedir.<sup>157</sup> Bu eşitlikle, hata terimlerinin aşırı değerlerine hassas olmayan tahminleyenler elde etmek için önerilen  $\psi\left(\frac{e_i}{s}\right)$  fonksiyonu ile,

$$\sum_{i=1}^n \Psi\left(\frac{e_i}{s}\right) x_i' = 0 \quad (2.15)$$

eşitliğine dönüşmektedir. Bu formüldeki  $\psi(\cdot)$  fonksiyonu,  $\rho(\cdot)$  fonksiyonunun türevi olup, *etki fonksiyonu (influence function)* olarak adlandırılmaktadır. En Küçük Kareler yönteminde ise, hata terimlerinin karelerinin toplamının minimizasyonu ile  $e_i$  terimi ile doğrusal olan etki fonksiyonu  $\psi(e_i) = e_i$  olmaktadır. Fakat bilindiği gibi, elde edilen tahminleyenler hata terimlerindeki veya gözlem değerlerindeki aşırı değerlerinden etkilenmektedirler.

(2.15)'deki denklem sisteminin çözülmesiyle elde edilen tahminleyenlere M-tahminleyenleri adı verilmektedir. Bu yöntemde amaç, hata terimlerinde aşırı değerlere sahip olan gözlem değerlerinin tahminleyenlere etkisini azaltmak için onlara düşük ağırlık verecek şekilde bir etki fonksiyonu belirlemektir. Literatürde sözü edilen etki fonksiyonları arasında en yaygın olarak kullanılanı Huber<sup>158</sup> tarafından önerilen,

$$\begin{aligned} \Psi(e_i^*) &= e_i^* & |e_i^*| &\leq r \\ &= r & e_i^* &> r \\ &= -r & e_i^* &< -r \end{aligned} \quad (2.16)$$

<sup>157</sup> Myers, a.g.e., s. 349.

<sup>158</sup> Peter J. Huber, "Robust Regression: Asymptotics, Conjectures, and Monte Carlo", **Annals of Statistics** 1, 1973, s. 799-821.

fonksiyondur.<sup>159</sup> Bu fonksiyonla, hata terimlerinin hiçbirinin tahminleyenlere etkisinin  $r$  gibi bir değerden büyük olmaması sağlanmaktadır. Uygulamalarda, genellikle  $r$ 'ye 1 bazen de 1,5 veya 2 değerleri verilmektedir.  $r=1$  olması durumunda hata terimlerinin mutlak değerlerinin standart sapmalarını aşması durumunda, fonksiyonun değeri 1 veya -1 olarak belirlenmekte, böylece aşırı değerlere hassas olmayan bir fonksiyon elde edilmektedir.

***M-Tahminleyenlerinin Hesaplanması:*** (2.15) ifadesinde uygun bir  $\psi(\cdot)$  fonksiyonu seçildikten sonra çözülmesi gereken denklem sistemi doğrusal olmayan bir sistem olduğu için  $b$  parametreleri tekrarlı (iterative) yöntemlerle çözülmektedir. Önerilen çözüm yöntemleri arasında en yaygın olanı bölüm (2.2.4.4.)'de bahsedilen Yeniden Ağırlıklandırılmış En Küçük Kareler yöntemidir. Bu yöntemde, hata terimlerinin standart sapmalarının tahmini değeri  $s$  yerine daha güçlü bir tahminleyen olan ve hata terimlerinin mutlak değerlerinin medyanı ile hesaplanan

$$\hat{\sigma} = 1,5 \text{med}|e_i|$$

değerinin kullanılması önerilmektedir. Bu tahmini değerle (2.15) ifadesi

$$\sum_{i=1}^n w_i e_i^* x_i = 0 \quad (2.17)$$

şekline dönüşmektedir. Bu formülde, ağırlıklar  $w_i = \psi(e_i^*)/e_i^*$  eşitlikleriyle belirlenmektedir. (2.17) ifadesi aynı zamanda  $\sum_{i=1}^n w_i (e_i^*)^2$  teriminin minimizasyonu ile elde edilebildiği için Ağırlıklı En Küçük Kareler yöntemi, bu yöntemin hesaplama aşamalarında kullanılmaktadır. Yöntem, şu şekilde uygulanmaktadır:<sup>160</sup>

1. Parametre tahminleyenleri için bir başlangıç değeri  $b_0$  ve bu değerlerle hesaplanan  $e_{i,0}$  başlangıç değerleri elde edilmektedir.

<sup>159</sup> Bu formülde  $e_i^* = e_i/s$ 'dir.

<sup>160</sup> Myers, **a.g.e.**, s. 351-352.

2. Başlangıç değerleri baz alınarak  $\hat{\sigma}_0$  ve başlangıç ağırlıkları  $w_{i,0}$  bulunur.

3. Ağırlıklı En Küçük Kareler yöntemi kullanılarak yeni güçlü tahminleyenler,

$$b_{R_0} = (X'W_0X)^{-1} X'W_0Y$$

ile hesaplanır.

4. Bu parametreler adım 1'deki başlangıç parametrelerinin yeni değerleri gibi ele alınarak, yeni  $\hat{\sigma}$  ve  $w_i$ 'ler belirlenir.

5. Adım 3'e dönülür.

Yöntem tahminleyenler arasındaki farklar çok küçük olana kadar yani gerçek değerlere yakınsayana kadar devam etmektedir. Yöntemin başarısı, seçilen etki fonksiyonuna da bağlıdır. (2.16)'daki fonksiyon seçildiğinde  $r=3$  veya daha büyük bir değer ile hesaplanan parametreler En Küçük Kareler yönteminin tahminleyenlerine yakın olacaktır. Bu yüzden, bu değer 1 veya 1,5 olarak seçilmelidir. Birinci adımda belirlenecek başlangıç değerlerini hesaplamada kullanılacak yöntem genellikle En Küçük kareler yöntemi olmaktadır. Ancak,  $\sum_{i=1}^n |y_i - y'_i|$  teriminin minimizasyonu ile elde edilen tahminleyenler de kullanılabilir.

## 2) *Least Median Squares(LMS) Tahminleyenleri*

Rousseeuw<sup>161</sup> tarafından geliştirilen Güçlü Regresyon yönteminde hata terimlerinin kareleri küçükten büyüğe sıralandıktan sonra belirlenen medyan

---

<sup>161</sup> Peter J. Rousseeuw, "Least Median of Squares", **Journal of the American Statistical Association**, Vol.79, 1984, s. 871-880.

terimi<sup>162</sup> minimize edilerek modelin parametreleri tahmin edilmektedirler. Literatürde LMS tahminleyenleri olarak bilinen tahminleyenler,

$$\text{Min } \text{med}_i(e_i)^2$$

ile elde edilmektedirler. İlk geliştirildiğinde oldukça dikkat çeken ve literatürde sıkça bahsedilen bu yöntem daha sonra Hawkins<sup>163</sup>, Rousseeuw ve Leroy tarafından tüm boyutlarıyla ele alınıp tartışılmıştır. Rousseeuw ve van Zomeren<sup>164</sup> çalışmalarında, En Küçük Kareler yöntemine alternatif olabilecek kadar güçlü bir yöntem olmadığını göstermişlerdir. Bu yöntemin tahminleyenlerinin asimptotik etkinliğinin 0 olmasının da, yöntemin zayıf bir noktası olduğu belirtilmektedir.<sup>165</sup>

### 3) *Least Trimmed Squares(LTS)Tahminleyenleri*

Tek bir gözlem değerinin minimizasyonu ile elde edilen LMS tahminleyenlerine alternatif olarak geliştirilen Least Trimmed Squares(LTS) tahminleyenleri, hata terimlerinin kareleri küçükten büyüğe sıralandıktan sonra,

$$h = [n/2]+1$$

kadar terimin toplamının minimizasyonu ile elde edilmektedirler.

$$\text{Min. } \sum_{i=1}^h e_i^2 \quad (2.18)$$

---

<sup>162</sup> n örnek büyüklüğündeki ve p kadar değişken sayısı bulunan bir örnek için medyan  $[(n/2)+(p+1)/2]$ 'ci sıradaki hata kareleri terimidir.

<sup>163</sup> D. M. Hawkins, "The Feasible Set Algorithm for Least Median of Squares Regression", **Computational Statistics and Data Analysis**, Vol. XVI, 1993, s. 81-101.

<sup>164</sup> Peter J. Rousseeuw, B. C. van Zomeren, "Unmasking Multivariate Outliers and Leverage Points", **Journal of American Statistical Association**, Vol.LXXXV, 1990a, s. 648-651.

<sup>165</sup> Rousseeuw, Leroy, **a.g.e.**, s. 178-183.



Rousseeuw ve Leroy<sup>166</sup> *ayar yüzdesi*(trimming percentage) olarak adlandırılan  $\alpha$  terimi ile belirlenen  $h$  değerinin (2.18) formülünde kullanılmasıyla daha etkin tahminleyenler elde edilebileceğini göstermişlerdir. Bu öneriye göre  $h$ ,

$$h = [n(1-\alpha)]+1$$

formülüyle belirlenmektedir. Veriye uygun LTS tahminleyenlerini elde etmek üzere geliştirilen algoritmalar arasında Ruppert<sup>167</sup> ve Hawkins<sup>168</sup> tarafından önerilen algoritmalar dikkat çekmektedirler.

#### 4) *S Tahminleyenleri*

Rousseeuw ve Yohai tarafından geliştirilen bu yöntemde tahminleyenler, hata terimlerinin bir fonksiyonu olarak ifade edilen  $S$  fonksiyonunun minimizasyonu ile elde edilmektedir.

$$\text{Min. } S(e_1(\beta), e_2(\beta), \dots, e_n(\beta))$$

Bu formülde  $e_1(\beta), e_2(\beta), \dots, e_n(\beta)$  terimleri verilen bir  $\beta$  matrisiyle elde edilen hata terimlerini göstermektedir.  $S(e_1(\beta), e_2(\beta), \dots, e_n(\beta))$ 'in dağılımı ise aşağıdaki denklemin çözümü ile tanımlanmaktadır.

$$\frac{1}{n} \sum_{i=1}^n \rho(e_i / s) = k$$

$\rho(x)$  fonksiyonu aynı yazarlar tarafından,

---

<sup>166</sup> Rousseeuw, Leroy, **a.g.e.**, s.134-135.

<sup>167</sup> David Ruppert, "Computing S-estimators for Regression and Multivariate Location/Dispersion", **Journal of Computational and Graphical Statistics**, Vol. I, 1992, s. 253-270.

<sup>168</sup> D. M. Hawkins, "The Feasible Solution Algorithm for Least Trimmed Squares Regression", **Computational Statistics and Data Analysis**, Vol. XVII, 1994, s. 185-196.

$$\rho(x) = \begin{cases} \frac{x^2}{2} - \frac{x^4}{2c^2} + \frac{x^6}{6c^4} & |x| \leq c \\ \frac{c^2}{6} & |x| > c \end{cases}$$

şeklinde önerilmektedir.  $k$ 'yı belirleyecek olan  $c$  değerinin seçimi yöntemin etkinliği ile yakından ilgilidir. S tahminleyenleri, LMS ve LTS tahminleyenleri kadar literatürde tartışılmamasına rağmen, Rousseeuw ve Leroy tarafından asimptotik etkinliği ayrıntılarıyla incelenmiştir. Stefanski<sup>169</sup> ise bu tahminleyenlerin etkinliklerinin özellikle küçük örnek verileri için oldukça zayıf olduğunu ortaya koymuştur.

## 2.4. Çoklu Doğrusal Bağlantının Belirlenmesi ve Çözümleme Yöntemleri

Çoklu doğrusal bağlantı, çok değişkenli bir doğrusal regresyon modelinde bütün veya bazı bağımsız değişkenler arasındaki kesin veya tam doğrusal ilişkiyi ifade etmektedir.<sup>170</sup> Bilindiği gibi En Küçük Kareler yöntemi bağımsız değişkenler arasında tam veya kesin bir bağlantı bulunmadığını varsaymaktadır. Dolayısıyla, EKK yöntemiyle elde edilen sonuçların geçerli olabilmeleri için modeldeki bağımsız değişkenler arasında çoklu bağımlılığın bulunup bulunmadığı saptanmalı, bulunması durumunda alternatif yöntemleri araştırılmalıdır.

*1. Değişkenler Arasında Tam İlişki Olması:* Eğer, değişkenler arasında tam bir çoklu doğrusal bağlantı varsa, bu değişkenler arasındaki basit korelasyon katsayıları  $\pm 1$  olacaktır. Bu durumda, parametrelerin tahmini için kullanılan normal denklemlerdeki bağımsız denklem sayısı, tahmin edilecek parametrelerin sayısından düşük olacaktır. Aynı zamanda, parametre tahminlerinde kullanılan  $X'X$  matrisinin

<sup>169</sup> Leonard A. Stefanski, "A Note on High-Breakdown Estimators", **Statistics and Probability Letters**, Vol.XI, 1991, s. 353-358.

<sup>170</sup> Akın, a.g.e., s. 348.

determinantı 0 olacağından, tersi hesaplanamayacak ve dolayısı ile parametreler tahmin edilemeyecektir. Bu durumda, regresyon katsayılarının standart sapmaları hesaplandığında sonsuz çıkacaktır.

2. Değişkenler Arasında Çoklu Doğrusal Bağlantı Olmaması: Modeldeki bağımsız değişkenler arasında doğrusal ilişki yoksa, basit korelasyon katsayıları 0 olacaklardır. Bu durumda, çoklu regresyon denkleminde kısmi regresyon katsayılarının, basit doğrusal regresyon katsayılarına eşit oldukları görülür. Bir başka ifadeyle  $X'X$  matrisindeki vektörler ortogondur. Sonuç olarak, bu matrisin tersi alınarak kısmi regresyon katsayıları hesaplandığında basit regresyon katsayılarına eşit çıkacaktır.

Her iki durum da, uç noktalar. Uygulamalarda, bağımsız değişkenler arasında tam ilişki olması veya hiç ilişki olmamasından çok, bağımsız değişkenler arasında önemli derecede çoklu doğrusal bağlantı ile karşılaşmaktadır.

#### **2.4.1. Çoklu Doğrusal Bağlantının Nedenleri**

Çoklu doğrusal bağlantının nedenleri aşağıdaki gibi sıralanabilir:<sup>171</sup>

- Modeldeki bağımsız değişkenlerin aynı trendde sahip olmaları,
- Modelin değişken sayısının gözlem sayısından çok olması,
- Kullanılan veri toplama yönteminin uygun olmaması.

#### **2.4.2. Çoklu Doğrusal Bağlantının Sonuçları**

Modelde kuvvetli bir çoklu doğrusal bağlantının bulunmasının En Küçük Kareler yönteminin bulgularına ve modelin istatistik özelliklerine etkisi aşağıdaki gibi sıralanabilir:<sup>172</sup>

---

<sup>171</sup> Douglas Montgomery, Elizabeth Peck, **Introduction to Linear Regression Analysis**, New York, Jon Wiley & Sons, 1982, s. 289-290.

- En Küçük Kareler tahminleyenleri en iyi doğrusal yansız tahminleyenlerdir. Ancak tahmini güçleştiren büyük varyanslara ve kovaryanslara sahip olmalarıdır.
- Bulunan kısmi regresyon katsayıları bir örnek küleden diğerine büyük değişiklikler göstermektedir. Bunun nedeni tahminleyenlerin varyanslarının büyük olmasıyla güven aralıklarının geniş olma eğilimi göstermesi ve dolayısıyla kısmi regresyon katsayılarının anlamlı ve güvenilir olamamasıdır. Bir başka ifade ile, b parametreleri test edildiğinde,  $H_0$  kabul yönünde sonuçlar verecektir (t istatistiği, paydası çoklu doğrusal bağlantı yüzünden büyüyeceği için değeri küçülecektir).
- Bir veya birden fazla katsayının t testi anlamsız çıktığı halde, bağımsız değişkenlerin aralarındaki basit korelasyon katsayıları ve F testi ile anlamlılığı test edilen çoklu korelasyon katsayısı çok yüksek çıkabilmektedir. Bunun nedeni, r'nin çoklu doğrusal bağlantı nedeniyle suni olarak yükselmesidir.
- Kısmi regresyon katsayılarının işaretleri beklenenden farklı çıkabilir.
- Örnek birim sayısının biraz arttırılması veya azaltılması gibi verilerde ufak bir değişiklik parametre tahminlerinde önemli etkilere neden olmaktadır.

### 2.4.3. Çoklu Doğrusal Bağlantının Belirlenmesi

Kmenta'ya göre, anlamlı bir ayırım çoklu doğrusal bağlantı varsayımının varlığı veya yokluğu üzerine değil çeşitli dereceleri arasında yapılmalıdır.<sup>173</sup>

<sup>172</sup> Gujarati, a.g.e., s. 327-335.

<sup>173</sup> Kmenta, a.g.e., s. 380.

Çoklu doğrusal bağlantının gücünü ölçmede tek bir yöntem yoktur. Bir önceki bölümde sözü edilen sonuçların tümü dikkate alınabileceği gibi, etkili olduğuna kesin karar verebilmek için aşağıda sıralanan biçimsel yaklaşımlar da ele alınabilmektedir.

### **2.4.3.1. Varyans Artış Faktörleri (VIF-Variance Inflation Factors)**

Varyans Artış Faktörleri, çoklu doğrusal bağlantının varlığı ile tahminleyenlerin varyanslarının nasıl genişlediğini gösterir. Çünkü, çoklu doğrusal bağlantının göstergelerinden biri de parametre tahminlerinin varyanslarından meydana gelen artıştır. Fakat varyanslarda meydana gelen artışları doğrudan çoklu doğrusal bağlantının sonucu olarak yorumlamak doğru değildir. Parametre tahminleyenlerinin varyansı (1.5) formülünden hesaplanacağı gibi,

$$V(b_i) = \frac{s^2}{1 - R_i^2}$$

formülüyle de belirlenebilmektedir. Burada s, örnek regresyon denkleminin standart hatası,  $R_i^2$  de incelenen modelde yer alan i. bağımsız değişkenin bağımlı değişken olup modeldeki diğer bağımsız değişkenlerin bağımsız değişken oldukları regresyon modelinin belirlilik katsayısıdır. Bu formülden de anlaşılacağı gibi, varyanslardaki büyüme bağımsız değişkenlerin birbirine yakın değerler alması ile  $R_i^2$  nin büyümesinden de kaynaklanabilir. Fakat bu büyümenin çoklu doğrusal bağlantıdan kaynaklanıp kaynaklanmadığını ve etkisinin önemli olup olmadığını belirlemek için VIF kriterinden yararlanılır. VIF kriteri, i. bağımsız değişken için,<sup>174</sup>

---

<sup>174</sup> Neter, Wasserman, Kutner, **a.g.e.**, s. 391.

$$VIF_i = \frac{1}{1 - R_i^2}$$

formülüyle hesaplanmaktadır. VIF, bir parametre tahmininin doğrusal bağlantı nedeniyle kesinlikten ne derece uzaklaştığının bir ölçüsüdür.  $R_i^2$  1'e yaklaştığında  $VIF_i$  değeri büyük olacaktır.  $R_i^2=1$  ise  $VIF_i \rightarrow \infty$  olacaktır. VIF'in alt limiti 1'dir ( $R_i^2=0$  olması durumu). Bu uç durumlar dışında, VIF kriterinin hangi değerleri için çoklu doğrusal bağlantının etkili olacağı konusunda kesin bir kural olmamasına rağmen, genellikle VIF değerleri 10 ile karşılaştırılmakta ve 10'dan büyük olduğunda ( $R_i^2 > 0,90$  olduğunda) çoklu doğrusal bağlantının önemli olduğu kabul edilmektedir.<sup>175</sup> İncelenen modelde kaç tane bağımsız değişken varsa o kadar VIF değeri hesaplanmakta, en az birinin 10'dan büyük olması durumunda modelde çoklu doğrusal bağlantının etkili olduğuna karar verilmektedir.

### 2.4.3.2. Yardımcı Regresyon Modelleri

Bu yöntemde VIF kriteri için hesaplanan belirlilik katsayılarından (bağımsız değişkenler arasındaki çoklu korelasyon katsayıları) yararlanılır. Modelde bulunan her bağımsız değişken için, diğer bağımsız değişkenlerle olan regresyon modelleri oluşturulur. Yardımcı regresyon modelleri adı verilen bu modeller:<sup>176</sup>

$$X_{1t} = a_1 + a_2 X_{2t} + \dots + a_{p-1} X_{p-1t} + u_t$$

$$X_{2t} = a_1 + a_3 X_{3t} + \dots + a_{p-1} X_{p-1t} + u_t$$

.

.

.

$$X_{p-1t} = a_1 + a_2 X_{2t} + \dots + a_{p-2} X_{p-2t} + u_t$$

<sup>175</sup> Kleinbaum, Kupper, Muller, Nizam, **a.g.e.**, s. 241.

<sup>176</sup> Gujarati, **a.g.e.**, s. 337.

şeklinde tanımlanırlar. Bu modellerin belirlilik katsayıları hesaplanarak her model için F test istatistiği aşağıdaki formülle hesaplanır.<sup>177</sup>

$$F_i = \frac{R_i^2 / (p - 2)}{(1 - R_i^2) / (n - p + 1)}$$

Burada  $R_i^2$ , i. bağımsız değişken  $X_i$ 'nin bağımlı değişken olduğu modelin belirlilik katsayısıdır. Bu testte temel hipotezler bağımsız değişkenler arasında ilişki olmaması durumunu, doğal olarak alternatif hipotezler de ilişki bulunması durumunu ifade etmektedirler. Hesaplanan test istatistiği,  $\alpha$  anlamlılık seviyesi, (p-2) ve (n-p+1) serbestlik dereceli F tablosundan bulunacak değer ile karşılaştırılır ve karar şu şekilde alınır:

$F_i \leq F_{(\alpha; p-2; n-p+1)}$  ise  $H_0$  kabul, bağımsız değişkenler arasında ilişki yoktur.

$F_i > F_{(\alpha; p-2; n-p+1)}$  ise  $H_0$  red, bağımsız değişkenler arasında anlamlı bir ilişki vardır.

### 2.4.3.3. Klein Kuralı

Modelde çoklu doğrusal bağlantının çözüm aramak gerekecek kadar etkili olup olmadığını araştırmada kullanılabilecek bir diğer biçimsel kriter de Klein<sup>178</sup> tarafından önerilmiştir. Bu kriterde, yardımcı regresyon modelleri tahmin edilerek bunlardan elde edilecek çoklu belirlilik katsayıları  $R_i^2$ 'ler modelin genel çoklu belirlilik katsayısı  $R^2$  ile karşılaştırılır. Klein kuralına göre,

$R_i^2 < R^2$  ise çoklu doğrusal bağlantı önemli değildir.

---

<sup>177</sup> Güriş, Çağlayan, **a.g.e.**, s. 557.

<sup>178</sup> Lawrence R. Klein, **Introduction to Econometrics**, New Jersey, Prentice-Hall, 1962, s. 101.

#### 2.4.3.4. Durum Sayısı (Condition Number)

Bu yöntemde, En Küçük kareler yönteminin tahminleyenlerinin hesaplanması için kullanılan  $X'X$  matrisinin determinantından faydalanarak çoklu doğrusal bağlantının modelde etkili olup olmadığına karar verilmektedir. Bilindiği gibi çoklu doğrusal bağlantı arttıkça bu matrisin determinant değeri mutlak olarak azalmakta ve tam çoklu doğrusal bağlantı durumunda 0 olmaktadır. Söz konusu determinant değeri çoklu doğrusal bağlantının derecesi hakkında tam bir fikir vermeyeceği için  $X'X$  matrisinin özdeğerlerinden yararlanılmaktadır. Bu matrisin en büyük özdeğeri  $\lambda_1$ , en küçük özdeğeri  $\lambda_2$  ise,

$$\text{Durum sayısı} = \sqrt{\frac{\lambda_1}{\lambda_2}}$$

formülüyle hesaplanır.<sup>179</sup> Genellikle, durum sayısı 10 ile 30 arasında değer alırsa orta derecede, 30'dan büyükse kuvvetli derecede çoklu doğrusal bağlantıdan söz edilebileceği kararı alınır. Bazı yazarlar durum sayısını, modelde genel olarak çoklu doğrusal bağlantının bulunup bulunmadığının en iyi göstergesi olarak kabul etmektedirler<sup>180</sup>, ancak bu görüş çok yaygın değildir.

#### 2.4.4. Çoklu Doğrusal Bağlantı için Çözüm Yolları

Modelde, çoklu doğrusal bağlantının etkin olduğuna bir önceki bölümde sözü edilen yöntemlerden biriyle saptanmışsa, En Küçük Kareler yönteminin varsayımlarından bir sağlanmamış olmaktadır. Bu durumu düzeltmek için tek bir yöntem bulunmamaktadır. Bu bölümde çoklu doğrusal bağlantının ortadan kaldırılması için literatürde en çok sözü edilen yollar<sup>181</sup> ve yöntemler ele alınacaktır.

---

<sup>179</sup> Maddala, **a.g.e.**, s. 274.

<sup>180</sup> David A. Belsley, Edwin Kuh, Roy E. Welsch, **Regression Diagnostics: Identifying Influential Data and Sources of Collinearity**, New York, John Wiley & Sons, 1980, s. 103-107.

<sup>181</sup> Güriş, Çağlayan, **a.g.e.**, s. 563-565.



1. Ön Bilgi Kullanma: Çoklu doğrusal bağlantıya neden olan bağımsız değişkenlere ait bilgiler doğrusal bağlantının daha az ciddi olduğu önceki deneyimsel çalışmalardan veya inceleme alanının gerisinde yatan teorik kavramlardan elde edilebiliyorsa, bu bilgiler kullanılarak çoklu doğrusal bağlantı düzeltilmeye çalışılabilir.

2. Çoklu Doğrusal Bağlantıya Neden Olan Değişken(ler)in Modelden Çıkarılması veya Yeni Değişken(lerin) Modele Eklenmesi: Çoklu doğrusal bağlantıya neden olan değişkenlere modelde yer verilmezse çoklu doğrusal bağlantı ortadan kaldırılabilir. Ancak modelden çıkartılan değişkenin teorik açıdan önemli bir değişken olması söz konusu ise tanımlama hatası bir başka ifade ile model kurma hatası yapılmış olacaktır. Bu durumda, bu değişkeni modelden atmak yerine diğer yolların denenmesi daha iyi olacaktır. Bazen de, modelden değişken atmak yerine yeni değişken(ler)in eklenmesi ile söz konusu olan bağımsız değişkenlerin varyanslarının azaldığı görülmektedir. Bunun nedeni teorik olarak modele dahil edilmesi gereken değişken(ler)in unutulmuş olmasıdır.

3. Ek veya Yeni Veri: Bazı durumlarda yalnızca örnek birim sayısını arttırmakla çoklu doğrusal bağlantının etkisi azaltılabilir. Burada önemli olan çoklu doğrusal bağlantının örnek birim sayısının artması ile değişkenler arasındaki doğrusal ilişkinin derecesinin azalmasıdır. Bazı çalışmalarda da, yeni bir örnek alınması veya incelenen dönemin değiştirilmesi ile çoklu doğrusal bağlantı ortadan kaldırılabilir. Fakat ek veya yeni veri toplanması her zaman kolay değildir veya incelenen dönemin değiştirilmesi çalışmanın amacına uygun olmayabilir.<sup>182</sup>

Çoklu doğrusal bağlantının düzeltilmesinde kullanılan yöntemler aşağıdaki sırada ele alınmaktadır.

---

<sup>182</sup> Judge, vd., **a.g.e.**, s. 625.

### 2.4.4.1. Bağımsız Değişkenlerin Modelden Çıkarılması-Adım Adım Regresyon (Stepwise Regression)

Model kurma yöntemlerinden biri olan Adım Adım Regresyon yöntemi, modelde bulunması anlamlı olan bağımsız değişkenleri, her adımda yeni bir değişken seçerek ve/veya daha önce seçilmiş bir değişkeni modelden atarak belirlemektedir. Yöntem uygulandığında, bağımlı değişkeni en çok etkileyen (modelde bulunması anlamlı) ve aynı zamanda aralarında kuvvetli ilişkiler bulunmayan bağımsız değişkenler belirlenmiş olmaktadır. Bu yüzden, modelde bulunması düşünülen tüm bağımsız değişkenler ile kurulan modelde çoklu doğrusal bağlantı saptanmışsa, bu duruma neden olan bağımsız değişken(ler)in elenmesi için kullanılacak bir yöntem olmaktadır. Ancak, modelde teorik olarak kesinlikle bulunması gereken değişkenler Adım Adım Regresyon yöntemiyle model dışı tutuluyorsa, diğer yöntemler uygulanarak model gerekli bağımsız değişken kombinasyonu ile oluşturulmalıdır. Bu yöntem aşağıdaki adımlarla uygulanmaktadır:<sup>183</sup>

1. İlk adımda, bağımlı değişkeni en çok etkileyen bağımsız değişken belirlenerek basit doğrusal regresyon modeli oluşturulmaktadır. Modele eklenebilecek (p-1) adet bağımsız değişken bulunduğu varsayımıyla, her bir basit doğrusal regresyon modelinin korelasyon katsayısının anlamlılığı F testi ile sınılanmaktadır. k. bağımsız değişken ile oluşturulan modelin F test istatistiği,

$$F_k^* = \frac{RKT(x_k)}{HKT(x_k)/(n-2)} = \frac{r_{yk_i}^2}{(1-r_{yk}^2)/(n-2)}$$

formülüyle hesaplanmaktadır. Tüm modellerin F değerleri karşılaştırılarak en yüksek  $F^*$  test değeri belirlenmekte, bu değer daha önce belirlenen sabit bir değerden  $F_{IN}$  (modele girecek değişken için, belirlenen bir  $\alpha$  anlamlılık düzeyine göre F tablosundan bulunacak

<sup>183</sup> Netter, Wasserman, Kutner, **a.g.e.**, s. 430-435.

değer) büyük ise bu değişken modele eklenmektedir. Böylece bağımsız değişkeni açıklamada en kuvvetli bağımsız değişken belirlenmektedir.  $F^*$  değeri belirlenen anlamlılık düzeyini aşmadığı taktirde algoritma durdurulmaktadır. Bu durumda, modele eklenecek anlamlı tek bir değişken dahi bulunmamaktadır.

2. İlk adımda, modele eklenen bağımsız değişkenin k. değişken  $X_k$  olduğu varsayımıyla, diğer tüm değişkenler tek tek modele eklenerek iki bağımsız değişkenli regresyon modelleri oluşturulur. Bu modellerin kısmi korelasyon katsayılarının F test istatistikleri hesaplanır. Modele i. değişkenin de eklendiği varsayımıyla  $F^*$  değeri,

$$F_i^* = \frac{RKT(x_i | x_k)}{HKT(x_i, x_k)/(n-3)} = \frac{r_{y,ik}^2 - r_{yk}^2}{(1 - r_{y,ik}^2)/(n-3)}$$

hesaplanır. Modele eklenecek ikinci bağımsız değişken adayı, en yüksek kısmi F test istatistik değerine sahip olandır. Bu değer daha önceki adımda olduğu gibi tablo değeri  $F_{IN}$  ile karşılaştırılmakta ve bu değeri aşması halinde anlamlı bulunarak ilgili bağımsız değişken modele eklenmekte aksi taktirde algoritma durdurulmakta ve model tek değişkenli olarak oluşturulmaktadır.

3. İkinci adımda i. değişkenin modele alındığı varsayımıyla, bu adımda algoritmanın ilk adımında eklenen k. değişkenin modelde kalmasının anlamlı olup olmadığı sınanmaktadır.

$$F_k^* = \frac{RKT(x_k | x_i)}{HKT(x_i, x_k)/(n-3)} = \frac{r_{y,ik}^2 - r_{yi}^2}{(1 - r_{y,ik}^2)/(n-3)}$$

Bu değer, sabit bir değer  $F_{OUT}$  (modelden çıkacak değişken için, belirlenen bir  $\alpha$  anlamlılık düzeyine göre F tablosundan bulunacak

değer) ile karşılaştırılmakta ve bu değerden küçükse k. değişken modelden atılmakta, aksi takdirde modelde kalmaya devam etmektedir.

4. Daha sonraki adımlarda öncelikle modele eklenecek anlamlı bir değişken bulunup bulunmadığı belirlenmektedir. Bu amaçla, her adımda modele yeni değişkenler teker teker eklenerek ilgili kısmi F test istatistiklerinden en yüksek olan belirlenmekte ve  $F_{IN}$  değerini aşması durumunda modele eklenmektedir. Modele eklenen her bir değişkenden sonra, modele daha önce eklenen değişkenlerden herhangi birinin modelden atılmasının anlamlı olup olmadığı 3. adımda anlatılan yaklaşımla sınanmaktadır. Modele eklenecek veya modelden atılacak bir değişken bulunmaması durumunda algoritma durdurulmaktadır. Bu noktada, algoritmanın belirlediği son model bağımlı değişkeni en iyi açıklayan model olarak saptanmaktadır.

Adım Adım Regresyon yönteminde, ilk adımlarında anlamlı bulunup modele eklenen bağımsız değişkenler daha sonraki adımlarda anlamsız bulunup modelden atılabilmektedir. Bunun nedeni diğer eklenen bağımsız değişkenler ile aralarındaki korelasyonun yüksek olması ve bağımlı değişkeni açıklamakta yetersiz kalmalarıdır.

#### **2.4.4.2. Tarafli Tahminleyenler Yöntemi (Ridge Regression)**

Hoerl ve Kennard<sup>184,185</sup> tarafından önerilen Tarfli Tahminleyenler yöntemi, çoklu doğrusal bağlantı problemini çözmek için kullanılabilen yöntemlerden biri olup En Küçük Kareler yöntemini değiştirerek, regresyon katsayılarının tarafli tahminlerini kullanmaya olanak sağlar. Bu yöntem, modelde çoklu doğrusal bağlantı bulunması durumunda uygulandığında, regresyon katsayılarında gerçekleşmesi beklenen sonuçlara ulaşmayı ve bu sonuçlara ulaşırken varyansların küçülmesini sağlamaktadır. Eğer bir regresyon katsayısının tarafli tahminleyeni, tarafsız

---

<sup>184</sup> A. E. Hoerl, R. W. Kennard, "Ridge Regression: Biased Estimation for Nonorthogonal Problems", **Technometrics**, Vol. XII, 1970a, s. 55-67.

<sup>185</sup> A. E. Hoerl, R. W. Kennard, "Ridge Regression: Applications to Nonorthogonal Problems", **Technometrics**, Vol. XII, 1970b, s. 69-82.

tahminleyeninden daha hassas ise (varyansı daha küçük ise) gerçek kısmi regresyon değerine, tarafsız katsayı tahminleyene oranla daha yakın demektir.<sup>186</sup>

Bu yöntemi üstün kılan önemli bir özelliği, Adım Adım Regresyon yöntemindeki gibi modeli oluştururken bağımlı değişkenlerin anlamlı olanlarını seçmek değil, söz konusu tüm değişkenlerin modele dahil edilmesidir. Teorik olarak, model düşünülen tüm değişkenler ile oluşturulacaksa bu yöntemin kullanılması önerilmektedir.

Tarafli tahminleyenler modelini elde etmek için öncelikle tüm değişkenlere korelasyon dönüşümü uygulanmaktadır. Bunun nedeni, çoklu doğrusal bağlantı bulunması durumunda, En Küçük Kareler tahminleyenlerini hesaplamada kullanılan ters matris işlemlerinde determinant değerinin 0'a çok yakın olması ve dolayısıyla tahminleyenlerin belirlenememesidir.

Dönüştürülmüş değişkenler, standardize değişkenler ile aşağıdaki formüller baz alınarak hesaplanır:

$$\begin{aligned} y' &= \frac{1}{\sqrt{n-1}} \cdot \left( \frac{y - \bar{y}}{s_y} \right) \\ x_1' &= \frac{1}{\sqrt{n-1}} \cdot \left( \frac{x_1 - \bar{x}_1}{s_1} \right) \\ &\cdot \\ &\cdot \\ &\cdot \\ x_{(p-1)}' &= \frac{1}{\sqrt{n-1}} \cdot \left( \frac{x_{(p-1)} - \bar{x}_{(p-1)}}{s_{(p-1)}} \right) \end{aligned} \quad (2.19)$$

<sup>186</sup> Neter, Wasserman, Kutner, **a.g.e.**, s. 394.

Bu deęişkenler kullanıldığında, çoklu regresyon modeli:

$$y' = b'_1x'_1 + b'_2x'_2 + \dots + b'_{(p-1)}x'_{(p-1)} + e'$$

denklemiyle ifade edilebilir.<sup>187</sup>

Bilindięi gibi, matris notasyonu kullanıldığında En Küçük Kareler tahminleyenleri,

$$b = (X'X)^{-1} X'Y$$

formülüyle elde edilmektedir. Çoklu doğrusal bağlantı söz konusu olduğunda,  $(X'X)^{-1}$  matrisinin köşegen elemanları çok büyük deęerler almaktadır. Bu sorunu çözmek için Hoerl ve Kennard En Küçük Kareler tahminleyenleri yerine, taraflı tahminleyenleri,

$$b_R = (X'X + cI)^{-1} X'Y$$

önermişlerdir. Burada,  $X'X$  matrisi, deęişkenlere (2.19) dönüşümleri uygulandığında deęişkenler arasındaki basit korelasyon matrisi olarak elde edilmektedir.

$$X'X = \begin{bmatrix} 1 & r_{12} & \dots & r_{1,p-1} \\ r_{21} & 1 & \dots & r_{21,p-1} \\ \vdots & \vdots & . & \vdots \\ r_{p-1,1} & \dots & \dots & 1 \end{bmatrix}$$

$X'Y$  matrisi ise, bağımlı deęişkenle her bir bağımsız deęişkenin aralarındaki basit korelasyon katsayılarını içermektedir.

---

<sup>187</sup> A.e. s. 395.

$$X'Y = \begin{bmatrix} r_{y1} \\ r_{y2} \\ \vdots \\ r_{y,p-1} \end{bmatrix}$$

Daha önce bahsedildiği gibi, Taraflı Tahminleyenler yönteminde (2.19) dönüşümlerinin önerilmesinin nedeni, modelde çoklu bağlantı bulunması durumunda,  $X'X$  matrisinin determinantının 0'a çok yakın olması ve bu durumun ters alma işlemlerinde yuvarlatma hatalarına neden olmasıdır. Korelasyon dönüşümü uygulandığında değişkenler standardize edilerek birimlerinden bağımsız hale getirilmekte ve böylece  $X'X$  ve  $X'Y$  matrislerinin elemanları basit korelasyon katsayıları ile ifade edilmekte, yani -1 ile +1 arasında değerler almaktadırlar.

Taraflı tahminleyenlerin taraflı olmalarının nedeni, köşegen elemanlarına ilave edilen ve *taraflılık sabiti (ridge parameter)*<sup>188</sup> olarak bilinen  $c$  değeridir. Bir diğer ifadeyle,

$$\begin{aligned} b_R &= (X'X + cI)^{-1} X'Y \\ &= (X'X + cI)^{-1} X'Xb \end{aligned}$$

olarak yazılırsa, taraflı tahminleyenin beklenen değeri

$$E(b_R) = (X'X + cI)^{-1} X'XE(b) = (X'X + cI)^{-1} X'X\beta$$

olacaktır. Görüldüğü gibi, beklenen değer anakütle parametresine eşit değil, onun taraflı bir tahminleyenidir.  $b_R$ 'nin varyans-kovaryans matrisi ise,

$$V(b_R) = \sigma^2 (X'X + cI)^{-1} X'X (X'X + cI)^{-1}$$

---

<sup>188</sup> Shrinkage parameter olarak da adlandırılmaktadır.

formülüyle elde edilmektedir.

Bu yöntemin en kritik noktası, bilinmeyen  $c$  değerini doğru olarak belirlemektir. Bu değer seçilirken ortalama hata karesini azaltmaya ve taraflılığı mümkün olduğu kadar küçük tutmaya özen gösterilmelidir. Hoerl ve Kennard, taraflı tahminleyenlerin hata karelerinin ortalamasının, En Küçük Kareler tahminleyenlerinin hata karelerinin ortalamasından küçük olmasını sağlayan bir  $c > 0$  sabitinin her zaman bulunabileceğini göstermişlerdir.<sup>189</sup>

$$\sum_{i=1}^{p-1} HKO(b_R) < \sum_{i=1}^{p-1} HKO(b)$$

Bu yüzden,  $c$  pozitif ve 0 ile 1 arasında değer alması önerilmektedir.  $c=0$  olması durumunda taraflı tahminleyenler, En Küçük Kareler tahminleyenlerine dönüşmektedir.  $c$  değerinin 0'dan büyük seçilmesi sonucunda ortaya çıkan toplam sapmanın karesi,

$$\sum_{i=1}^{p-1} [E(b_{iR}) - \beta_i]^2 = c^2 \beta' [XX + cI]^{-2} \beta \quad (2.20)$$

değerine eşit olmaktadır. Taraflı tahminleyen yönteminin başarısı, seçilen bir  $c$  değerine karşı hesaplanan parametre varyanslarındaki azalmanın (2.20) ifadesindeki yanlı terimden büyük olmasıyla ölçülmektedir.<sup>190</sup>

**Taraflılık Parametresinin Seçimi:** Taraflılık sabiti, bilinmeyen parametre ve varyansın fonksiyonu olduğu için seçilen değer doğruluğundan emin olmak için bazı şartların sağlanması gerekmektedir. Bu şartlar, modelin seçilen  $c$  değeri ile dengede olması, elde edilen parametrelerin teorik beklentilere cevap vermesi ve hata kareleri toplamının yükselen değerlere sahip olmaması olarak sıralanabilirler.<sup>191</sup>

---

<sup>189</sup> Maddala, **a.g.e.**, s. 281.

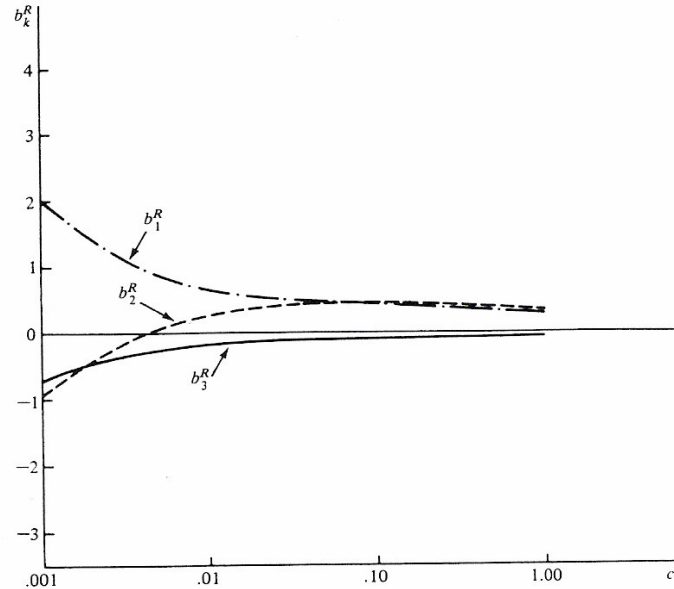
<sup>190</sup> Myers, **a.g.e.**, s. 395.

<sup>191</sup> Güriş, Çağlayan, **a.g.e.**, s. 567.



Uygulamalarda  $c$ 'nin değeri, 0'dan başlanarak 0,01 gibi küçük değer artışları ile 1'e kadar arttırılmakta ve bu değerlere karşı düşen taraflı tahminleyenler hesaplanmaktadır. En doğru  $c$ 'yi seçmek için önerilen en yaygın yollar şu şekilde sıralanabilirler:

1. Ridge İzi (Ridge Trace): Çoklu bağlantının etkisini görebilmek için grafiksel bir yöntem olan Ridge İzi'nden yararlanılır. Ridge İzi, yatay ekseninde  $c$  parametresi, düşey ekseninde de  $c$ 'nin değerlerine karşı düşen regresyon katsayılarını gösterecek şekilde iki boyutlu uzayda elde edilir. Bu grafiği çizmekteki amaç, En Küçük Kareler'den daha küçük hata kareleri ortalamasını veren  $c$  değerini bulmak ve kararlı katsayıları belirlemektir. Kararlı katsayılar, verilerdeki küçük değişikliklere duyarlı olmayanlardır. Bağımsız değişkenler arasında kuvvetli ilişkiler varsa  $c$ 'nin küçük değerleri için katsayılar çok hızlı değişecek fakat  $c$ 'nin değeri büyüdükçe derece derece kararlı olacaklardır. Burada amaç, Ridge İzi'nin görsel olarak sunduğu katsayılardaki değişmelerin kararlı olmaya başladıkları  $c$  değerini seçmektir.<sup>192</sup>



**Şekil (2.8): Ridge İzi**

<sup>192</sup> Orhan İpek, "Ridge Regresyon Üzerine Bir Çalışma", (Çevrimiçi) <http://idari.cu.edu.tr/sempozyum/bil28.htm>, 28 Ocak 2005.

2. c değeri belirlenirken, Varyans Artış Faktörleri VIF değerlerinden de yararlanılması önerilmektedir. Uygulamalarda, her bir c değerine karşı gelen tüm bağımsız değişkenlerin VIF değerleri hesaplanır. Tüm VIF değerlerinin 10'dan küçük olduğu c değerinin seçilmesine dikkat edilmelidir.<sup>193</sup>

3. c'nin seçiminde,

$$c = \frac{\sigma^2}{b'b}$$

formülü de önerilmektedir.<sup>194</sup>

Bu yöntemler dikkate alınarak seçilen c değeriyle hesaplanan dönüştürülmüş değişkenlerle regresyon modeli oluşturulmakta ve daha sonra değişkenlere aşağıdaki geri dönüşümler uygulanarak orijinal regresyon modeli belirlenmektedir.

$$b_1 = \frac{s_y}{s_1} b'_1$$

$$b_2 = \frac{s_y}{s_2} b'_2 \quad (2.21)$$

.

.

.

$$b_{(p-1)} = \frac{s_y}{s_{p-1}} b'_{p-1}$$

$$b_0 = \bar{y} - b_1 \bar{x}_1 - b_2 \bar{x}_2 - \dots - b_{(p-1)} \bar{x}_{(p-1)}$$

<sup>193</sup> D. W. Marquardt, R. D. Snee, "Ridge Regression in Practice", **The American Statistician**, Vol. XXIX, No. 1, 1975, s. 4-9.

<sup>194</sup> Güriş, Çağlayan, **a.g.e.**, s. 569.

Shi ve Wang<sup>195</sup>, verilerde etkili (influential) gözlemlerin bulunması durumunda bu gözlemlerin çoklu doğrusal bağlantı ve taraflı tahminleyenlere etkilerini incelemiştir.

Hawkins ve Yin<sup>196</sup> ise, regresyon analizinde parametre sayısının gözlem sayısından büyük olma durumunu ele alarak çoklu doğrusal bağlantı durumunda, taraflı tahminleyenleri elde etmede gerekli hesaplama sayısını azaltmak için bu yöntemde bazı değişiklikler önermişlerdir.

Ngo, Kemeny ve Deak<sup>197</sup>, taraflı tahminleyen yöntemini üç değişik çoklu regresyon modelinde incelemiştir. Değişkenlerin büyüklük olarak çok yakın oldukları ilk modelde, taraflı tahminleyenler yöntemi önce değişkenlerin özgün değerleriyle, daha sonra korelasyon dönüşümü kullanılarak uygulanmış ve özgün değerleriyle elde edilen tahminleyenlerin diğerine göre daha iyi olduğu sonucuna varılmıştır. Değişkenlerin büyüklük olarak çok farklı oldukları ikinci modelde ise dönüştürülmüş değişkenlerle elde edilen sonuçların en iyi olduğu gösterilmiştir. Ele alınan son örnekte ise taraflı tahminleyenler yöntemi genişletilerek doğrusal olmayan bir regresyon modeline uygulanmıştır.

#### **2.4.4.3. Asal Bileşenler Yöntemi (Principal Components Method)**

Asal Bileşenler yöntemi, aralarında yüksek korelasyonlar bulunan değişkenleri birbirleriyle ilişkisiz değişkenlere dönüştürmek, bir veri kümesinde değişkenliği yüksek veya düşük olan doğrusal kombinasyonu saptamak ve modeli en

---

<sup>195</sup> Lei Shi, Xueren Wang, "Local Influence in Ridge Regression", **Computational Statistics & Data Analysis**, Vol. XXXI, 1999, s. 341-353.

<sup>196</sup> Douglas M. Hawkins, Xiangrong Yin, "A Faster Algorithm for Ridge Regression", **Computational Statistics & Data Analysis**, Vol. XL, 2002, s. 253-262.

<sup>197</sup> S.H. Ngo, S. Kemeny, A. Deak, "Performance of the Ridge Regression Method as Applied to Complex Linear and Nonlinear Models", **Chemometrics and Intelligent Laboratory Systems**, Vol. LXVII, 2003, s. 69-78.

az sayıda deęişken ile tanımlamak gibi üç temel amaç için kullanılan bir yöntemdir.<sup>198</sup>

Çoklu doğrusal bağlantı söz konusu olduğunda ise, bir dönüşümden yararlanılarak birbirleriyle ilişkili X bağımsız deęişkenleri yerine onların asal bileşenleri olan ve birbirleriyle ilişkileri olmayan Z deęişkenlerinin kullanılması yöntemidir. Bu yöntem Tarafalı Tahminleyenler yöntemi gibi yanlı bir tahmin yöntemidir.

Çoklu regresyon modeli matris notasyonu ile,

$$Y = Xb + e$$

ifadesiyle ele alındığında, ve deęişkenlere (2.19) dönüşümleri<sup>199</sup> uygulandığında  $X'X$ ,  $X'Y$  matrisleri korelasyon katsayılarını içermektedirler. Aynı model için<sup>200</sup>,

$$V'(X'X)V = \Lambda \quad \text{ve} \quad V'V = VV' = 1 \quad (2.22)$$

şartlarını sağlayan bir V matrisi bulunmaktadır, öyle ki  $\Lambda$ ,  $X'X$  matrisinin sıralı özdeęerlerini içeren bir köşegen matrisidir. Özdeęerler  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{(p-1)}$  ile ifade edilmektedirler.

Bilindięi gibi bir matrisin özdeęerleri ve özvekörleri şu şekilde elde edilmektedir:

A matrisi  $n \times n$  boyutunda simetrik bir matris ise,  $x'Ax$  karesel ifadeyi  $x'x=1$  şartını sağlayacak şekilde minimize etmek için,  $\lambda$  Lagranj çarpanının da ifadeye girmesiyle oluşan

---

<sup>198</sup> Bernhard Flury, **Common Principal Components & Related Multivariate Models**, New York, John Wiley & Sons, 1988, s.5.

<sup>199</sup> Tarafalı Tahminleyenler yönteminde korelasyon dönüşümlerinin uygulanma nedenleri bu yöntem için de geçerlidir.

<sup>200</sup> Chatterjee, Price, **a.g.e.**, s. 207.

$$x'Ax - \lambda(x'x - 1)$$

teriminin minimize edilmesi gerekmektedir. Bu amaçla, bu terimin  $x'$ e göre türevi alınarak 0' eşitlenmesiyle

$$(A - \lambda I)x = 0$$

denklemini elde edilir. Bu denklemin 0'dan farklı çözümleri için

$$|A - \lambda I| = 0$$

eşitliği çözülerek,  $n$  tane  $\lambda$  değeri belirlenir. Bu değerler,  $x'Ax$  ifadesini minimum yapan kökler olup  $A$  matrisinin *özdeğerleri* olarak adlandırılır: Her bir  $\lambda_i$  değeriyle

$$(A - \lambda_i I)x = 0$$

denklemini çözülerek bulunan  $x$ 'ler de  $A$  matrisinin *özvektörleridir*.

(2.22)'deki eşitliklerde  $V$  matrisinin sütunları,  $X'X$  matrisinin,  $\lambda_1, \lambda_2, \dots, \lambda_{(p-1)}$  özdeğerlerine karşı düşen normalize edilmiş ( $V'V = VV' = 1$  şartını sağlayan) özvektörleridir. Bu  $V$  matrisinden hareketle bağımsız değişkenlerin çeşitli doğrusal kombinasyonları oluşturulabilmektedir.  $Z$  ile gösterilen bu asal bileşenler,

$$Z = XV$$

ile ifade edilmektedirler. Çoklu doğrusal regresyon modeli, asal bileşenler ile

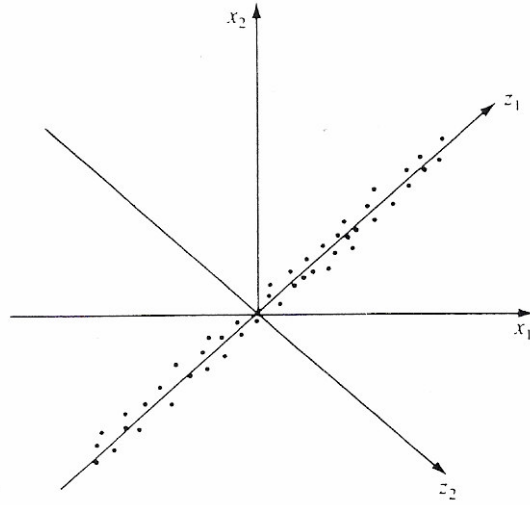
$$Y = Za + e$$

ifadesiyle yeniden tanımlanabilir. Burada,  $a = V'b$  matrisidir. Böylece bağımlı değişken ( $p-1$ ) adet  $x$  değişkeni ile açıklanırken, ( $p-1$ ) adet  $z$  değişkeni ile açıklanır hale gelir. Asal bileşenler varyanslarına göre sıraya dizildiklerinde,

$$\text{var}(z_1) > \text{var}(z_2) > \text{var}(z_3) > \dots > \text{var}(z_{p-1})$$

varyansı en büyük olan  $z_1$  ilk asal bileşen, ikinci büyük değişken ikinci asal bileşen vb. olarak adlandırılırlar. Bu asal bileşenler aşağıdaki özellikleri taşımaktadırlar:<sup>201</sup>

- $\text{var}(z_1) + \text{var}(z_2) + \dots + \text{var}(z_{p-1}) = \text{var}(x_1) + \text{var}(x_2) + \dots + \text{var}(x_{p-1})$
- $X$ 'lerden farklı olarak,  $Z$ 'ler ortogonal ve birbirlerinden bağımsızdırlar. Aşağıdaki şekilden de görüldüğü gibi  $X_1$  ve  $X_2$  arasında güçlü bir doğrusal ilişki mevcutken, dönüştürülmüş değişkenler  $Z_1$  ve  $Z_2$  arasında tesadüfi bir dağılım mevcuttur.



**Şekil (2.9): İki Bağımsız Değişkenli Modelde Asal Bileşenler**

Bu yöntemde elde edilen birinci asal bileşen özgün veri setindeki varyansın en büyük bölümünü açıklarken, ikinci asal bileşen ikinci büyük bölümü ve sonuncu

<sup>201</sup> Maddala, **a.g.e.**, s. 284.

asal bileşen varyansın en düşük kısmını açıklamaktadır. Bu yüzden, genel olarak ilk birkaç asal bileşen varyansın büyük çoğunluğunu oluşturur ve geri kalan bileşenlerin toplam varyans üzerindeki katkısı ihmal edilebilir büyüklükte olduğundan modelden çıkarılabilirler. Böylece, bir çok bağımsız değişkenden oluşan bir model, aynı bilgilerin büyük çoğunluğunu saklayacak daha az sayıdaki bağımsız değişkenle açıklanabilmektedir. Ancak ilk birkaç değişken toplam varyansın büyük bir çoğunluğunu oluşturmuyorsa bu yöntemi kullanmanın bir avantajı kalmayacaktır. Ayrıca bu yöntemin diğer dikkat edilmesi gereken noktaları şu şekilde sıralanmaktadır:<sup>202</sup>

- En yüksek varyansa sahip birinci asal bileşen  $Z_1$ , Y'ye göre en yüksek korelasyona sahip değişken olmayabilir. Bu durumda X'lere uygulanan Adım Adım Regresyon yöntemiyle modeli açıklayacak Z'ler seçilmelidir.
- X'lerin doğrusal kombinsyonu olan Z'ler ekonomik açıdan bir anlam ifade etmezler. Birimleri farklı olan X'lerin kombinasyonlarının biriminden söz edilemez. Bu yöntemde bu nedenle yorum zorluğu ortaya çıkmaktadır.
- X'lerin birimleri değiştirildiği zaman asal bileşenler de değişeceklerdir. Bu sorunu ortadan kaldırmak için, daha önce de bahsedildiği gibi değişkenlere korelasyon dönüşümleri uygulanarak standardize edilmektedirler.

Webster<sup>203</sup>, Asal Bileşenler yöntemini üniversiteleri sıralamak (en iyi başta olmak üzere) için kullanılacak sıralama kriterlerinin, bu işleme katkılarını ölçmek üzere uygulamıştır. Araştırmanın başında önerilen 11 kriterden, birkaçının gerçekte bu sıralama üzerinde etkili olduğu ve bunun nedeninin bu etkin kriterler ile diğer kriterler arasında doğrusal ilişkiler bulunması olarak gösterilmiştir.

---

<sup>202</sup> A.e., s. 285.

<sup>203</sup> Thomas J. Webster, "A Principal Component Analysis of the U.S. News & World Report Tier Rankings of Colleges and Universities", **Economics of Education Review**, Vol. XX, 2001, s. 235-244.

Öztürk ve Akdeniz<sup>204</sup> çalışmalarında, verilerdeki küçük deęişikliklere sonuçların hassas olması durumunu (kötü konumlanmış- ill-posed) ele almışlar ve çoklu doğrusal bağlantı durumunda tahminleyenlerin kararsız olma durumunu kötü konumlama problemi olarak incelemiş ve düzeltilmesi için önerilen matematik ve istatistik yöntemlerin sonuçlara etkilerini tartışmışlardır.

---

<sup>204</sup> Fikri Öztürk, Fikri Akdeniz, "Ill-Conditioning and Multicollinearity", **Linear Algebra and Its Applications**", Vol. CCCXXI, 2000, s. 295-305.



### **3. REGRESYON ANALİZİ İÇİN YENİ BİR BİLGİSAYAR PROGRAMI**

Bilgisayar teknolojilerindeki yazılım ve donanımdaki hızlı gelişmelere paralel olarak programlama dillerinin de işlem hacmi, fonksiyonellik ve veri işleme gibi özelliklerinde büyük gelişmeler sağlanmış ve bu da değişik amaçlar için hazırlanan ileri seviyeli paket programlarda çeşitliliğe neden olmuştur.

Günümüzde tahmin yöntemleri arasında en yaygın olarak kullanılan regresyon analizi, SPSS, SAS, STATISTICA, STATA, MINITAB gibi birçok istatistik paket programında yer almaktadır. Bu paket programların hemen hemen hepsinde, En Küçük Kareler yöntemiyle regresyon denkleminin parametreleri, standart hataları, korelasyon katsayıları ve ilgili testleri gibi regresyon analizinin temel aşamalarına benzer şekilde regresyon analizi başlığı altında yer verilmekte, ancak varsayımların saptanmasına ve çözümlenmesine yönelik yöntemlere farklı başlıklar altında karşılaşılmakta veya SPSS'te olduğu gibi Macro yazılımlarla ulaşılabilmektedir. Ancak yaptığımız araştırma sonucunda, regresyon ve korelasyon analizini model kurma adımlarından başlayarak, tüm varsayımların geçerliliğini araştırmak ve varsayımlardan sapmalar gözlemlendiğinde yaygın olarak kullanılan çözüm yöntemlerini de bir bütün olarak sunan programa rastlamak mümkün olmamıştır.

Çalışmanın amacı, sadece regresyon ve korelasyon analizine odaklanmak isteyen araştırmacıların kullanımına hizmet verecek ve yukarıda belirtilen özellikleri taşıyacak bir bilgisayar programı hazırlamaktır. Bu amaçla, (3.1.) alt bölümünde tanıtilen JAVA programlama dili kullanılarak oluşturulan programda dosyalama işlemleri, regresyon modeli oluşturma işlemleri, En Küçük Kareler yöntemiyle regresyon ve korelasyon analizi ile varsayımların saptanmasına yönelik birer test ile çözümleme yöntemleri sunulmaktadır. Bu programı diğerlerinden farklı kılan diğer bir özelliği de, istatistik programlarında genellikle kullanıcıya bırakılan sonuç

yorumlarının otomatik olarak yapılarak sunulmasıdır. Bu amaçla, sonuçları kesin kurallarla belirlenebilen test ve yöntemlere ait yorumlar verilmekte, yorumlanmasında kesin kurallar bulunmayan sonuçların değerlendirilmesi araştırmacıya bırakılmaktadır. Program, diğer kullanılan istatistik programları ile bütünlük oluşturmak amacıyla İngilizce dilinde hazırlanmıştır.

### 3.1. JAVA Programlama Dili

Java programlama dilinin oluşumu, ortam değişikliklerine uyum sağlamak ve programlama dillerindeki iyileştirmeleri gerçekleştirmek gibi iki temel nedene bağlıdır.<sup>205</sup> Java, yazım kuralı özellikleri açısından C programlama dilinden, nesne-yönelimli özellikleri açısından ise C++ dilinden etkilenmiştir. Nesnelere, ilgili değişken ve yöntemleri içeren yazılım grupları olarak tanımlanmaktadır.<sup>206</sup>

C dili 1970'lerde programcılar tarafından tasarlanıp geliştirilerek ortaya atıldığında, güçlülüğü, verimli ve yapısal özellikleriyle dikkat çekmiştir. 1980'lerin başında, programcıların kullanmayı en çok sevdikleri bir dil haline gelmiştir. Günümüzde de kullanılan bu dilin, en zayıf noktası karmaşıklığıdır. Bu karmaşıklığı düzenlemenin bir gereksinimi olarak ortaya atılan C++ dili programcılara geniş programları idare edebilme olanağı sağlamaktadır. 1990'ların başında C++ kullanarak nesne yönelimli programlar etkin hale gelmiştir. Ancak, birkaç yıl içinde World Wide Web ve Internet'in büyük kitlelere ulaşması programlama dilinde bir başka devrimin hazırlanmasına neden olmuştur.

Java 1991'de Sun Microsystems'da çalışan dört kişi tarafından düşünülüp geliştirilmiştir. İlk başta "Oak" olarak anılan bu dile 1995'de "Java" adı verildi. Bu dilin geliştirilmesi için ilk motivasyon elektronik tüketim eşyalarının içerisinde yerleştirilmesi düşünülen yazılımın platformdan bağımsız olması gereksinimiydi. Bu dil, değişik CPU(Central Processing Unit)'lar üzerinde çalışabilecek kodlar üretmek

<sup>205</sup> Herbert Schildt, **Java 2**, Çev. Taner Akbaş, İstanbul, Alfa, 2001, s. 4.

<sup>206</sup> "The Java Tutorial, Object-Oriented Programming Concepts", (Çevrimiçi) <http://java.sun.com/docs/books/tutorial/concepts/index.html>, 6 Haziran 2005.

için kullanılabilir, taşınabilir, başka bir ifade ile platformdan bağımsız bir dil olmalıydı. Java'nın yaratılmasına neden olan bu gereksinime paralel olarak gelişen ve bu dilin geleceğinde önemli bir rol oynayan ikinci bir faktör de Web idi. Çünkü Web de taşınabilir programlara gereksinim duyuyordu. Böylece Java, C++ tarafından kullanılan nesne yönelimli paradigmayı genişleterek, Internet'e bağlı değişik platformlara dinamik olarak indirilecek programların taşınabilir ve yürütülebilir kodlar üretilmesini sağlayan en güçlü araç haline geldi.

Java'nın bu güçlü özelliklerinin yanısıra diğer önemli özellikleri de dilin son haline sokulmasında etkili olmuştur.<sup>207</sup> Bunlar basit, güvenli, sağlam, çok kanallı ve dinamik oluşu, yorumlama kolaylığı, yüksek performansı gibi özellikleridir.

En yaygın olarak kullanılan Java programları uygulamalar ve applet'lerdir. Uygulamalar yükledikleri bilgisayarın işletim sistemiyle çalışan bağımsız programlardır. Appletler de uygulamalara benzerdirler, ancak onlar gibi bağımsız olarak çalıştırılmazlar. Appletler Internet üzerinden nakledilebilecek ve Java uyumlu bir Web tarayıcıda çalışabilecek, kullanıcı girdisine göre tepki veren dinamik olarak değişen uygulamalardır.<sup>208</sup>

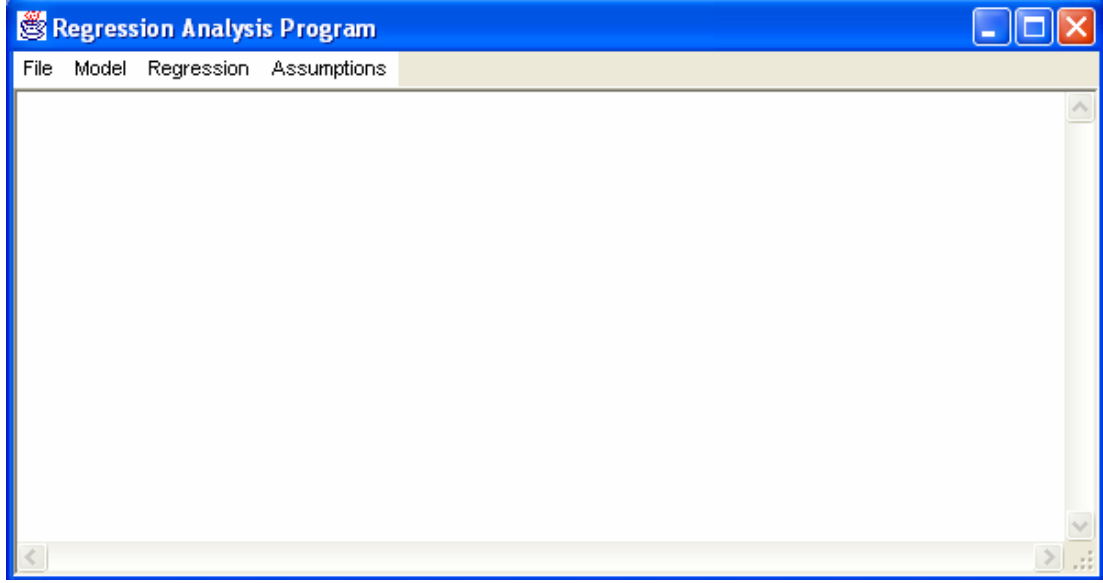
### 3.2. Regresyon Analizi Programı

Daha önce de bahsedildiği gibi, Java programlama diliyle yazılan bu programda dört ana başlık bulunmaktadır. "Regression Analysis Program" adı verilen programın ana başlıkları, "File", "Model", "Regression" ve "Assumptions" dır.

---

<sup>207</sup> Schildt, **a.g.e.**, s. 10.

<sup>208</sup> "The Java Tutorial, Getting Started", (Çevrimiçi) <http://java.sun.com/docs/books/tutorial/getStarted/index.html>, 6 Haziran 2005.

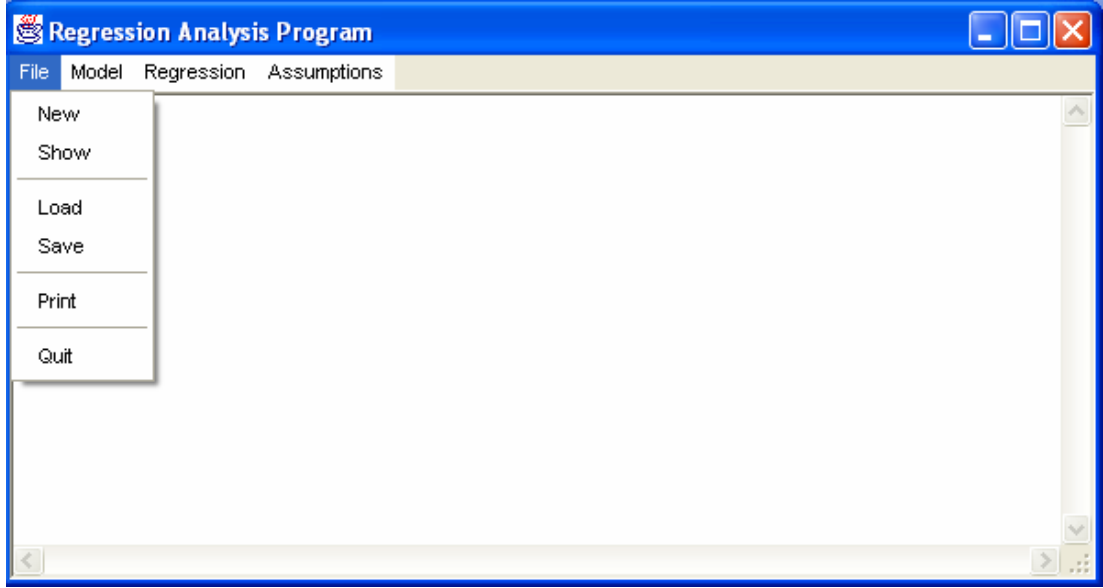


**Şekil (3.10): Regresyon Analizi Programı'nın Ana Ekranı**

Bu bölümde, tüm başlıklar alt başlıklarıyla birlikte ayrıntılarıyla açıklanmakta ve girilmiş olan örnek verilerle analiz sonuçları gösterilmektedir. Programın ürettiği sonuçların doğruluğunu test etmek amacıyla, aynı uygulamalar SPSS ile de gerçekleştirilmiş ve sonuçları (Ek: 6-18)'de sunulmuştur.

### **3.2.1. “File” Menüsü**

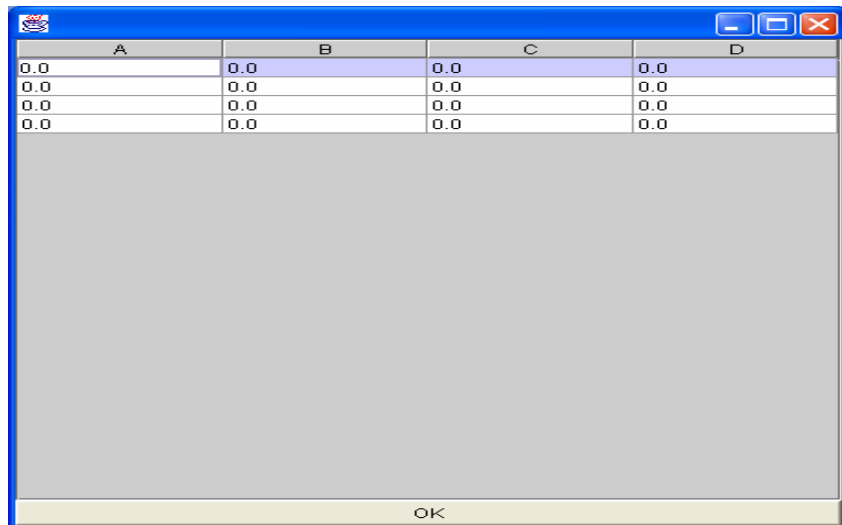
Diğer Windows tabanlı programlarda olduğu gibi çeşitli dosyalama işlemleri, “File” menüsü altındaki alt başlıklarda yapılmaktadır. Bu işlemler yeni bir dosya yaratmak, mevcut bir dosyayı açmak, bir dosyayı basmak ve programdan çıkmak gibi seçeneklerdir.



Şekil (3.11): “File” Menüsü

“File” seçeneğine basıldığında açılan listedeki seçenekler ve işlevleri şu şekildedir:

- “New” seçeneğine basıldığında yeni veri girişi için gerekli satır ve sütun sayıları<sup>209</sup> program tarafından sorulmakta ve girilen değerlere uygun bir veri giriş sayfası sunulmaktadır. Veri girişi yapıldıktan sonra “OK” tuşu ile onay verilmektedir.



Şekil (3.12): “New”Seçeneği

<sup>209</sup> Satır sayısı için gözlem, sütun sayısı için ise değişken sayısı girilmelidir.

- “Show” seçeneği “New” ile yaratılan veya “Load” ile yüklenen bir veri dosyasını matris formatında göstermektedir. Aşağıdaki şekilde gözüken veriler (Ek:4)’deki örneğe “Show” seçeneğinin uygulanmasıyla elde edilmiştir.

A	B	C	D	E
2.301	6.7	62.0	81.0	2.59
2.0043	5.1	59.0	66.0	1.7
2.3096	7.4	57.0	83.0	2.16
2.0043	6.5	73.0	41.0	2.01
2.7067	7.8	65.0	115.0	4.3
1.9031	5.8	38.0	72.0	1.42
1.9031	5.7	46.0	63.0	1.91
2.1038	3.7	68.0	81.0	2.57
2.3054	6.0	67.0	93.0	2.5
2.3075	3.7	76.0	94.0	2.4
2.5172	6.3	84.0	83.0	4.13
1.8129	6.7	51.0	43.0	1.86
2.9191	5.8	96.0	114.0	3.95
2.5185	5.8	83.0	88.0	3.95
2.2253	7.7	62.0	67.0	3.4
2.3365	7.4	74.0	68.0	2.4
1.9395	6.0	85.0	28.0	2.98
1.5315	3.7	51.0	41.0	1.55
2.3324	7.3	68.0	74.0	3.56
2.2355	5.6	57.0	87.0	3.02
2.0374	5.2	52.0	76.0	2.85
2.1335	3.4	83.0	53.0	1.12
1.8451	6.7	26.0	68.0	2.1
2.2424	5.8	67.0	86.0	2.4

**Şekil (3.13): “Show” Seçeneği**

- “Load” seçeneği daha önce yaratılıp saklanan “txt” uzantılı veri dosyalarının programda kullanmak için yüklenmesini sağlamaktadır. Ancak bu dosyaların ilk satırında gözlem sayısı, ikinci satırında ise değişken sayısı (bağımlı değişken dahil) yazılmalı, ardından veriler matris formatında girilmelidir. Bu matrisin ilk sütunu her zaman bağımlı değişken, diğer sütunları ise bağımsız değişken değerlerini içermelidir. (Ek:1)
- “Save”, “New” komutu ile girilen veya “Load” ile programa yüklenen verilerin “txt” uzantılı bir dosyada saklanması için kullanılmaktadır.

- “Print” seçeneği sonuçların yazıldığı ekranın yazıcıdan basılmasını sağlamaktadır.
- “Quit” programı sonlandırmak için kullanılmaktadır.

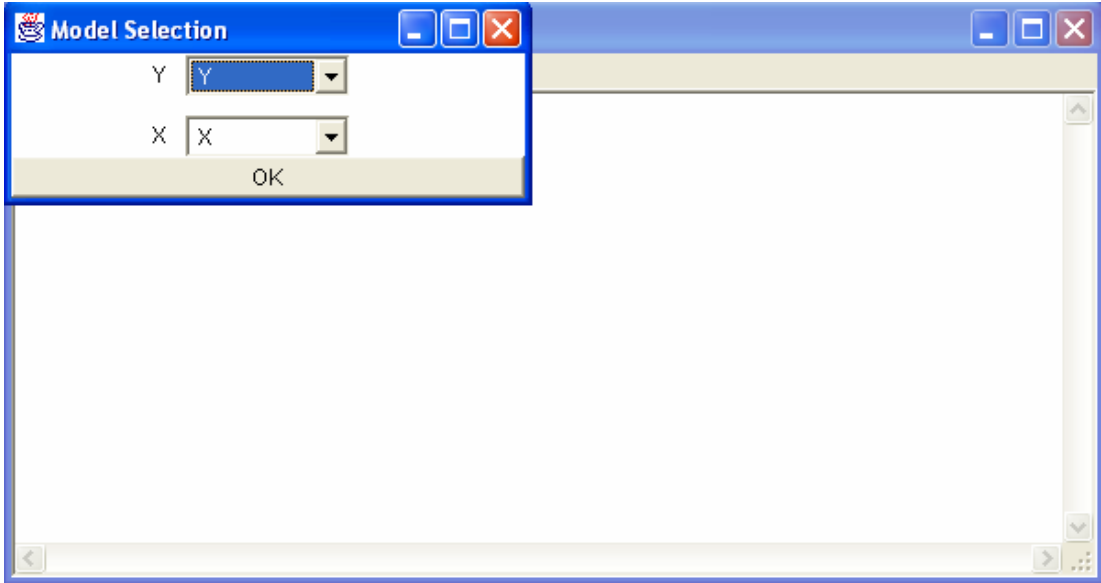
### 3.2.2. “Model” Menüsü

Bu menü ile girilen değişkenler doğrudan kullanılabilceği gibi, değişkenlere çeşitli dönüşümler de uygulanabilmektedir. Bu menü seçildiğinde gözüken “Selection” düğmesine basıldığında açılan pencerede, bağımlı değişken dahil tüm değişkenler listelenmektedir. Her bir değişkenin yanında gözüken ok işareti bu değişkenlere ait çeşitli seçenekler bulunduğunu göstermektedir. Bilindiği gibi, regresyon analizinde değişkenler arasındaki ilişki her zaman doğrusal olmayabilmektedir. Bu durumda ilişki doğrusal olmayan bir denklemle ifade edilmekte veya değişkenlerin bazılarında bazen de hepsine uygulanan dönüşümlerle model doğrusal forma dönüştürülerek uygunluğu araştırılmaktadır. Yaygın olarak uygulanan dönüşüm fonksiyonları doğal logaritma, karekök ve hiperbolik fonksiyonlardır. Değişkenlere basıldığında açılan listeler ile kullanıcı bu dönüşümlerden uygun olanı kolayca seçebilmektedir. Bu seçenekler modelin basit veya çoklu olmasına bağlı olarak değişmektedir.

Basit doğrusal regresyon analizi dikkate alındığında, doğal olarak menüde iki değişken listelenecektir. Bağımlı değişken seçildiğinde, açılan seçenekler listesinde bağımlı değişkenin kendisi ( $y$ ), doğal logaritması ( $\ln(y)$ ), karekökü ( $\sqrt{y}$ ) ve tersi ( $1/y$ ) bulunmaktadır. Otomatik olarak seçili gözüken seçenek bağımlı değişkenin doğrudan kendisi ( $y$ ) olmakta, ayrıca diğer seçeneklerden herhangi biri tıklanarak dönüşüm uygulanabilmektedir. İkinci dereceden eğrisel regresyon modeli genellikle basit regresyon modellerinde söz konusu olduğundan, bağımsız değişken için sunulan seçenekler çoklu modellerinkinden farklı olacaktır. Bağımsız değişken seçildiğinde listede kendisi ( $x$ ), eğrisel model için sunulan ( $x$  ve  $x^2$ ) seçeneği ile diğer dönüşüm seçenekleri ( $\ln(x)$ ,  $\sqrt{x}$  ve  $1/x$ ) bulunmaktadır.

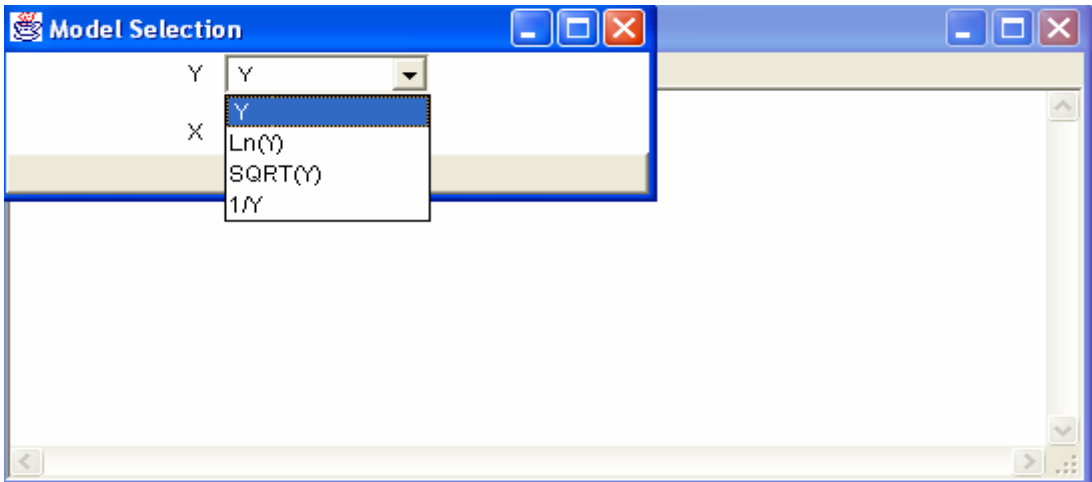
Sonuç olarak arařtırmacı, doğrusal, eğrisel veya uygulanan en yaygın dönüşümlerle doğrusal forma kavuşturulmuş modelleri kolaylıkla oluşturabilmekte ve diğer menüler aracılığıyla hepsi için, ayrı ayrı regresyon analizini uygulayabilmektedir.

“Load” seçeneđi ile programa yüklenen (Ek:2)’deki tek bağımsız deđişkenli veriler için “Selection” seçeneđine basıldığında ařađıdaki pencere açılacaktır:



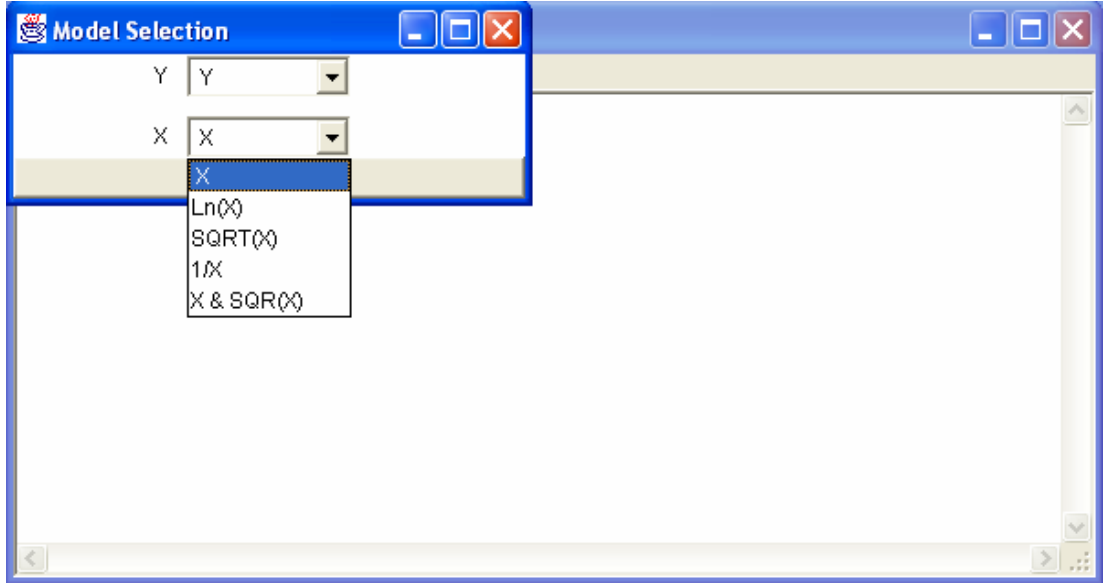
**Şekil (3.14): Tek Bağımsız Deđişkenli Model için “Selection” Seçeneđi**

Y ve X için seçenekler listesinin görünümü şu şekilde olacaktır:



**Şekil (3.15): Tek Bağımsız Deđişkenli Modellerde “Y” için Sunulan Seçenekler**

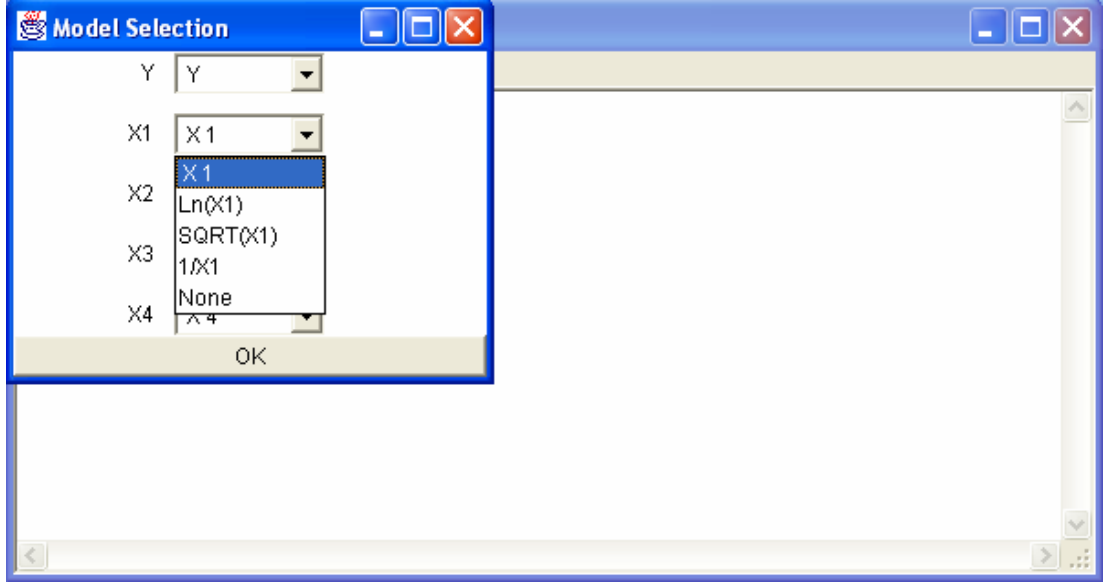




**Şekil (3.16): Tek Bağımsız Değişkenli Modellerde “X” için Sunulan Seçenekler**

Çoklu regresyon modelleri söz konusu olduğunda ise, listede “File” menüsündeki “New” veya “Load” seçenekleri ile programa tanıtılmış olan tüm değişkenler otomatik olarak listelenecektir. Bağımlı değişken için, basit modeldeki bağımlı değişken için tanımlanan seçeneklerin aynıları geçerli olmakta, bağımsız değişkenler için ise beş değişik seçenek bulunmaktadır. Bunlardan biri “none” seçeneği olup bu seçenek ile bağımsız değişkenin model dışı tutulması sağlanabilmektedir. Diğerleri, basit regresyon modelinde de bulunan  $x$ ,  $\ln(x)$ ,  $\sqrt{x}$  ve  $1/x$  seçenekleridir. Böylece çoklu doğrusal regresyon modelleri, girilen tüm değişkenlerle veya bazı değişkenler model dışı tutularak oluşturulabilmektedir. Ayrıca değişkenlerine çeşitli dönüşümler uygulanarak doğrusal forma dönüştürülmüş modeller de bu seçenekler aracılığıyla kolaylıkla söz konusu olabilmektedir.

(Ek:4)’de verilen çoklu regresyon verileri programa yüklendiğinde, “Selection” seçeneği ile modele dahil edilecek değişken seti aşağıda gözükün pencere ile belirlenmektedir.



**Şekil (3.17): Çoklu Regresyon Modellerinde Model Belirleme**

Çoklu regresyon modellerinde de, kolaylıkla oluşturulan modellere diğer menülerde sunulan çeşitli seçenekler aracılığıyla regresyon analizi uygulanabilmekte ve varsayımların varlığı araştırılmaktadır.

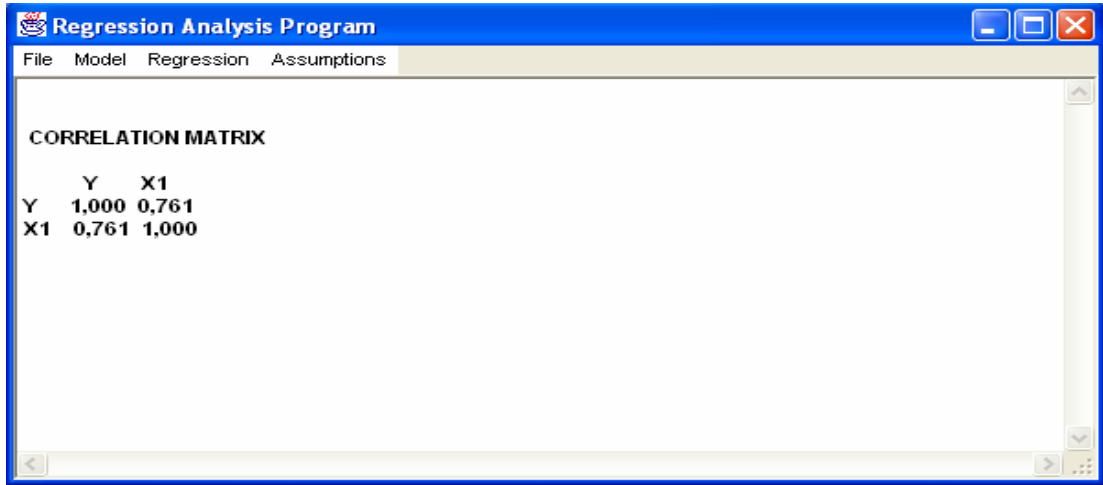
Basit ve çoklu modellerde ulaşılan sonuçları karşılaştırarak veriye uygun modele karar vermek araştırmacıya bırakılmıştır. Bu amaçla, programın herhangi bir noktasında “Model” menüsüne geri dönüp seçilen değişken setinde yukarıda bahsedilen değişiklikleri yapıp, analizi baştan uygulamak olası hale getirilmiştir.

### **3.2.3. “Regression” Menüsü**

Regression menüsü altında “Correlations” ve “Regression Equation” alt başlıkları bulunmaktadır. Her iki seçeneğin de sonuç oluşturabilmeleri için “Model” menüsünden bir model belirlenmesi ve “OK” düğmesiyle onaylanması gerekmektedir.

“Correlations” seçeneği ile modeldeki tüm değişkenlerin aralarındaki basit korelasyon katsayıları gösterilmektedir. Correlations’ın ilk sunulan seçenek olmasının nedeni, analize başlamadan önce araştırmacıların genellikle seçtikleri bağımsız değişkenlerin amaç değişkenleri olan bağımlı değişken ile aralarındaki ve ayrıca da birbirleriyle olan basit korelasyon katsayılarını incelemeleri gerekmesidir. Bilindiği gibi regresyon analizinde, bağımlı değişken ile seçilen bağımsız değişkenler arasında belirli bir ilişki olması, bağımsızlar arasında ise ilişki olmaması gerekmektedir. Sadece başlangıçta değil, program kullanılırken herhangi bir anda bu seçenek seçildiğinde korelasyon matrisi çıktı raporuna eklenmektedir.

(Ek:2)’deki basit regresyon modeline basit doğrusal korelasyon katsayılarını elde etmek için, “Selection” seçeneğinde değişkenleri olağan hallerinde seçerek “OK” düğmesine, ardından “Correlations” seçeneğine tıklamak gerekmektedir.



The screenshot shows a window titled "Regression Analysis Program" with a menu bar containing "File", "Model", "Regression", and "Assumptions". The main area displays the following correlation matrix:

	Y	X1
Y	1,000	0,761
X1	0,761	1,000

**Şekil (3.18): Basit Regresyon Modellerinde “Correlations” Seçeneği**

Benzer şekilde, (Ek:4)’deki çoklu regresyon modelinde, tüm değişkenleri hiçbir dönüşüm uygulamadan doğrudan seçtiklerinde “Correlations” seçeneği aşağıdaki sonuçları vermektedir:

The screenshot shows a window titled "Regression Analysis Program" with a menu bar containing "File", "Model", "Regression", and "Assumptions". The main area displays a "CORRELATION MATRIX" with the following data:

	Y	X1	X2	X3	X4
Y	1,000	0,346	0,593	0,665	0,726
X1	0,346	1,000	0,090	-0,150	0,502
X2	0,593	0,090	1,000	-0,024	0,369
X3	0,665	-0,150	-0,024	1,000	0,416
X4	0,726	0,502	0,369	0,416	1,000

Şekil (3.19): Çoklu Regresyon Modellerinde “Correlations” Seçeneği

“Regression Equation” seçeneği, regresyon analizinin değişkenlere uygulanıp sonuçların raporda gösterilmesini sağlamaktadır. Burada, öncelikle “Model” menüsüyle seçilen modelin katsayıları ( $b_0, b_1, \dots, b_{p-1}$ ) hesaplanarak regresyon denklemi elde edilmektedir. Daha sonra, çoklu korelasyon katsayısı ( $R$ ), belirlilik katsayısı ( $R^2$ ) ile modelin standart hatası ( $S$ ) ve modelin varyans analizi sonuçları ( $F$  testi) verilmekte, raporun sonunda ise regresyon katsayılarının standart hataları ve  $t$  testleri sonuçlarına yer verilmektedir. Bu adımda diğer yazılımlardan farklı olan kısım, test sonuçlarının kullanıcı için %1 ve %5 anlamlılık düzeyleri için yorumlanmasıdır.  $t$  test değerleri her iki düzey için  $t$  tablosundan (Ek:23) uygun serbestlik dereceleri için bulunan kritik değerler ile karşılaştırılmakta ve parametrelerin anlamlı olup olmadıkları belirtilmektedir. Benzer şekilde, çoklu korelasyon katsayısının  $F$  testinde, test değeri aynı anlamlılık düzeyleri için  $F$  tablosu (Ek:24-25) değerleriyle karşılaştırılarak yorumlanmaktadır.

Şekil (3.9)’da korelasyon katsayıları hesaplanan basit doğrusal regresyon modeli için “Regression Equation” seçeneği aşağıdaki sonuçları verecektir.

**MODEL SUMMARY**

$Y=7,3374 + 57,5174 X1$

R    R2    S  
0,7605 0,5784 9,4366

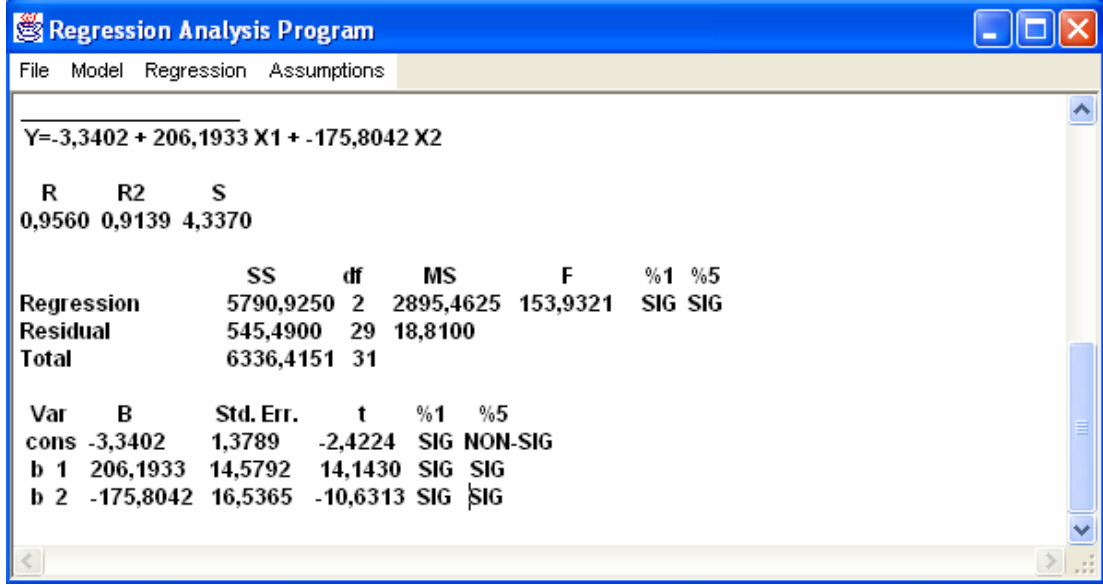
	SS	df	MS	F	%1	%5
Regression	3664,9509	1	3664,9509	41,1567	SIG	SIG
Residual	2671,4642	30	89,0488			
Total	6336,4151	31				

Var	B	Std. Err.	t	%1	%5
cons	7,3374	2,0556	3,5694	SIG	SIG
b 1	57,5174	8,9656	6,4153	SIG	SIG

**Şekil (3.20): (Ek:2)'deki Verilere Uygulanan Basit Doğrusal Regresyon Analizi Sonuçları**

Bu örnekte, hesaplanan korelasyon katsayısı bağımsız değişkenin, bağımlı değişkendeki değişkenliği %76 oranında açıklayabildiğini göstermektedir. Doğrusal regresyon modelinin F testi ile regresyon katsayısının t testi sonuçları, yaygın olarak kullanılan %1 düzeyinde anlamlı bulunmuştur. Daha yüksek bir orana ulaşmak için, modelin doğrusal olmayan ikinci derece bir fonksiyonla ifade edilip edilemeyeceği veya dönüşümler uygulanarak doğrusal forma dönüştürülüp dönüştürülemeyeceği araştırılabilmektedir.

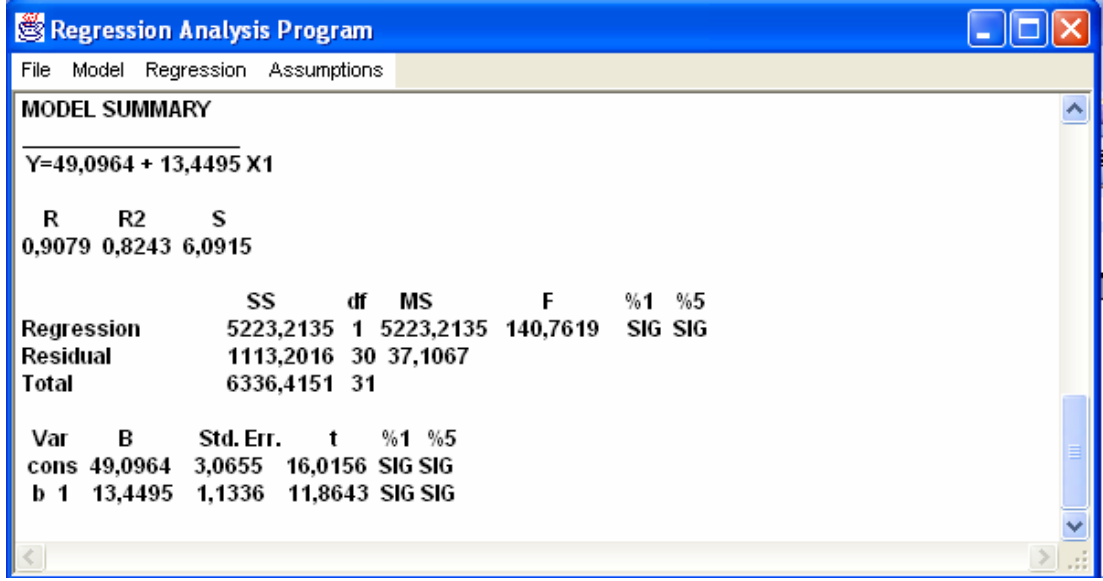
Bu amaçla, yapılması gerekenler “Model” menüsündeki “Selection” seçeneğine geri dönerek, x için “X & sqrt(X)” seçeneğine basmak ve “Regression Equation” ile regresyon analizi sonuçlarına ulaşmaktır.



**Şekil (3.21): (Ek:2)'deki Verilere Uygulanan Eğrisel Regresyon Analizi Sonuçları**

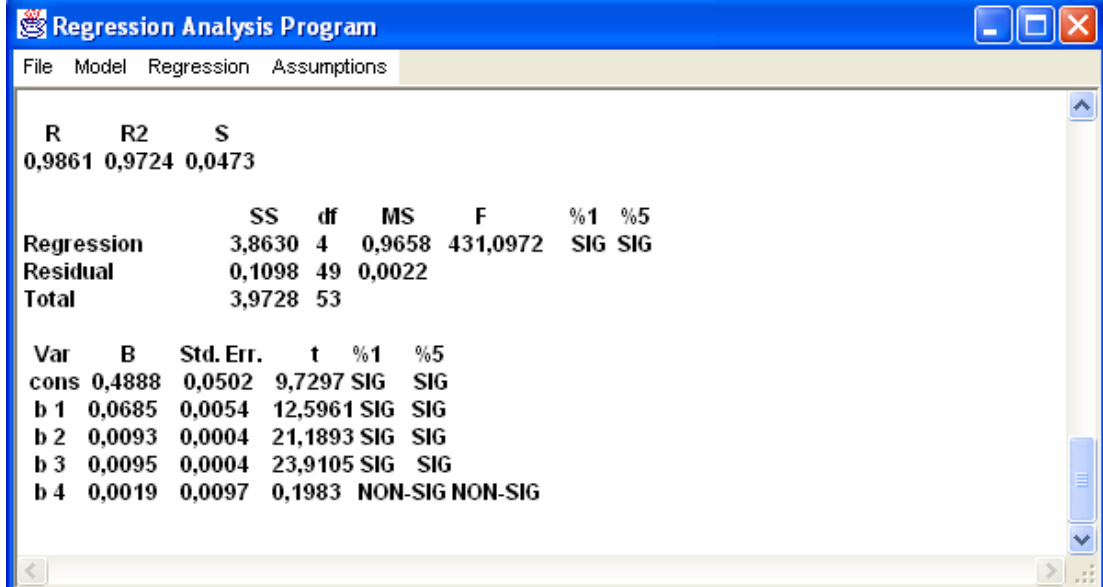
Bu sonuçlara göre, elde edilen yeni modelde de iki regresyon katsayısı ve korelasyon katsayısının test sonuçları anlamlı bulunmuş ve daha yüksek bir korelasyon katsayısına ulaşılmıştır. Demek ki, bağımlı değişken ile bağımsız değişken arasında ikinci dereceden bir ilişki söz konusudur.

Bağımsız değişkene uygulanabilecek tüm dönüşümler denendiğinde ise, elde edilen modeller arasında en yüksek korelasyon katsayısına sahip model  $\ln(x)$  dönüşümü ile şu şekilde elde edilmektedir.



**Şekil (3.22): (Ek:2)'deki Verilere Uygulanan  $\ln(x)$  Dönüşümü ile Ulaşılan Regresyon Analizi Sonuçları**

(Ek:4)'deki çoklu regresyon örneği için "Regression Equation" ile elde edilen sonuçlar ise aşağıda gösterilmektedir.



**Şekil (3.23): Çoklu Regresyon Modeli için Doğrusal Regresyon Analizi**

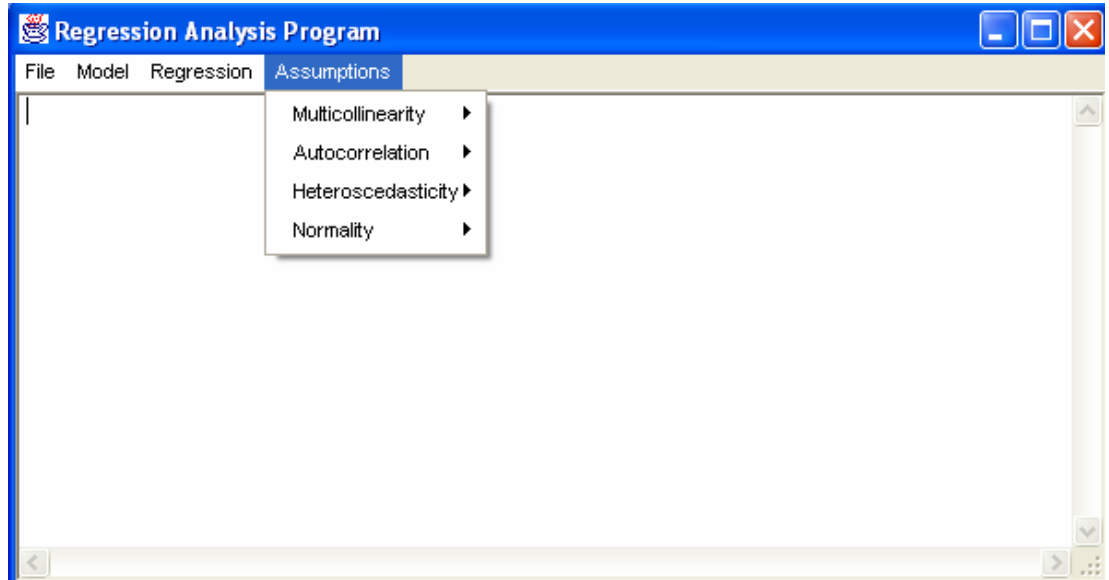
Bu örnekte, bağımsız değişkenlerin anakütlede bağımlı değişkenin değişkenliğini 0,986 oranında açıklayabildikleri saptanmakta, modelin F testi ile  $b_1$ ,  $b_2$  ve  $b_3$  katsayılarının t testleri anlamlı, ancak  $b_4$  katsayısının t test sonucu her iki

düzey için de anlamsız bulunmaktadır. Bu durumda, başta çoklu doğrusal bağlantı olmak üzere diğer varsayımların varlığı da araştırılarak doğru modele karar verilmesi gerekmektedir.

### 3.2.4. “Assumptions” Menüsü

Çalışmanın teorik bölümünde ayrıntılı olarak açıklandığı üzere seçilen modelin kullanılabilmesi için dört varsayımın geçerliliğinin araştırılması gerekmektedir. Geliştirilmeye çalışılan bu programda odaklanılan amaç, regresyon analizinin yanında bu varsayımların geçerliliklerini de test etmek ve geçerli olmamaları durumunda seçilecek uygun yöntemler ile yeni regresyon modelleri elde etmektir.

“Assumptions” menüsü araştırmacılara bu olanağı sağlamakta ve her varsayımla ilgili işlemlerin sunulduğu “Multicollinearity”, “Autocorrelation”, “Heteroscedasticity” ve “Normality” alt başlıklarını içermektedir.

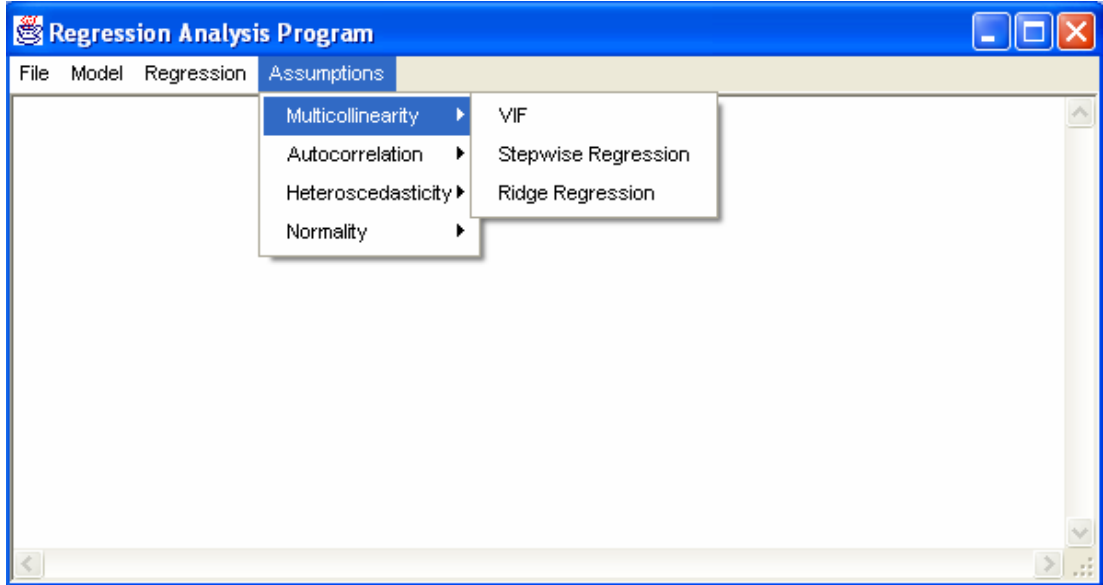


Şekil (3.24): “Assumptions” Menüsü



### 3.2.4.1. “Multicollinearity” Alt Başlığı

“Multicollinearity” alt başlığı seçildiğinde yanda açılan menüde üç seçenek bulunmaktadır.



Şekil (3.25): “Multicollinearity” Alt Başlığı

Bunlardan ilki çoklu doğrusal bağlantıyı saptamakta sıkça kullanılan ölçütlerden biri olan VIF (Variance Inflation Factors) değerlerinin hesaplanmasını sağlayan seçenektir. Böylece, seçilen model için tüm bağımsız değişkenlere ait VIF değerleri rapora eklenmektedir. Regresyon analizi sonuçlarıyla birlikte bu değerler de dikkate alınarak, hangi bağımsız değişkenlerin aralarında çoklu doğrusal bağlantının önemli sayılacak bir düzeyde olduğu ve dolayısıyla hangi değişkenlerin model dışında tutulması gerektiği araştırılabilmekte veya En Küçük Kareler yöntemi dışında bir yöntemin uygulanması gerektiği anlaşılabilir.

(Ek:4)’deki çoklu model için “Regression Equation” ile regresyon analizi sonuçları elde edildikten sonra modelin, çoklu doğrusal bağlantı içerip içermediğini araştırmak için VIF değerleri aşağıdaki gibi elde edilmektedir.

The screenshot shows the 'Regression Analysis Program' window. The 'Assumptions' tab is active, displaying the following data:

Var	B	Std. Err.	t	%1	%5
cons	0,4888	0,0502	9,7297	SIG	SIG
b 1	0,0685	0,0054	12,5961	SIG	SIG
b 2	0,0093	0,0004	21,1893	SIG	SIG
b 3	0,0095	0,0004	23,9105	SIG	SIG
b 4	0,0019	0,0097	0,1983	NON-SIG	NON-SIG

VIF	
X1	1,799
X2	1,289
X3	1,678
X4	2,555

**Şekil (3.26): Çoklu Regresyon Modelinde VIF Değerleri**

Programın bu bölümünde teorik bölümde de bahsedildiği gibi çoklu doğrusal bağlantının bulunması durumunda, bağımlı değişken üzerinde anlamlı etkisi olan ancak diğer bağımsız değişkenlerle arasında ilişki olmayan değişkenlerin seçiminde en yaygın olarak kullanılan ve hemen hemen tüm paket programlarda mevcut olan “Stepwise Regression” (Adım Adım Regresyon) yöntemi uygulanabilmektedir. İkinci alt başlıkta yer alan bu seçenek, bölüm (2.4.4.1.)’de bahsedilen algoritmaya<sup>210</sup> göre saptanan değişkenlerle yeni modeli oluşturmaktadır. Sonuç olarak, “Regression Equation” menüsünde olduğu gibi burada da modelin korelasyon ve belirlilik katsayıları, standart hatası, varyans analizi tablosu ve regresyon katsayıları ile t testi sonuçları elde edilmektedir.

Yukarıda ele alınan çoklu regresyon modelinde, elde edilen VIF değerleri 10’dan büyük olmamasına karşın “Correlations” seçeneği ile elde sonuçlara göre bazı bağımsız değişkenlerin aralarındaki basit korelasyon katsayıları oldukça yüksektir. Bu durumda, bağımlı değişkeni anakütlede açıklayabilecek en iyi değişken setine ulaşabilmek için “Stepwise Regression” seçeneğine basıldığında aşağıdaki sonuçlar elde edilmektedir.

<sup>210</sup> Algoritmadaki  $F_{IN}$  değeri 4,  $F_{OUT}$  değeri de 3.99 olarak seçilmiştir.

**MODEL SUMMARY**

$Y=0,4836 + 0,0692 X1 + 0,0093 X2 + 0,0095 X3$

R	R2	S
0,9861	0,9723	0,0469

	SS	df	MS	F	%1	%5
Regression	3,8629	3	1,2876	586,0431	SIG	SIG
Residual	0,1099	50	0,0022			
Total	3,9728	53				

Var	B	Std. Err.	t	%1	%5
cons	0,4836	0,0426	11,3450	SIG	SIG
b 1	0,0692	0,0041	16,9755	SIG	SIG
b 2	0,0093	0,0004	24,2992	SIG	SIG
b 3	0,0095	0,0003	31,0817	SIG	SIG

**Şekil (3.27) Çoklu Regresyon Modelinde “Stepwise Regression” Yöntemi**

Bu menünün üçüncü ögesi ise, çoklu doğrusal bağlantı durumunda değişken eklemek yerine modeldeki tüm değişkenlerin de modele alındığı En Küçük Kareler yöntemine alternatif olarak ortaya atılmış olan, tahminleyenleri taraflı fakat gerçek anakütle değerlerine daha yakın olarak tahmin eden “Ridge Regression” yöntemidir.

Bu seçenekte, program yeni bir pencere açmakta ve yöntemde uygun tahminleyenlerin seçilmesine olanak sağlayan c katsayısının değerini 0’dan 0.99’a kadar her adımda 0.01 aralıklarla arttırarak, bu değerlere karşı gelen modelin regresyon katsayılarının tahminleyenlerini hesaplayarak yazmaktadır. Araştırmacının bu değerleri inceleyerek tahminleyenlerin durağanlaşmaya başladığı uygun c değerini belirlemesi ve pencerenin alt satırına “Enter c” başlığına karşı gelen kutuya girmesi gerekmektedir. “Find Model” seçeneğine basıldığında ise, seçilen tahminleyenlere bölüm (2.4.4.2.)’de bahsedilen (2.21)’deki dönüşümler uygulanmakta ve yeni regresyon modeli oluşturularak doğrudan rapora yazılmaktadır.

(Ek:4)’deki veriler için bu yöntem uygulandığında, c katsayısı ile b parametrelerin değerleri aşağıdaki şekilde elde edilmektedir.

c	B1	B2	B3	B4
0,00	0,4012	0,5713	0,7355	0,0075
0,01	0,3863	0,5597	0,7180	0,0263
0,02	0,3728	0,5489	0,7018	0,0432
0,03	0,3604	0,5389	0,6868	0,0585
0,04	0,3490	0,5294	0,6728	0,0724
0,05	0,3385	0,5205	0,6596	0,0851
0,06	0,3288	0,5121	0,6473	0,0966
0,07	0,3198	0,5042	0,6357	0,1072
0,08	0,3115	0,4967	0,6248	0,1169
0,09	0,3037	0,4896	0,6144	0,1258
0,10	0,2964	0,4827	0,6046	0,1340
0,11	0,2896	0,4762	0,5953	0,1415
0,12	0,2832	0,4700	0,5864	0,1485
0,13	0,2772	0,4641	0,5779	0,1549
0,14	0,2715	0,4583	0,5698	0,1608
0,15	0,2662	0,4529	0,5621	0,1663
0,16	0,2611	0,4476	0,5546	0,1714
0,17	0,2564	0,4425	0,5475	0,1762
0,18	0,2518	0,4376	0,5406	0,1806
0,19	0,2475	0,4328	0,5340	0,1847
0,20	0,2434	0,4282	0,5276	0,1885
0,21	0,2395	0,4238	0,5215	0,1920
0,22	0,2358	0,4195	0,5155	0,1953
0,23	0,2322	0,4153	0,5098	0,1984
0,24	0,2288	0,4112	0,5042	0,2012
0,25	0,2255	0,4073	0,4989	0,2039
0,26	0,2224	0,4034	0,4936	0,2064
0,27	0,2194	0,3997	0,4886	0,2087
0,28	0,2165	0,3961	0,4837	0,2108

Enter c: 0.40 Find Model Close

**Şekil (3.28): Ridge Regression Yönteminde “c” Katsayısının Belirlenmesi**

c için 0.40 değeri girilip “Find Model” düğmesine basıldığında (2.21)’deki geri dönüşümler uygulanarak elde edilen model otomatik olarak rapora eklenmektedir.

	SS	df	MS	F	%1	%5
Regression	3,8630	4,0000	0,9658	431,0972	SIG	SIG
Residual	0,1098	49,0000	0,0022			
Total	3,9728	53,0000				

Var	B	Std. Err.	t	%1	%5
cons	0,4888	0,0502	9,7297	SIG	SIG
b 1	0,0685	0,0054	12,5961	SIG	SIG
b 2	0,0093	0,0004	21,1893	SIG	SIG
b 3	0,0095	0,0004	23,9105	SIG	SIG
b 4	0,0019	0,0097	0,1983	NON-SIG	NON-SIG

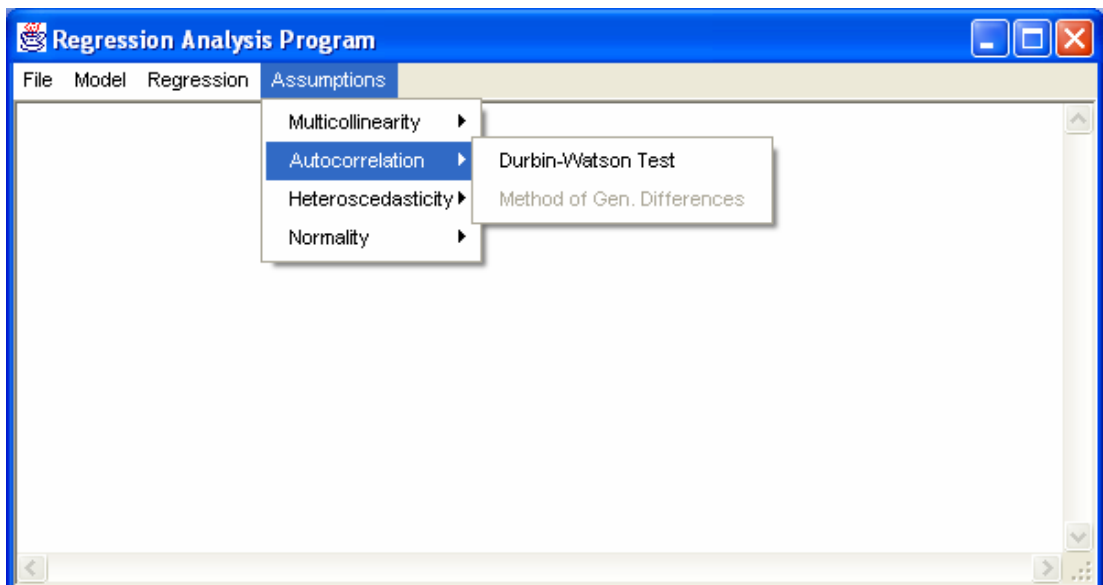
**MODEL SUMMARY**

Y=1,0615 + 0,0323 X1 + 0,0058 X2 + 0,0056 X3 + 0,0581 X4

**Şekil (3.29): Ridge Regression Yöntemiyle Modelin Belirlenmesi**

### 3.2.4.2. “Autocorrelation” Alt Başlığı

Bu başlık seçildiğinde, yanda açılan listede iki seçenek bulunmaktadır. Bunlardan ilki otokorelasyonun saptanmasında kullanılan en yaygın test olan Durbin-Watson testi, ikincisi ise otokorelasyon bulunması durumunda önerilen yöntemlerden biri olan Genelleştirilmiş Farklar yöntemidir.(Bakınız (2.1.5.1.1.) alt bölümü)



**Şekil (3.30): “Autocorrelation” Alt Başlığı**

Durbin-Watson testini uygulamadan önce seçilen modele regresyon analizi “Regression Equation” alt başlığı ile veya “Stepwise Regression” yöntemiyle uygulanarak öncelikle sonuçlar incelenmelidir. Program, en son sonuçları gösterilen modele testi uygulamaktadır. Diğer paket programlarda olduğu gibi modele ait Durbin-Watson test istatistik değeri hesaplanarak rapora eklenmektedir. Programda ek olarak bu değer, %1 ve %5 anlamlılık düzeyleri için tablo değerleriyle (Ek:26a ve 26b) karşılaştırarak bölüm (2.1.4.3.)’de bahsedildiği gibi yorumlanmaktadır.

Sonuçlar %1 ve %5 için ayrı ayrı şu şekilde yazılmaktadır:

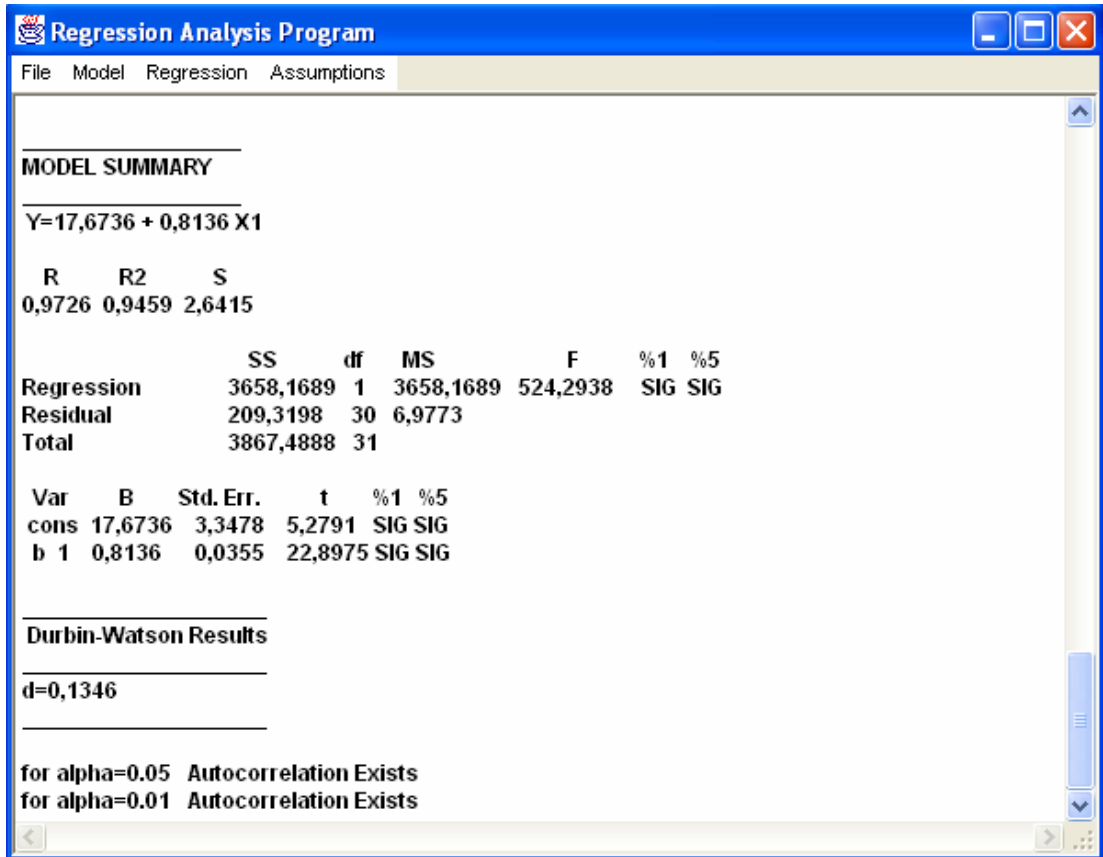
$d < d_L$ veya $d > 4-d_L$	ise “Autocorrelation Exists”
$d_U < d < 4-d_U$	ise “Autocorrelation Does Not Exist”
$d_L < d < d_U$ veya $4-d_U < d < 4-d_L$	ise “No Decision”

Modelde en az bir anlamlılık düzeyi için otokorelasyon saptanmış veya belirsizlik durumu söz konusu ise ikinci seçenek olan Genelleştirilmiş Farklar yöntemi otomatik olarak aktif olmaktadır. Bir başka ifade ile bu yöntem, Durbin-Watson testi uygulanmamışsa veya uygulandıktan sonra otokorelasyon bulunmamış ise aktif hale gelmemektedir. Araştırmacılara büyük kolaylık sağlayan bu olanak incelenen diğer paket programlarda yer almamaktadır.

Bölüm (2.1.5.1.1.)’de bahsedildiği gibi, Genelleştirilmiş Farklar yöntemi değişkenlere dönüşüm uygulandıktan sonra En Küçük Kareler yöntemiyle modelin parametrelerini tahmin etmektedir. Böylece yeni modelin otokorelasyondan arınabileceği varsayımı vardır. Bu varsayımdan hareketle program, bu seçenek seçildiğinde değişken dönüşümünü kendi içinde otomatik olarak uygulayarak genelleştirilmiş fark denkleminin katsayılarını hesaplamakta ve sonuç penceresine eklemektedir. Yeni oluşturulan modelde otokorelasyonun bulunup bulunmadığını saptamak için tekrar Durbin-Watson testi otomatik olarak uygulanarak sonuçlar aynı kriterlere göre yorumlanmaktadır. Böylece araştırmacı, tek bir seçenek ile hem genelleştirilmiş farklar yönteminin modele uygulanmasını sağlamakta hem de yeni

modelin otokorelasyondan arındırılıp arındırılmadığını görebilmektedir. Bu da diğer programlarda yer almayan önemli bir olanaktır.

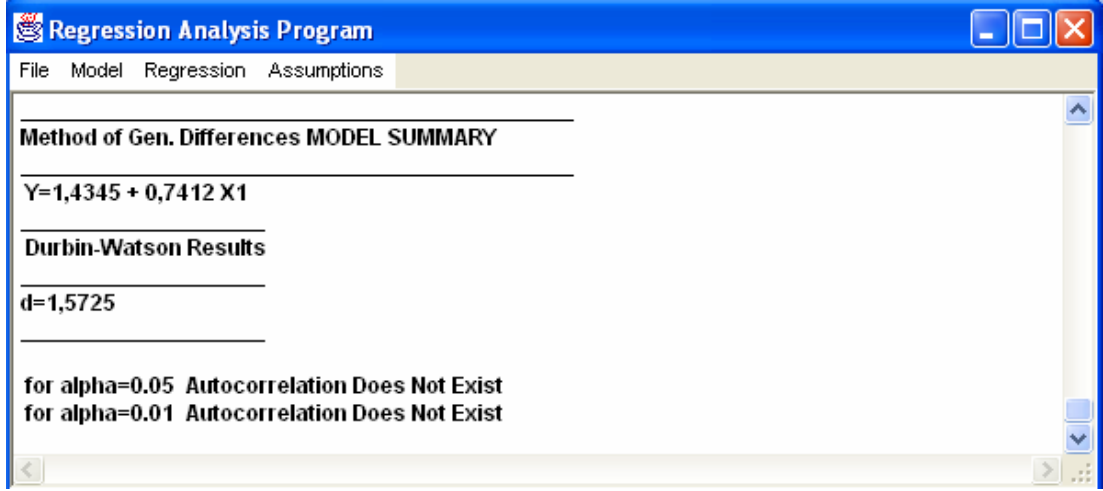
(Ek:3)'deki basit regresyon verilerine “Regression Equation” seçeneği ile uygulanan regresyon analizinden sonra otokorelasyonun varlığını araştırmak üzere “Durbin-Watson Test” seçeneğine basıldığında d test istatistik değeriyle birlikte verilen yorumlarda otokorelasyonun bulunduğu görülmektedir.



**Şekil (3.31): Durbin-Watson Testi**

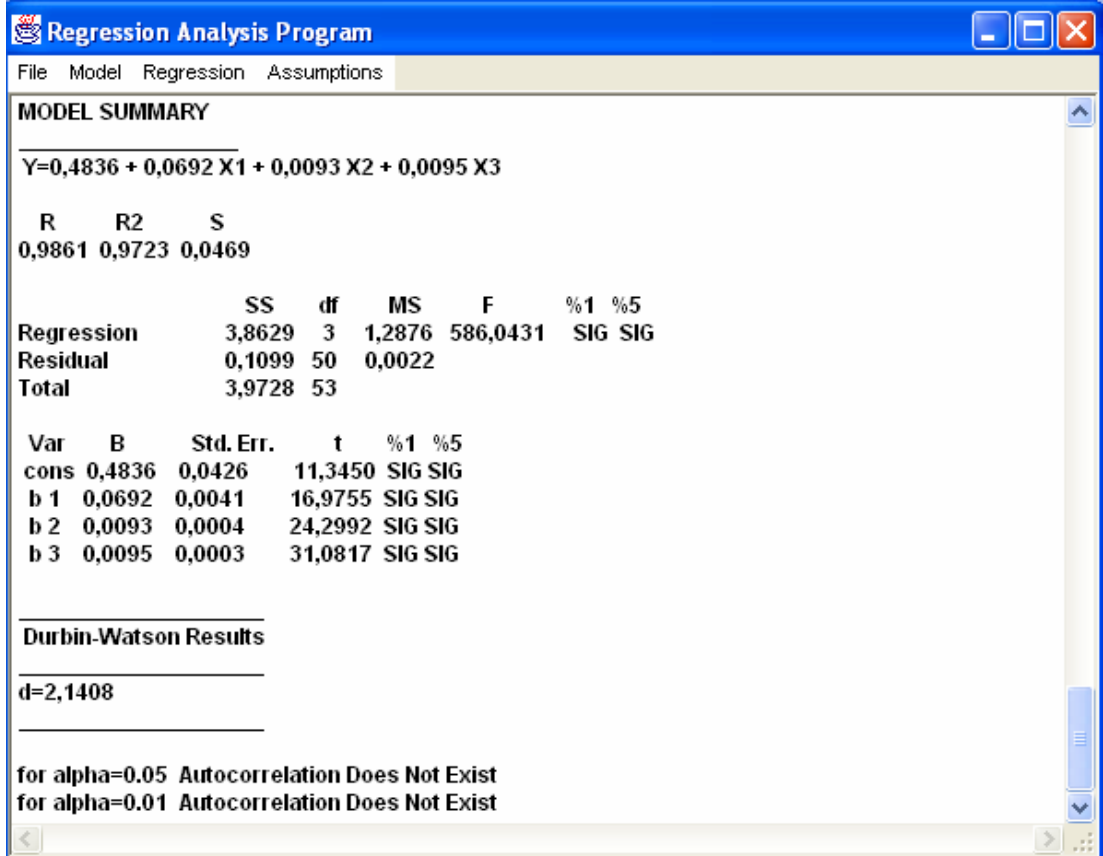
Bu durumda ikinci seçenek otomatik olarak aktif olmakta ve basıldığında genelleştirilmiş fark denklemini ve bu modelin Durbin-Watson test istatistik sonucunu vermektedir.<sup>211</sup>

<sup>211</sup> SPSS çıktısı verilere gerekli dönüşümler “Transform” ve “Compute” seçenekleri ile uygulandıktan sonra elde edilmiştir.



**Şekil (3.32): Genelleştirilmiş Farklar Yöntemi**

(Ek:4)'deki çoklu regresyon analizi için aynı işlemler uygulandığında “Stepwise Regression” seçeneği ile belirlenen model için Durbin-Watson test sonucu şu şekilde elde edilir.



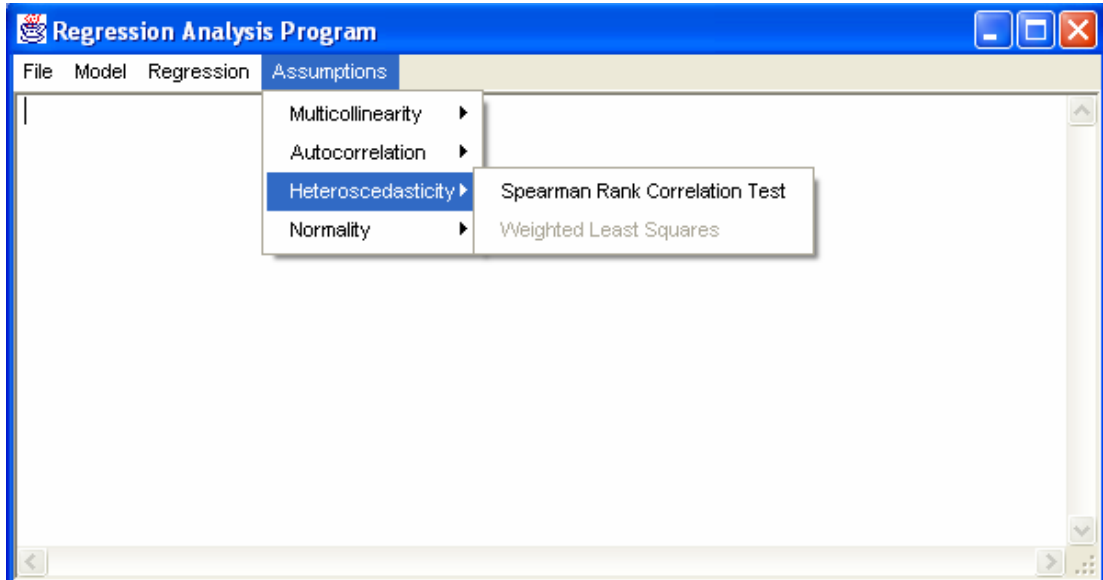
**Şekil (3.33): Çoklu Regresyon Modelinde Durbin-Watson Test Sonuçları**



Durbin-Watson test deęeri (Ek:26)'daki kritik deęerler ile karřılařtırıldıęında programın verdięi yorumların %1 ve %5 anlamlılık dzeyleri iin doęru olduęu gzkmektedir. Modelde otokorelasyon saptanmadıęı iin Genelleřtirilmiř Farklar yntemi aktif olmamaktadır.

### 3.2.4.3. ‘‘Heteroscedasticity’’ Alt Bařlıęı

Eřit varyanslılık varsayımının geerlilięini arařtırmak zere eklenen bu seenekte ncelikle Spearman Sıra Korelasyon testi uygulanmaktadır.



řekil (3.34): ‘‘Heteroscedasticity’’ Alt Bařlıęı

Blm (2.2.3.)'de bahsedilen testler ierisinden bu testin seilmesinin nedeni, farklı varyanslılık durumunda, bu duruma neden olan baęımsız deęiřkenlerin doęrudan saptanabilmesidir. Program geliřtirildięinde dięer testlere de yer verilebilecektir. Bilindięi zere bu test, eřit varyanslılık varsayımının varlıęına dair genel bir yorumdan ok, modelin hata terimleriyle iliřkili baęımsız deęiřkenlerini belirleyebilmektedir.

Bu seçenek seçildiğinde her bir bağımsız değişken ile hata terimi arasındaki Spearman Sıra Korelasyon katsayısı ve test istatistiği hesaplanarak yazılmakta ve %1, %5 anlamlılık düzeyleri için t tablosundaki kritik değerlerle (Ek:23) karşılaştırılarak anlamlılıkları yorumlanmaktadır.

Tüm hesaplanan Spearman Sıra Korelasyon katsayılarının anlamsız olması, bir başka ifade ile hata terimleriyle bağımsız değişkenler arasında anlamlı bir ilişki saptanmamış olması durumunda eşit varyanslılık varsayımı geçerli olacağı için bu başlığın ikinci seçeneği olan “Weighted Least Squares” (AEKK) seçeneği aktif olmamaktadır. Aksi takdirde, yani en az bir anlamlı katsayının bulunması durumunda bu yöntem otomatik olarak aktif hale gelmektedir. Bu olanak da araştırmacılara önemli zaman tasarrufu sağlayacaktır.

Bölüm (2.2.4.2.)’de sözü edilen AEKK yöntemi hata terimleriyle ilişkisi anlamlı olan bağımsız değişkeni ağırlık olarak seçerek regresyon katsayılarını tahmin etmektedir. Aktif halde iken, bu yöntemin seçilmesi durumunda program yeni bir pencere açarak, burada bir önceki adımda Spearman Sıra korelasyon katsayısı en az bir anlamlılık düzeyi için anlamlı bulunan tüm bağımsız değişkenleri sıralamaktadır. Kullanıcı hangi bağımsız değişkeni ağırlık olarak seçmesi gerektiğine karar vererek “OK” düğmesine bastığı takdirde program, bu bağımsız değişkenin 9 değişik değerini ağırlık olarak alarak, AEKK yöntemini uygulamaktadır. Bu değerler:

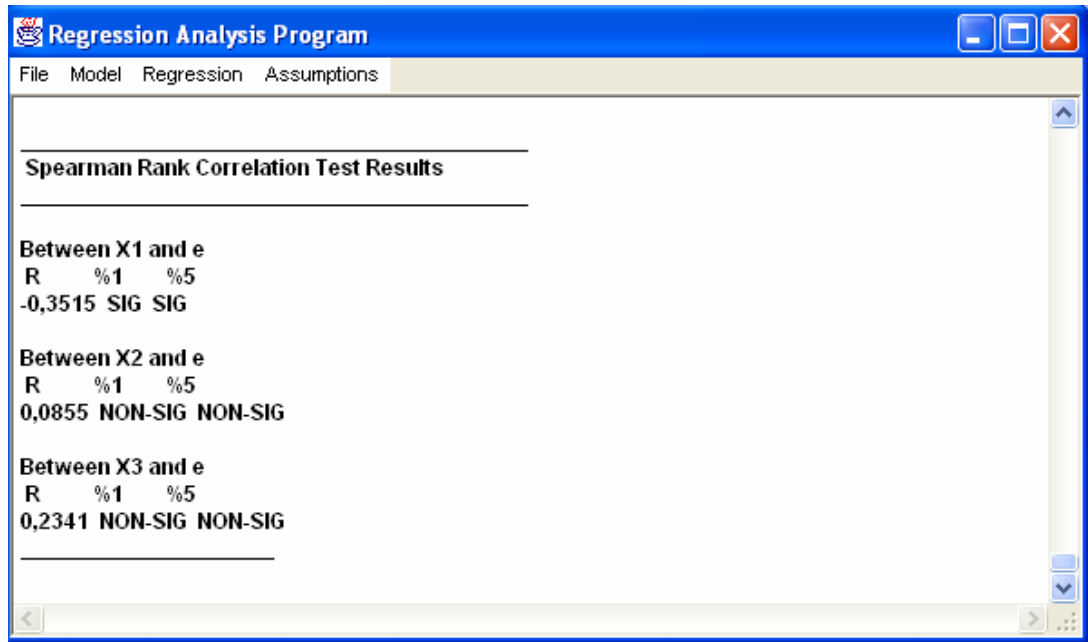
$$1/x^{-2}, 1/x^{-1.5}, 1/x^{-1}, 1/x^{-0.5}, 1/x^0, 1/x^{0.5}, 1/x^1, 1/x^{1.5}, 1/x^2$$

olarak tanımlanmıştır. Bir başka ifade ile bu işlem, bölüm (2.2.4.4.)’de bahsedilen  $\gamma$ ’nın -2’den başlanarak +2’ye ulaşınca kadar 0,5’lik aralıklarla arttırılmasıyla AEKK yönteminin uygulanmasıdır. Aynı bölümde bahsedildiği gibi, farklı ağırlık değerleriyle elde edilen 9 farklı model arasından hangisinin en iyi tahmin olduğuna karar verebilmek için (2.10)’daki formül ile hesaplanan En Büyük Olabilirlik fonksiyonuna bakmak gerekmektedir. Program otomatik olarak, bu işlemleri gerçekleştirerek bu fonksiyonu en büyükleleyen modeli rapora eklemektedir. Bu

seçenek de arařtırmacılara diđer programlarda yer almayan çok önemli bir olanak ve zaman tasarrufu sağlamaktadır.

Arařtırmaacı, anlamlı bulunarak listeye eklenen her bağımsız deęişken için Ağırlıklı En Küçük Kareler yöntemini uygulayabilmektedir. Bu bağlamda, en uygun modelin seçimi arařtırmaacıya bırakılmıştır.

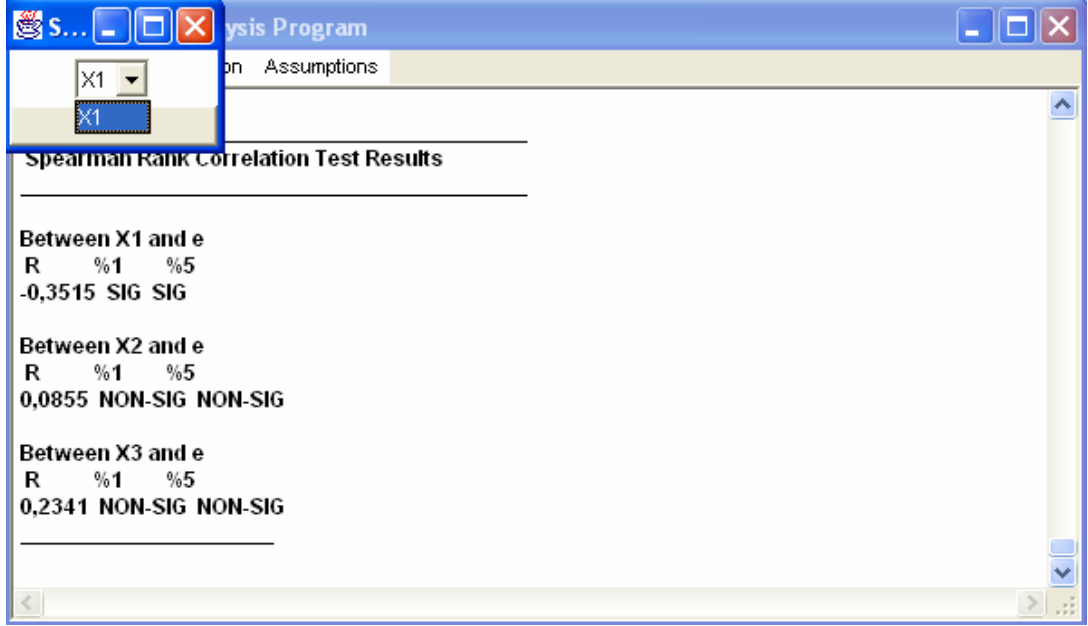
(Ek:4)'deki çoklu modelde eşit varyanslılık varsayımını arařtırmak üzere "Stepwise Regression" ile oluşturulan model için "Spearman Rank Correlation Test" seçeneğine basıldığında şu sonuçlar rapora eklenmektedir.<sup>212</sup>



**Şekil (3.35): Spearman Sıra Korelasyon Katsayıları**

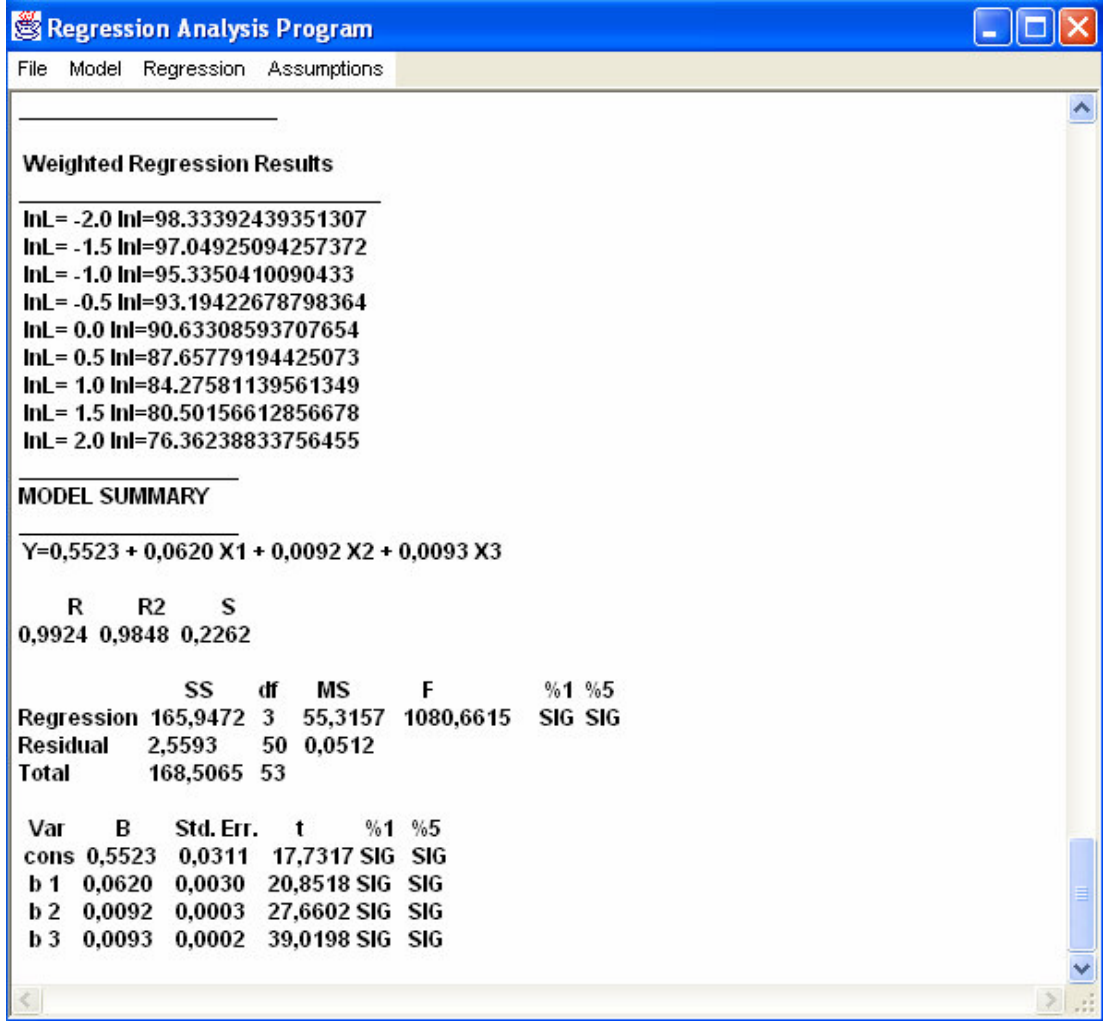
Buradan görüldüğü üzere, sadece  $X_1$ 'in hata terimleriyle ilişkisi %1 ve %5 anlamlılık düzeyleri için anlamlı bulunmaktadır. Bu durumda "Weighted Least Squares" seçeneęi otomatik olarak aktif hale gelmiştir ve seçilmesi durumunda, açılan pencerede sadece "X1" listelenmektedir.

<sup>212</sup> Bu çıktıyı elde edebilmek için, SPSS'de "Stepwise" seçeneęi ile elde edilen modelin hata terimleri kaydedilmiş ve bu deęerlerin mutlak deęerleri hesaplanmıştır.



**Şekil (3.36): Ağırlıklı En Küçük Kareler Yönteminde Ağırlık Değerlerinin Seçimi**

Bu değişkenle AEKK yöntemi uygulandığında elde edilen 9 farklı model için En Çok Olabilirlik fonksiyonun değerleriyle, bu değerlerden en büyüğüyle oluşturulan modelin denklemi yazılmaktadır. Ayrıca elde edilen yeni modele ait varyans analizi, korelasyon ve belirlilik katsayılarıyla modelin standart hatası ve parametre değerleri ile anlamlılıklarının %1 ve %5 anlamlılık düzeyleri için yorumlanarak verilmektedir. Daha önce de belirtildiği gibi bu varsayımdan sapma gözlemlendiğinde de program diğerlerinde olduğu gibi otomatik olarak model değişimi yapabilmektedir.



Şekil (3.37): AEKK Yöntemiyle Elde Edilen Model

### 3.2.4.4. “Normality” Alt Başlığı

Hata terimlerinin normal dağılması gerektiği varsayımını araştırmak üzere eklenen en son alt başlık “Normality” seçeneğidir. Bu seçenekte, bölüm (2.3.2.)’de bahsedilen testler arasından Jarque-Bera testi seçilmiş, varsayımın geçerli olmaması durumunda ise uygulanabilecek yaygın olarak kullanılan algoritmik bir yöntem saptanamadığı için programa eklenememiştir.

Belirlenen modelin hata terimlerinin normal dağılıp dağılmadıklarını araştırmak için, “Regression Equation” veya “Stepwise Regression” seçeneklerinden biri seçilmeli ve regresyon analizi yapılmalıdır. Daha sonra “Jarque-Bera Test” seçeneğine basıldığında, program (2.14) formülüyle JB test istatistiğini hesaplamakta, ve diğer testlerde olduğu gibi bu değeri, %1 ve %5 anlamlılık düzeyleri için tablo değerleriyle karşılaştırarak sonucu yorumlamaktadır. Bölüm (2.3.2.4.)’de bahsedildiği gibi JB değeri 2 serbestlik dereceli Ki-Kare ( $X^2$ ) dağılımına uymaktadır. Bu durumda program kritik değerleri;

%1 için 9,21

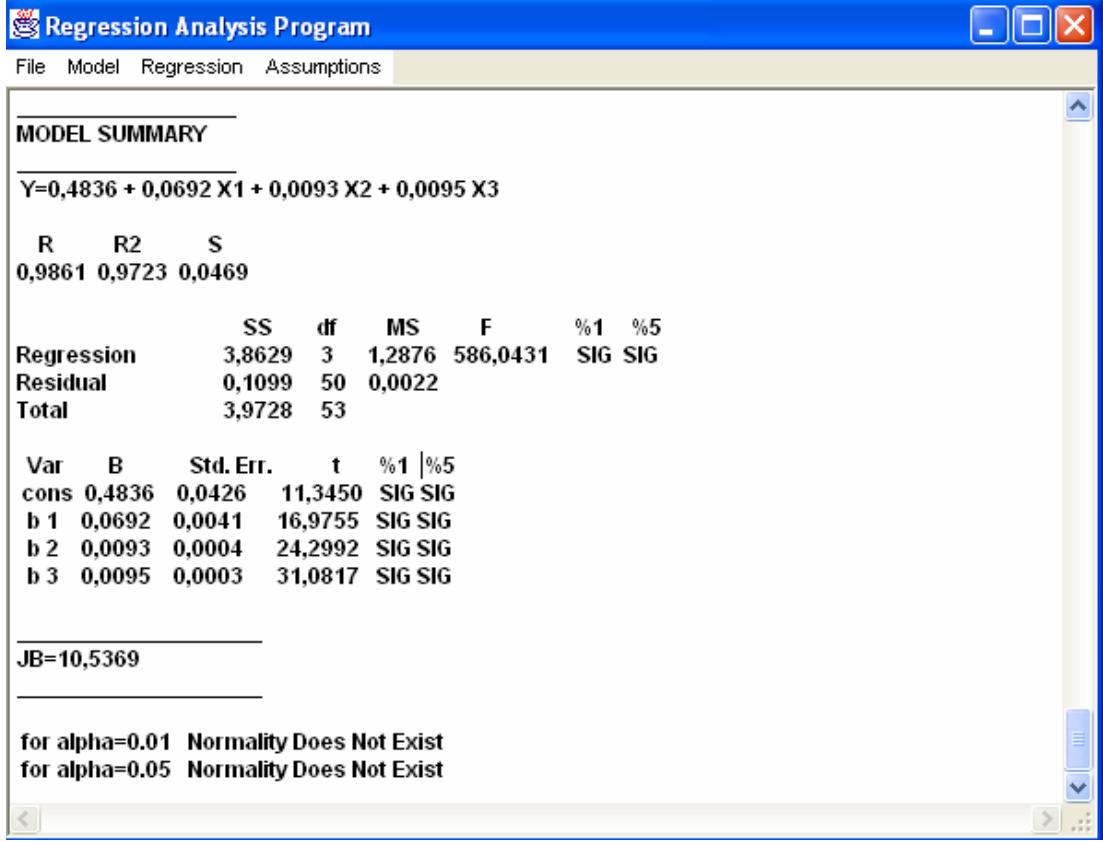
%5 için 5,99

olarak almaktadır. Test istatistik değerinin bu değerlerden küçük olması durumu, hataların normal dağıldığını ( $H_0$  kabul), tersi ise dağılmadığını ( $H_0$  red) göstermektedir. İlk durumda program, “Normality Exists”, ikinci durumda ise “Normality Does Not Exist” mesajlarını vermektedir.

(Ek:4)’deki çoklu regresyon verileri ele alındığında, “Stepwise” seçeneğinden sonra “Normality” alt başlığı ve “Jarque-Bera” testi seçilerek şu değere ulaşılmaktadır:<sup>213</sup>

---

<sup>213</sup> SPSS programında sadece (2.12) ve (2.13) formülleryle belirlenen asimetri (S) ve basıklık (K) ölçülerine ulaşılmaktadır. Bu değerler ile (2.14) formülüyle JB değeri Excel ile hesaplandığında programın verdiği sonuca ulaşılmaktadır.



**Şekil (3.38): Jarque-Bera Testi**

Görüldüğü gibi %1 ve %5 anlamlılık düzeyleri için hata terimleri normal dağılmamaktadır.

Hata terimlerinin normal dağılıma uygun olarak dağılmadıkları durumda bölüm (2.3.3.)’de sözü edilen yöntemlerden yaygın olarak kullanılan Yeniden Ağırlıklandırılmış En Küçük Kareler ve Güçlü Regresyon yöntemleri uygulanabilmektedir. Bunlardan, Ağırlıklandırılmış En Küçük Kareler yönteminin programa ilave edilmesi çalışmanın ileriye dönük hedefleri arasında yer almaktadır.

## 4. TÜRKİYE’NİN 500 BÜYÜK SANAYİ KURULUŞU VERİLERİYLE İŞGÜCÜ VERİMLİLİĞİ ÜZERİNE BİR UYGULAMA

Çalışmanın bu aşamasında, geliştirilen program kullanılarak Türk İmalat Sanayiinde önemli bir performans göstergesi olan işgücü (emek) verimliliği analizi yapılmıştır. Amaç, hem işgücü verimliliğinin hangi ekonomik değişkenlere bağımlı olduğunu regresyon ve korelasyon analiziyle araştırmak, hem de programın bir bütün olarak işleyişini göstermektir.

### 4.1. İşgücü Verimliliği

Türk İmalat Sanayiinde işgücü verimliliğinin uygulama konusu olarak seçilmesinin nedeni, işgücü verimliliğinin işletmelerin gelişmesinde çok önemli bir gösterge olmasıdır. Gelişmekte olan bir ülke olan Türkiye’de kıt üretim kaynaklarının rasyonel kullanımı olan verimlilik, doğal olarak ayrı bir önem taşımaktadır.

Bir endüstri yabancı rakiplerinin seviyesinde ya da daha üzerinde toplam faktör verimliliğine sahipse o endüstri rekabet edebilir bir endüstridir.<sup>1</sup>1980’li yıllardan itibaren bir çok ekonomist, uluslararası rekabet gücünün temelde ihraç fiyatları tarafından belirlendiğini ve bunların da ağırlıklı olarak ücret maliyeti tarafından ileri sürüldüğünü belirtmektedir. Bu yaklaşım, uluslararası rekabet gücünün belirlenmesinde ölçümlerin, ağırlıklı olarak ücret maliyetinin ölçülmesi ve işgücü verimliliği üzerinde odaklaşılması gerektiğini göstermektedir. Bir başka ifade

---

<sup>1</sup> J. Markusen, **Productivity, Competitiveness, Trade Performance and Real Income: The Nexus Among Four Concepts**, Ottawa, 1992, s. 8.



ile, birçok şirket rekabet edebilme ve yeniden yapılandırma problemlerinin çözümü için verimlilik yönetimine başvurmuşlardır.<sup>2</sup>

Verimliliğin ekonomik planlamadaki önemi de yadsınmaz. Verimliliğin düşük olduğu bir ülkede mal ve hizmet üretimini arttırabilmek için işgücü kaynaklarının daha geniş ölçüde kullanılması sağlandığında, ücret gelirlerindeki artış nedeniyle tüketim malları talebinde üretimle karşılanamayacak bir genişlemeye yol açacak ve arz-talep dengesi bozulacaktır. Böyle bir durum ancak kaynakların daha iyi kullanılması ve daha yüksek bir verimlilik ile yeterli üretim artışı sağlanmasıyla önlenmiş olmaktadır.<sup>3</sup>

Bu açıklanmalardan da anlaşılacağı gibi verimlilik, rekabet gücünün ve rasyonel çalışmanın bir göstergesidir ve ölçülmesi gerekmektedir.

Verimlilik en geniş şekliyle birim girdi ünitesine karşı gelen üretim olarak tanımlanır.<sup>4</sup> Uygulamalarda en çok kullanılan verimlilik ölçüleri işgücü ve sermaye verimliliğidir. İşgücü verimliliği genel olarak, çıktı üretiminde kullanılan bir birimlik işgücü girdisi başına üretilen net katma değer olarak tanımlanmaktadır.<sup>5</sup>

$$\text{İşgücü Verimliliği} = \frac{\text{NetKatmaDeğer}}{\text{İşgücüSayısı(İsgunu / İssaatlerisayisi)}}$$

Sermaye verimliliği ise,

$$\text{Sermaye Verimliliği} = \frac{\text{NetKatmaDeğer}}{\text{SermayeMiktari(ToplamNetAktifler)}}$$

---

<sup>2</sup> Robert E. Sibson, **Maximizing Employee Productivity: A Manager's Guide**, New York, Amacom, 1994, s. 1.

<sup>3</sup> Kenan Gürtan, **Türkiye'nin Ekonomik Yapı Problemleri**, İstanbul, Güryay Matbaacılık, 1984, s. 173.

<sup>4</sup> **A. e.**, s. 170-171.

<sup>5</sup> "Verimlilik ve Maliyet Göstergeleri", (Çevrimiçi) <http://www.canaktan.org/yeni-trendler/yeni-rekabet/olcme.htm>, 27 Temmuz 2005.

formülüyle hesaplanmaktadır.

Yukarıdaki formülle hesaplanabilen işgücü verimliliğinin artırılabilmesi için hangi ekonomik değişkenlerden etkilendiğinin bilinmesi gerekmektedir. Çok sayıda faktöre bağımlı olan işgücü verimliliğini başta işletme büyüklüğü, sermaye yoğunluğu-teknoloji düzeyi, kapasite kullanım oranı ve işgücünün kalitesi gibi değişkenler etkilemektedir. Yayınlanmış verilere dayanan araştırmada bu göstergeler arasından ulaşılabilenler kullanılabilmiştir.

## 4.2. Çalışmanın Değişkenleri

Analizde bağımlı değişken olarak tanımlanan işgücü verimliliği

$$Y = \frac{\text{BrütKatmaDeğer}}{\text{ÇalışanSayısı}}$$

formülüyle hesaplanmaktadır.<sup>6</sup>

İşletme büyüklük ölçüsü olarak satış değeri, katma değer, özsermaye ve net aktifler toplamı gibi göstergeler yaygın olarak kullanılmaktadır. Çalışmada,

büyükölçü ölçüsü göstergesi olarak,

$$X_1 = \text{Satış}$$

sermaye yoğunluğu göstergesi olarak,

$$X_2 = \frac{\text{Makine, Tesis ve Cihazlar}}{\text{Çalışan Sayısı}}$$

---

<sup>6</sup> Net Katma Değer yerine Brüt Katma Değer değeri kullanılmıştır.

satış karlılığı,

$$X_3 = \frac{\text{Kar}}{\text{Satış}}$$

finansal oranlardan borçluluk oranı,

$$X_4 = \frac{\text{ToplamBorçlar}}{\text{ToplamAktifler}}$$

ihracatın satışa oranı,

$$X_5 = \frac{\text{İhracat}}{\text{Satış}}$$

değişkenleri seçilmiştir. Ayrıca, işletmelerin özel veya yabancı sermayeli oluşunun işgücü verimliliği üzerinde etkili olup olmadığının incelenmesi için, yabancı sermaye oranını veren bir değişken dikkate alınabilmektedir.

$$X_6 = \text{Yabancı Sermaye Oranı}$$

İşletmelerin farklı sektörlerde buldukları düşünüldüğünde sektör farkının bu değişken üzerinde etkili olup olmadığını araştırmak üzere bir kukla değişken araştırmaya eklenmiştir. Sektörler tüketim malı üretenler ve diğerleri (yatırım ve aramalı üretenler) gibi iki grup olarak tek bir kukla değişkenle tanımlanmıştır. Bu durumda,

$$X_7 = 1 \text{ tüketim malı üretenleri}$$

$$X_7 = 0 \text{ ise diğerlerini}$$

temsil etmektedir.

### 4.3. Çalışmanın Verileri

Uygulamanın verilerinin büyük bir çoğunluğu, İstanbul Sanayi Odası'nın her yıl geleneksel olarak hazırladığı Türkiye'nin 500 Büyük Sanayi Kuruluşu çalışmasının 2003 yılı yayınlarından<sup>7</sup> tesadüfi olarak seçilen kuruluş bilgilerinden oluşmaktadır. Bu yayında kuruluşların toplam satışları, üretimden satışları, yarattıkları brüt katma değer, özsermaye, ve aktif toplamları, kar ve zarar rakamları, ihracat tutarları, çalışan sayıları ve her şirketin kamu, özel veya yabancı sermaye oranları verilmektedir.

Seçilen 60 şirket için, bir önceki bölümde tanımlanan bağımlı Y ve bağımsız değişkenlerden  $X_1$ ,  $X_3$ ,  $X_5$ ,  $X_6$  ile bağımlı değişken bu yayındaki bilgilerden doğrudan veya gerekli işlemler yapılarak kullanılmıştır. Bunun dışında araştırmada söz konusu olan bağımsız değişkenlerden  $X_2$ 'yi hesaplamak için gerekli olan "Makine, Tesis ve Cihazlar" ile  $X_4$  için gerekli olan "Toplam Borçlar" değerlerini her şirket için saptamak üzere, şirketlerin 2003 yılı bilançoları İstanbul Menkul Kıymetler Borsası'nın Web sayfasından<sup>8</sup> erişilerek elde edilmiştir. Son değişken olarak seçilen ve şirketlerin tüketim sektörüne ait olup olmadıklarını gösteren  $X_7$  değerlerini belirlemek için ise Capital dergisinin Web sayfasından şirket sektör bilgilerinden yararlanılmıştır.<sup>9</sup>

Böylece üç kaynak kullanılarak biraraya getirilen, Türkiye'nin 500 Büyük Sanayi Kuruluşu arasına giren 60 sanayi kuruluşu verileri (Ek:5)'de sunulmaktadır.

### 4.4. Yeni Program ile Regresyon Analizinin Uygulanması

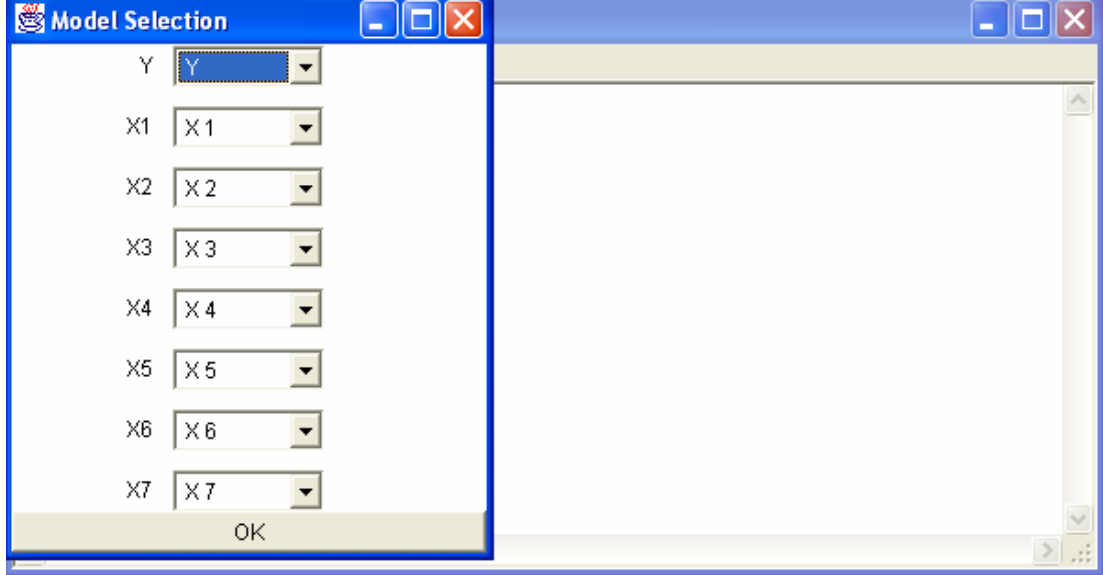
Örnek verilere regresyon analizinin uygulanabilmesi için öncelikle veriler "txt" uzantılı bir dosyaya kopyalanarak kaydedilmiştir. "File" menüsündeki "Load"

<sup>7</sup> Türkiye'nin 500 Büyük Sanayi Kuruluşu, Sanayi Odası Yayınları 2003, Ağustos 2004.

<sup>8</sup> "Mali Tablolar", (Çevrimiçi) <http://www.imkb.gov.tr/bilanço/mtablodonem.htm>, 25 Mayıs 2005.

<sup>9</sup> "Capital 500", (Çevrimiçi) <http://www.capital500.net/capital/ca03-1.asp>, 30 Mayıs 2005.

komutu ile programa yüklendikten sonra “Model” ve ardından “Selection” seçeneği ile tüm değişkenler listelenmektedir.

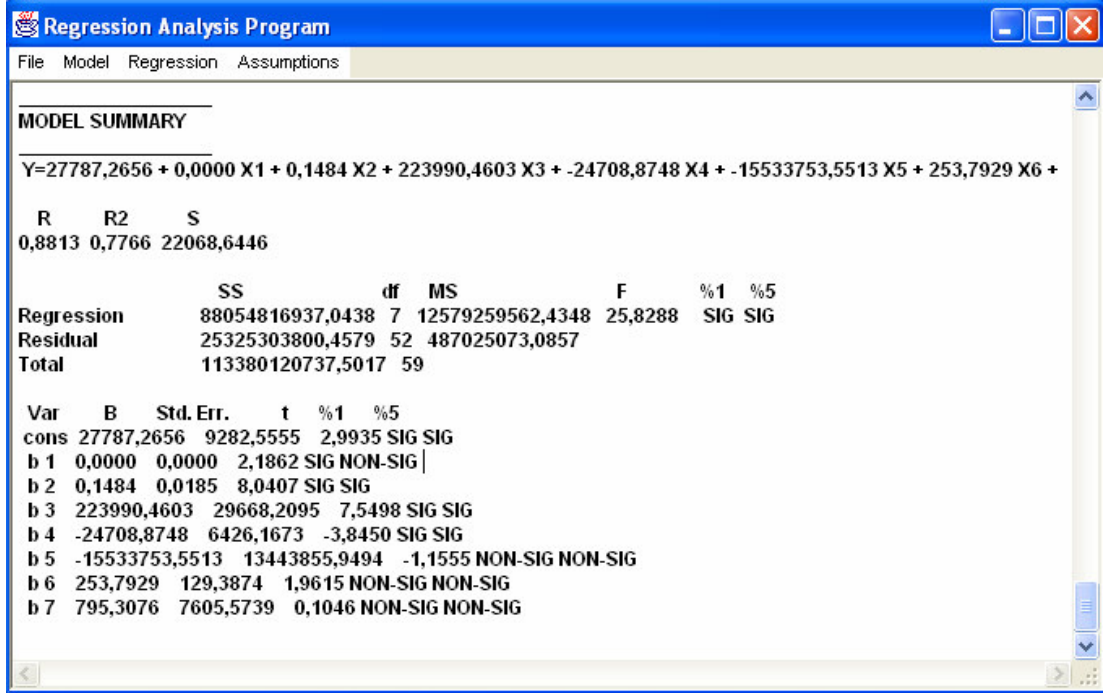


Şekil (4.1): İşgücü Verimliliği Verileri için Model Seçimi

Öncelikle, tüm değişkenler doğrudan seçilerek “OK” düğmesine tıklanarak model seçilmiş, “Correlations” seçeneği ile tüm değişkenler arasındaki basit korelasyon katsayıları incelenmiş ve ardından “Regression Equation” seçeneği ile de regresyon analizi uygulanmış ve sonuçlar Şekil (4.2) ve (4.3)’de iki ayrı çıktı olarak verilmiştir. (Analizin SPSS çıktıları (Ek:19-22)’de verilmektedir.)

	Y	X1	X2	X3	X4	X5	X6	X7
Y	1,000	0,243	0,577	0,437	-0,301	-0,189	0,241	-0,429
X1	0,243	1,000	-0,000	0,070	0,010	0,276	0,430	0,068
X2	0,577	-0,000	1,000	-0,178	-0,225	-0,123	-0,023	-0,387
X3	0,437	0,070	-0,178	1,000	0,212	-0,096	0,077	-0,260
X4	-0,301	0,010	-0,225	0,212	1,000	0,133	-0,017	0,230
X5	-0,189	0,276	-0,123	-0,096	0,133	1,000	0,102	0,017
X6	0,241	0,430	-0,023	0,077	-0,017	0,102	1,000	0,026
X7	-0,429	0,068	-0,387	-0,260	0,230	0,017	0,026	1,000

Şekil (4.2): İşgücü Verimliliği Verileri için Korelasyon Matrisi



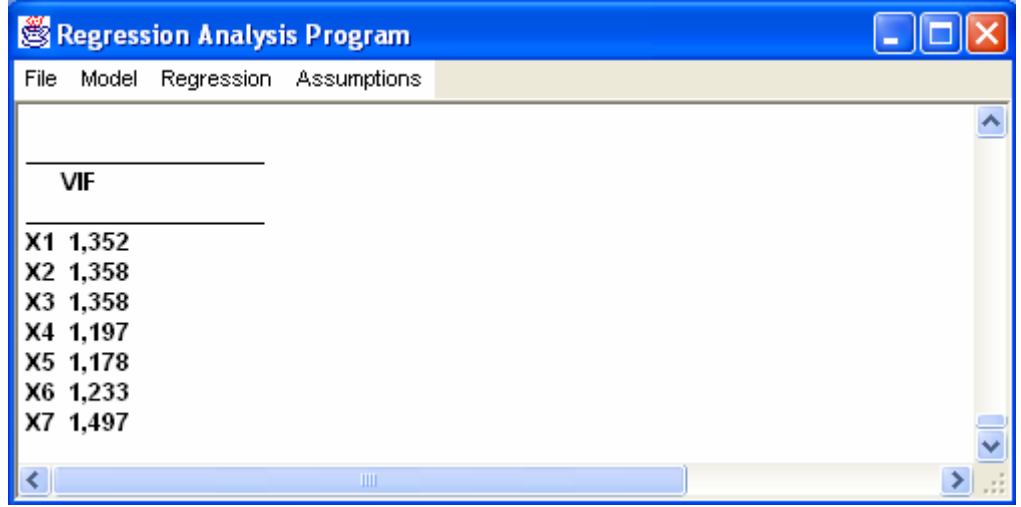
**Şekil (4.3): İşgücü Verimliliği Verileri ile Regresyon Analizi**

Korelasyon matrisinden görüldüğü gibi, bağımlı değişken ile bağımsız değişkenlerden çoğu ile arasında belirli bir ilişki gözlenmektedir. Bağımsız değişkenlerin arasındaki ilişkilere bakıldığında ise, yüksek sayılabilecek korelasyon katsayılarına  $X_1$  ile  $X_6$  ve  $X_2$  ile  $X_7$  arasında rastlanmakta, diğerleri için belirli bir ilişkiden söz edilememektedir.

Regresyon analizi sonuçlarına bakıldığında ise, çoklu korelasyon katsayısının yüksek sayılabilecek 0,88 değerine sahip olduğu, modelin F test sonucunun anlamlı bulunduğu, katsayıların t testinde ise  $b_5$ ,  $b_6$  ve  $b_7$  katsayılarının her iki anlamlılık düzeyinde de anlamsız çıktığı görülmektedir. Bu durumda, uygun değişken setine karar verebilmek için varsayımlardan sapmaların belirlenmesi gerekmektedir.

*Çoklu Doğrusal Bağlantı Bulunmaması Varsayımının İncelenmesi:*

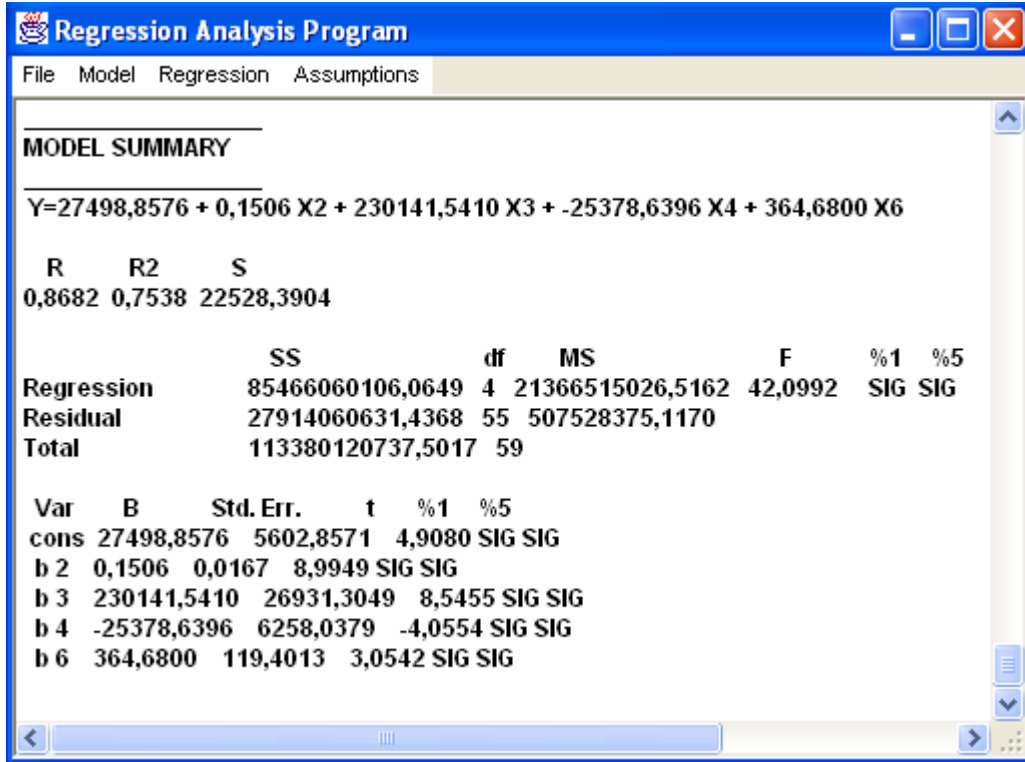
Değişkenler arasında çoklu doğrusal bağlantının bulunup bulunmadığını saptamak için hesaplanan VIF değerleri aşağıda verilmektedir.



VIF	
X1	1,352
X2	1,358
X3	1,358
X4	1,197
X5	1,178
X6	1,233
X7	1,497

**Şekil (4.4): İşgücü Verimliliği Verileri için VIF Değerleri**

Tüm değerler 10'un altında ve hatta 1'e çok yakın olduğu saptandığı için veriler arasında önemli bir çoklu doğrusal bağlantının söz konusu olmadığı anlaşılmaktadır. Ancak regresyon analizi sonuçlarında bazı parametrelerin anlamsız çıkması, başka bir ifade ile %1 ve %5 anlamlılık düzeyleri için, bu bağımsız değişkenlerin anakütlede bağımlı değişkeni anlamlı bir şekilde etkilemedikleri saptandığı için Adım Adım Regresyon yöntemiyle anlamlı değişkenlerden oluşan bağımsız değişken setini belirlemek için, "Stepwise Regression" seçilerek Şekil (4.5)'deki model elde edilmiştir.



**Şekil (4.5): İşgücü Verimliliği Verilerine Adım Adım Regresyon Yönteminin Uygulanması**

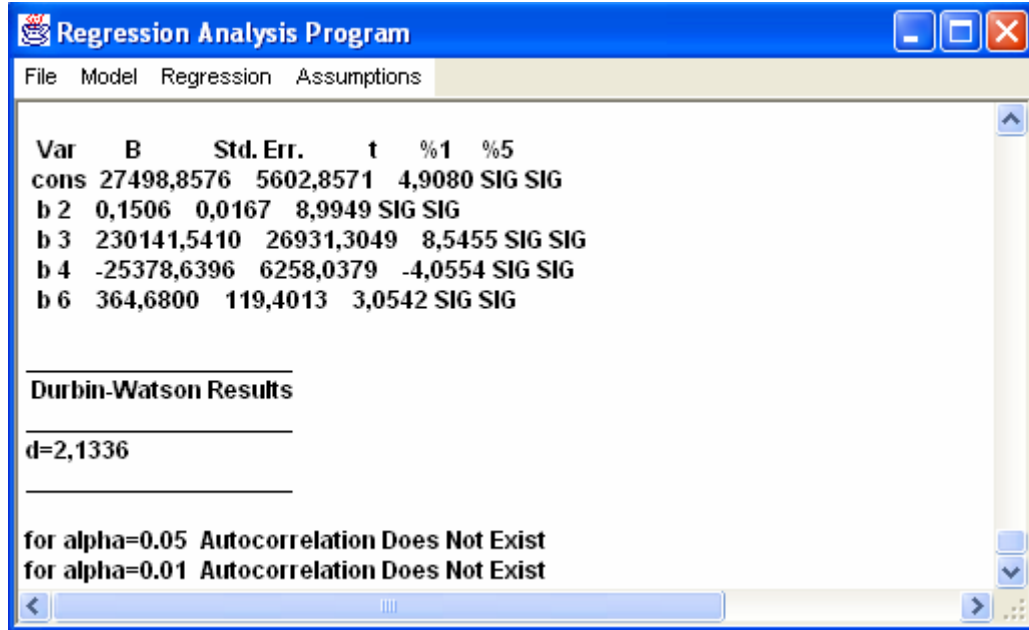
Böylece, modele girmesi gereken bağımsız değişkenler sermaye yoğunluğu ( $X_2$ ), satış karlılığı ( $X_3$ ), borçluluk oranı ( $X_4$ ) ve yabancı sermaye oranı ( $X_6$ ) olarak belirlenmektedir.

Daha önce de açıklandığı gibi bu menünün son seçeneği olan Ridge Regression, çoklu doğrusal bağlantı bulunması durumunda tüm değişkenleri de modele alarak regresyon katsayılarını yanlış olarak tahmin etmektedir. Bu varsayımın geçerliliği saptandığı için “Ridge Regression” yöntemi uygulanmamıştır. Ancak, bilindiği gibi varsayımdan sapma olduğunda geliştirilen program bu aşamada araştırmacıya hemen çözüm yöntemi sunabilmektedir.

#### Otokorelasyon Bulunmaması Varsayımının İncelenmesi:

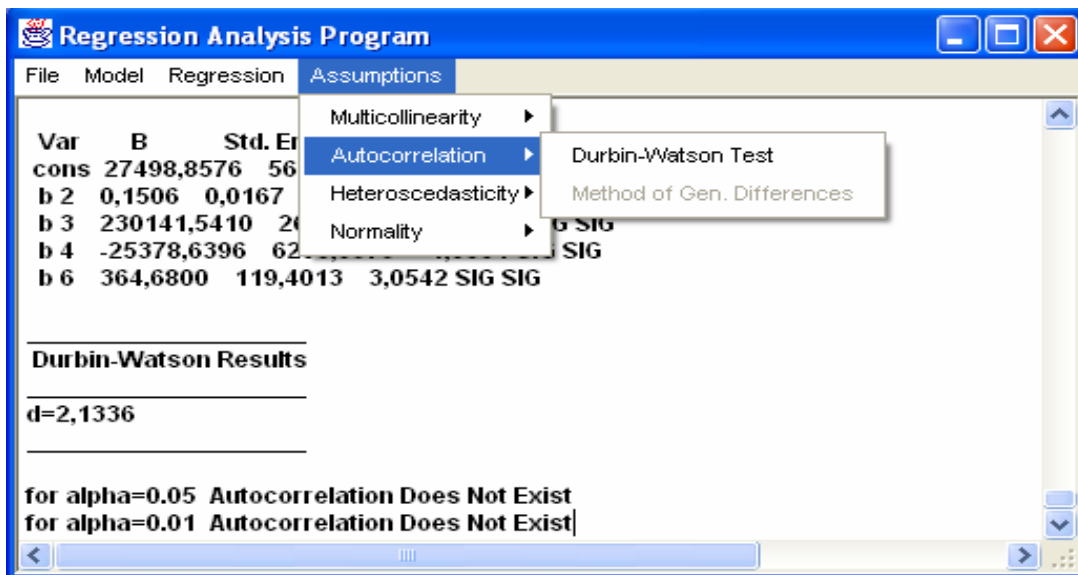
Adım Adım Regresyon yöntemiyle belirlenen yeni modelde otokorelasyonun bulunup bulunmadığının araştırılması için “Autocorrelation” menüsünden “Durbin-Watson” testi seçilerek aşağıdaki sonuçlar elde edilmiştir.





**Şekil (4.6): İşgücü Verimliliği Verilerine Uygulanan Durbin-Watson Testi Sonuçları**

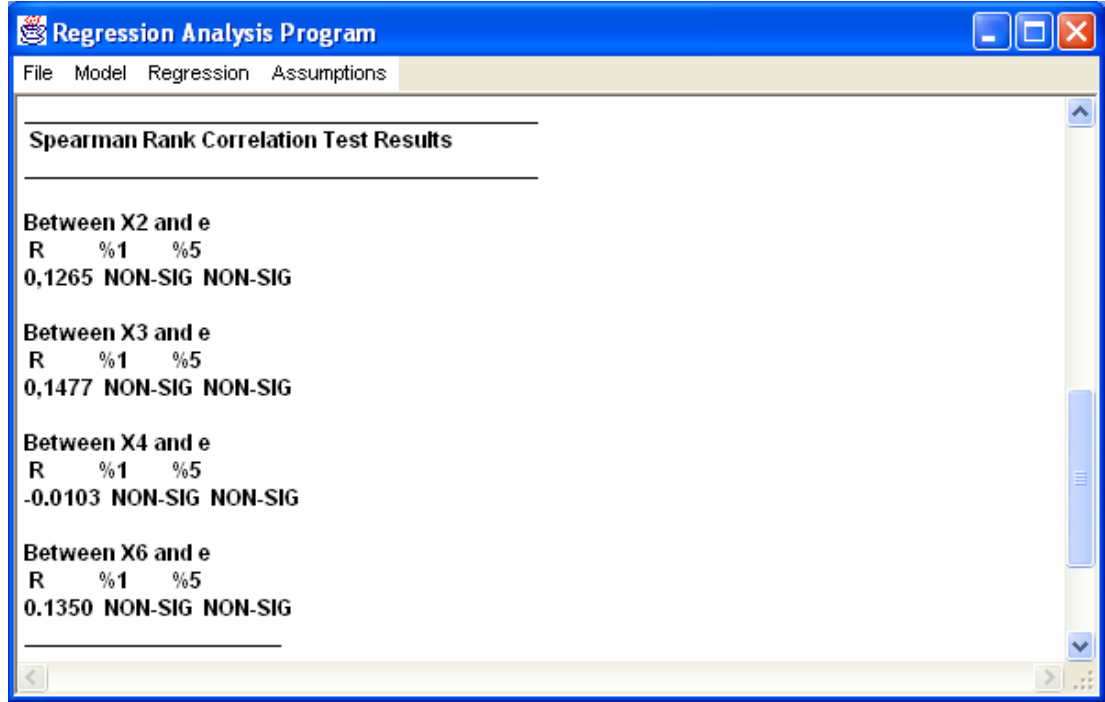
Her iki anlamlılık düzeyinde de modelde anlamlı bir otokorelasyon bulunmadığı anlaşıldığı için ikinci seçenek olan Genelleştirilmiş Farklar yöntemini çalıştıran “Method of Gen. Differences” seçeneği aktif duruma gelmemektedir (Şekil (4.7)). Şayet otokorelasyon çıksaydı programla hemen bu çözüm yöntemi uygulanabilecekti.



**Şekil (4.7): Genelleştirilmiş Farklar Yönteminin Aktif Olmaması Durumu**

### Eşit Varyanslılık Varsayımının İncelenmesi:

Bu aşamada Adım Adım Regresyon yöntemiyle belirlenen modelde eşit varyanslılık varsayımının geçerliliği “Spearman Rank Correlation Test” seçeneğiyle araştırılmış ve aşağıdaki bulgulara ulaşılmıştır.

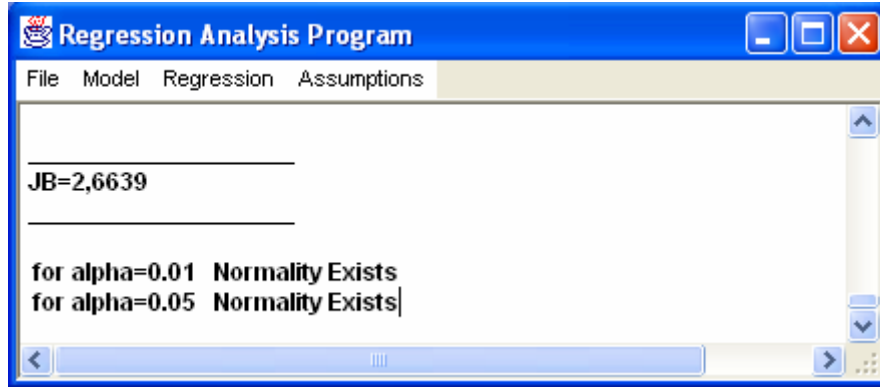


**Şekil (4.8): İşgücü Verimliliği Verilerine Spearman Sıra Korelasyon Testinin Uygulanması**

Şekil (4.8)'den, tüm değişkenler ile hata terimleri arasında her iki anlamlılık düzeyinde de anlamlı bir ilişki saptanmadığı anlaşılmaktadır. Bir başka ifadeyle, tüm sıra korelasyon katsayıları anlamsız bulunduğu için modelde eşit varyanslılık varsayımının geçerli olduğuna karar verilmektedir. Bu nedenle, “Heteroscedasticity” alt başlığına tıklandığında “Weighted Least Squares” seçeneği aktif hale gelmemektedir.

### Normallik Varsayımının İncelenmesi:

Bu araştırma için, “Stepwise Regression” seçeneğinden sonra “Normality” alt başlığındaki “Jarque-Bera” testi seçildiğinde program tarafından hesaplanan test istatistik değeriyle anlamlılıkları aşağıdaki gibi elde edilecektir.



**Şekil (4.9): İşgücü Verimliliği Verilerine Uygulanan Jarque-Bera Testinin Sonuçları**

Normallik varsayımının modelde geçerli olduğu açıkça görülmektedir.

Sonuç olarak, işgücü verimliliği üzerinde yapılan araştırmada analize sokulan 7 bağımsız değişken arasından dördü anlamlı bulunarak, anakütlede işgücü verimliliğinin değişkenliğini 0,87'lik bir oranla açıklayabildikleri gösterilmiştir. Regresyon analizinin varsayımları incelendiğinde, elde edilen bu modelde tüm varsayımların geçerli oldukları uygulanan testler ile saptanmış ve bu yüzden geliştirilen programda hiç bir model değişikliğine gerek duyulmamıştır. Çoklu Korelasyon katsayısının da oldukça yüksek olduğu dikkate alındığında, işgücü verimliliğini anakütlede sermaye yoğunluğu, satış karlılığı, borçluluk oranı ve yabancı sermaye oranının etkilediği, diğer değişkenlerin önemli bir etkisinin olmadığı söylenebilmektedir. Bu durumda, tüm varsayımların geçerli olduğu Adım Adım Regresyon yöntemiyle elde edilen bu modelin, işgücü verimliliğinin ileriye yönelik tahmin ve politika belirlenmesinde kullanılabileceği kararı alınabilmektedir.

Ancak, çoklu korelasyon katsayısının daha da yüksek olduđu bir modele ulaşmak için modele yeni deđişkenler ekleme, bazen de gözlem sayısını arttırma yoluna gidilebilmektedir.

## SONUÇ

Bilindiği gibi günümüzde karar verme durumunda olan her yönetici için olduğu gibi, tüm bilim dallarında yapılan araştırmalarda geleceği tahmini ve/veya çeşitli konularda politika belirlemek büyük önem taşımaktadır. İstatistik yöntemler arasında bu amaca yönelik yöntem ise “İlişki Analizi-Regresyon ve Korelasyon Analizi”dir. Regresyon analizinde, öncelikle seçilen bağımsız değişkenlerin bağımlı değişken arasındaki basit doğrusal korelasyon katsayıları ile bağımsız değişkenler arasında basit doğrusal korelasyon katsayıları hesaplanarak bağımsız değişkenlerin modele girebilirliğinin irdelenmesi gerekmektedir. Daha sonra, bağımlı değişkenle anlamlı ilişkisi olan ancak diğer bağımsızlarla anlamsız ilişkisi olan bağımsız değişkenlerle çoklu regresyon denkleminin yazılması ve regresyon katsayıları ile çoklu korelasyon katsayısının test edilmesi aşamasına geçilmektedir. Bu aşamalardan sonra, modelin En Küçük Kareler yönteminin temel varsayımlarını sağlayıp sağlamadığının kontrol edilmesi gerekmektedir. İkinci bölümde ayrıntılı olarak ele alınan bu varsayımlar, modelde bağımsız değişkenler arasında çoklu doğrusal bağlantı bulunmaması, hata terimleri arasında otokorelasyon bulunmaması, her bir bağımsız değişken değerine karşı gelen hata terimleri varyansının eşit olması ve hata terimlerinin normal dağılımlarıdır.

Bu çalışmada geliştirilen nesne yönelimli bilgisayar programı, araştırmacının regresyon analizinin tüm aşamalarını kolaylıkla ve kısa zamanda uygulayabilmesini sağlamaktadır.

Sonuçları, yaygın olarak kullanılan istatistik programı SPSS ile doğrulanmış olan bu programın en kuvvetli yönlerinden biri, regresyon analizi için gerekli tüm test ve yöntemlerin tek bir menü altında toplanmış olup hiç bir ek hesap veya işleme gerek olmadan kolaylıkla uygulanabilir olmasıdır. Bazı istatistik programlarda tüm işlemlerin komutlarla çalıştırıldığı da dikkate alındığında, programın bu özelliği ile araştırmacılara zaman kazandırarak sonuçlara daha hızlı odaklanmalarını sağladığı

söylenbilir. Programın bir önemli üstünlüğü de, doğrusal olarak açıklanamayan modellere regresyon analizini uygulamak üzere, yaygın olarak uygulanması gereken dönüşüm fonsiyonlarının, değişkenlerin hemen altında açılan listelerde toplanmasıdır. Böylece, her değişken için farklı fonsiyonlar kolaylıkla seçilmekte veya çoklu regresyon modellerinde bağımsız değişkenler aynı yolla model dışında tutulabilmektedir. Analizin herhangi bir aşamasında yapılabilen model değişiklikleri ile regresyon analizi ve varsayımdan sapmaların saptanması aşamaları ard arda hızlı bir biçimde gerçekleştirilerek, veri için en uygun model seçilebilmektedir. Programın önemli bir diğer üstünlüğü de, diğer programların verdikleri sonuçlara ek olarak test sonuçlarını yorumlaması ve araştırmacının daha sonraki adımda yapması gerekli analizlere yol göstermesidir. Bu amaçla, tüm testlerde test istatistikleri çeşitli bilim dallarında yaygın olarak kullanılan %1 ve %5 anlamlılık düzeyleri için gerekli tablo değerleriyle otomatik olarak karşılaştırılarak test sonuçları yorumlanmaktadır.

Varsayımların araştırılmasına gelindiğinde çoklu doğrusal bağlantı varsayımında, öncelikle VIF değerleri hesaplanıp gösterilmekte, Adım Adım Regresyon yöntemi uygulanarak model için anlamlı bağımsız değişkenlerle tahminde kullanılacak fonsiyon yazılabilmektedir. Çoğu programda doğrudan kullanılmayan çoklu doğrusal bağlantı durumunda çözüm getiren Ridge Regression yöntemi, bu alt başlıkta bir seçenek olarak hemen sunulmaktadır. Otokorelasyon varsayımında, Durbin-Watson testinin sonuçları otokorelasyon bulunduğunu göstermesi halinde, çözümleyici yöntemlerden biri olan Genelleştirilmiş Farklar yöntemi hemen uygulanabilmekte, model belirlenmekte ve ardından yeni model için gerekli test sonuçları otomatik olarak verilmektedir. Benzer yaklaşım eşit varyanslılık varsayımı için de geçerlidir. Spearman Sıra Korelasyon testinin sonucuna göre, gerekiyorsa (farklı varyanslılık durumu) Ağırlıklı En Küçük Kareler yöntemi hemen uygulanabilmektedir. Bu yöntemde ağırlık olarak kullanılması uygun olacak bağımsız değişkenler ayrı bir pencerede otomatik olarak listelenmektedir. Hata terimlerinin normal dağılıp dağılmadıklarının sınanması için ise Jarque-Bera testi aynı seçenek içinde uygulanmakta ve sonuçlar yorumlanarak verilmektedir.

Geliştirilen bu programın işleyişini test etmek için kullanılan küçük örneklerden sonra programın kapsamlı bir araştırmada uygulanması uygun görülmüş ve Türk İmalat Sanayiinde verimlilik analizi yapılmasına karar verilmiştir. Verimlilik analizi ve imalat sanayiinin seçilmesinin nedenleri arasında verimliliğin çok önemli bir performans göstergesi olması, ülke ekonomisindeki önemli katkısı ve imalat sanayiinin sanayi sektörü göstergesi sayılmasıdır. Sanayi Odası yayınlarından Türkiye'nin En Büyük 500 Kuruluşu dergisinin 2003 yılı çalışmalarından derlenen 60 işletmeye ait bilgilerle uygulanan işgücü verimliliği analizlerinin sonuçlarına göre bu değişkeni anaküttele etkileyen faktörler, sermaye yoğunluk göstergesi olan makine, tesis ve cihazlar değerinin çalışan sayısına oranı, satış karlılığı, borçluluk oranı ve yabancı sermaye oranı olarak belirlenmiştir. Bu bağımsız değişkenlerin anaküttele işgücü verimliliğini 0,87'lik bir oranda açıklayabildikleri gösterilmiştir. Varsayımdan sapmalara gelindiğinde, modelde çoklu doğrusal bağlantının söz konusu olmadığı, otokorelasyonun bulunmadığı, eşit varyanslılık ve normallik varsayımlarının geçerli oldukları uygulanan testlerle saptanmıştır. Hiç bir model değişikliğine gerek duyulmadan, Adım Adım Regresyon yöntemiyle elde edilen bu modelin, işgücü verimliliği üzerine tahmin ve politika belirlemede kullanılabileceği kararı alınmıştır. Geliştirilen programın sonuçları SPSS ve Excel 2003 ile bir kez daha test edilmiştir. Böylece hem programın işleyişi bir kez daha kontrol edilmiş hem de imalat sanayiinde önemli olan bir verimlilik analizi gerçekleştirilmiştir.

İstatistik yöntemleri arasında yaygın kullanımı olan İlişki-Regresyon ve Korelasyon-Analizini bir bütün olarak tek bir başlık altında gerçekleştirmeyi amaçlayan bu programın başta karar vericiler olmak üzere çeşitli bilim dallarındaki araştırmacılara yararlı olması temenni edilmektedir. Diğer gelişmiş istatistik paket programlarında farklı menü ve alt menülerle gerçekleştirilen bu analiz böylece daha kolay ve kısa zamanda gerçekleştirilebilecektir. Kuşkusuz programın bu üstünlükleri yanında bazı geliştirilmesi gereken seçenekleri de vardır. Bunlar arasında, varsayımlardan sapmaların araştırılmasının, yaygın olarak kullanılan testler arasından seçilen tek testlerle gerçekleştirilmesi ve varsayımlardan sapma durumunda yine yaygın olarak kullanılan çözüm yollarından birer tanesinin seçilmesi sayılabilir. Son olarak belirtmek gerekir ki, programı kullanacak olanların görüşlerini

de alarak iliřki analiziyle ilgili yeni geliřtirilen test ve yntemleri de eklemek ve bylece programı daha fazla geliřtirmek alıřmanın ileriye dnk hedefleri arasında yer almaktadır.



## KAYNAKÇA

- Akın, Fahamet: **Ekonometri**, Bursa, Beta Basım A.Ş., 2002.
- Anderson, David R.,  
Dennis J. S., Thomas A.  
W.: **Statistics for Business and Economics**, eighth edition,  
Ohio, South-Western Thomson Learning, 2002.
- Belsley, David A., Edwin  
Kuh, Roy E. Welsch: **Regression Diagnostics: Identifying Influential Data  
and Sources of Collinearity**, New York, John Wiley &  
Sons, 1980.
- Bera, Anil K., Carlos M.  
Jarque: “Efficient Tests for Normality, Heteroscedasticity and  
Serial Independence of Regression Residuals: Monte  
Carlo Evidence”, **Economic Letters**, Vol.VII, 1981, s.  
313-318.
- Berry, William D.: **Understanding Regression Assumptions**, Newbury Park,  
CA, Sage Publications, Inc., 1993.
- Breusch, Trevor, Adrian  
Pagan: “A Simple Test for Heteroscedasticity and Random  
Coefficient Variation”, **Econometrica**, Vol.XLVII, 1979,  
s. 1287-1294.
- Capital 500: “Capital 500”, (Çevrimiçi)  
<http://www.capital500.net/capital/ca03-1.asp>, 30 Mayıs  
2005.
- Carroll, Raymond J.,  
Daren B. H. Cline: “An Asymptotic Theory for Weighted Least Squares with  
Weights Estimated by Replication”, **Biometrika**,  
Vol.LXXV, 1988, s. 35-43.
- Carrol, Raymond J.,  
David Ruppert: “Robust Estimation in Heteroscedastic Linear Models”,  
**Annals of Statistics**, Vol.X, 1982, s. 429-441.
- Carrol, Raymond, J.,  
David Ruppert: **Transformation and Weighting in Regression**, New  
York, Chapman and Hall, 1988.
- Chatterjee, Samprit,  
Bertram Price: **Regression Analysis By Example**, second edition, New  
York, John Wiley & Sons, Inc., 1991.
- Davidian, Marie: “Estimation of Variance Components in Assays with  
Possibly Unequal Replication and Nonnormal Data”,  
**Biometrika**, Vol.LXXVII, 1990, s. 43-54.

- Deaton, M.L., Marion R. Reynolds, Raymond H. Myers: “Estimation and Hypothesis Testing in Regression in the Presence of Nonhomogeneous Error Variances”, **Communications in Statistics, Simulation and Computation**, Vol.XII, 1983, s. 45-66.
- Draper, Norman R., Harry Smith: **Applied Regression Analysis**, New York, John Wiley & Sons, 1981.
- Durbin, J.: “Estimation of Parameters in Time-Series Regression Models”, **Journal of the Royal Statistical Society**, Vol.XXII, 1960, s. 139-153.
- Durbin, James, Graham Watson: “Testing for Serial Correlation in Least Squares Regression-I”, **Biometrika**, Vol. XXXVII, 1950, s. 409-428.
- Durbin, James, Graham Watson: “Testing for Serial Correlation in Least Squares Regression-II”, **Biometrika**, Vol. XXXVIII, 1951, s. 159-178.
- Edwards, Allen L.: **Introduction to Linear Regression and Correlation**, second edition, New York, W.H. Freeman and Company, 1984.
- Ertek, Tümay: **Ekonometriye Giriş**, 2. bs., İstanbul, Beta, 2000.
- Farebrother, Richard W.: “The Durbin-Watson Test For Serial Correlation When There Is No Intercept in the Regression”, **Econometrica**, Vol.XLVIII, 1980, s. 1553-1563.
- Flury, Bernhard: **Common Principal Components & Related Multivariate Models**, New York, John Wiley & Sons, 1988.
- Fox, John: **Applied Regression Analysis, Linear Models, and Related Methods**, California, Sage Publications, 1997.
- Geary Roy C.: “Relative Efficiency of Count of Sign Changes for Assesing Residual Autocorrelation in Least Regression”, **Biometrika**, Vol.LVII, 1970, s. 123-127.
- Glejser, Herbert: “A New Test for Heteroscedasticity”, **Journal of the American Statistical Association**, Vol.LXIV, 1969, s. 316-323.

- Godfrey, Leslie: “Testing for Multiplicative Heteroscedasticity”, **Journal of Econometrics**, Vol.VIII, 1987, s. 227-236.
- Goldberger, Arthur: **A Course in Econometrics**, Cambridge, Harvard University Press, 1991.
- Goldfeld, Stephen, Richard Quandt: “Some Tests for Heteroscedasticity”, **Journal of the American Statistical Association**, Vol.LX, 1965, s. 539-547.
- Greene, William H.: **Econometric Analysis**, New York, Macmillan Publishing Company, 2000.
- Gujarati, Damodar N.: **Temel Ekonometri**, Çev. Ümit Şenesen, Gülay Günlük Şenesen, İstanbul, Literatür, 1999.
- Günel, Alptekin: “Regresyon Denkleminin Başarısını Ölçmede Kullanılan Belirleme Katsayısı ve Kritiği”, **Doğuş Üniversitesi Dergisi**, C.IV, 2003, s. 133-140.
- Gürüş, Selahattin, Ebru Çağlayan: **Ekonometri Temel Kavramlar**, İstanbul, Der Yayınevi Yayın No:282, 2000.
- Gürtan, Kenan: **Türkiye’nin Ekonomik Yapı Problemleri**, İstanbul, Güryay Matbaacılık, 1984.
- Harvey, Andrew, Garry Phillips: “A Comparison of the Power of Some Tests for Heteroscedasticity in the General Linear Model”, **Journal of Econometrics**, Vol.II, 1974, s. 307-316.
- Hawkins, D. M.: “The Feasible Set Algorithm for Least Median of Squares Regression”, **Computational Statistics and Data Analysis**, Vol. XVI, 1993, s. 81-101.,
- Hawkins, D. M.: “The Feasible Solution Algorithm for Least Trimmed Squares Regression”, **Computational Statistics and Data Analysis**, Vol. XVII, 1994, s. 185-196.
- Hawkins, Douglas M., Xiangrong Yin: “A Faster Algorithm for Ridge Regression”, **Computational Statistics & Data Analysis**, Vol. XL, 2002, s. 253-262.
- Hoerl, A. E. , R. W. Kennard: “Ridge Regression: Biased Estimation for Nonorthogonal Problems”, **Technometrics**, Vol. XII, 1970a, s. 55-67.
- Hoerl, A. E., R. W. Kennard: “Ridge Regression: Applications to Nonorthogonal Problems”, **Technometrics**, Vol. XII, 1970b, s. 69-82.

- Huber, Peter J.: “Robust Regression: Asymptotics, Conjectures, and Monte Carlo”, **Annals of Statistics** **1**, 1973, s. 799-821.
- Huber, Peter J.: **Robust Statistics**, New York, Wiley-Interscience, 2004.
- İMKB: “Mali Tablolar”, (Çevrimiçi) <http://www.imkb.gov.tr/bilanço/mtablodonem.htm>, 25 Mayıs 2005.
- İpek, Orhan: “Ridge Regresyon Üzerine Bir Çalışma”, (Çevrimiçi) <http://idari.cu.edu.tr/sempozyum/bil28.htm> 28 Ocak 2005.
- İSO: “Türkiye’nin 500 Büyük Sanayi Kuruluşu”, Sanayi Odası Yayınları 2003, Ağustos 2004.
- Jarque, Carlos M., Anil K. Bera: “A Test for Normality of Observations and Regression Residuals”, **International Statistical Review**, Vol.LV, 1987, s. 163-172.
- Jeong, Jinook, Kyoungwoo Lee: “Bootstrapped White’s Test for Heteroscedasticity in Regression Models”, **Economics Letters**, Vol. LXIII, 1999, s. 261-267.
- Johnston, Jack: **Econometric Methods**, 3. edition, McGraw-Hill, New York, 1984.
- Judge, G., C. Hill, W. Griffiths, T. Lee, H. Lutkepol: **Introduction to the Theory and Practice of Econometrics**, New York, John Wiley & Sons, 1982.
- Kendall, Maurice G., William R. Buckland: **A Dictionary of Statistical Terms**, New York, Hafner Publishing Company, 1971.
- Klein, Lawrence R.: **Introduction to Econometrics**, New Jersey, Prentice-Hall, 1962.
- Kleinbaum, David G., L. L. Kupper, K. E. Muller, A. Nizam: **Applied Regression Analysis and Other Multivariable Methods**, 3rd edition, Washington, Duxbury Press, 1998.
- Kmenta, Jan: **Elements of Econometrics**, New York, Macmillan Publishing, 1971.
- Koenker, Roger: “A Note on Studentizing a Test for Heteroscedasticity”, **Journal of Econometrics**, Vol.XVII, 1981, s. 107-112.

- Koenker, Roger, Gilbert Basset: “Robust Tests for Heteroscedasticity Based on Regression Quantiles”, **Econometrica**, Vol.L, 1982, s. 43-61.
- Leslie, J. R., M.A. Stephens, S. Fotopoulos: “Asymptotic Distribution of the Shapiro-Wilk W For Testing for Normality”, **The Annals of Statistics**, Vol. XIV, 1986, s. 1497-1506.
- Long , J. Scott: **Regression Models for Categorical and Limited Dependent Variables**, California, Sage Publications, Inc., 1997.
- Maddala, G. S.: **Introduction to Econometrics**, second edition, New York, Macmillan, 1992.
- Markusen, J.: **Productivity, Competitiveness, Trade Performance and Real Income: The Nexus Among Four Concepts**, Ottawa, 1992.
- Marquardt, D. W., R. D. Snee: “Ridge Regression in Practice”, **The American Statistician**, Vol. XXIX, No. 1, 1975, s. 4-9.
- Miles, Jeremy, Mark Shevlin: **Applying Regression & Correlation**, London, SAGE Publications, 2001.
- Montgomery, Douglas, Elizabeth Peck: **Introduction to Linear Regression Analysis**, New York, Jon Wiley & Sons, 1982.
- Myers, Raymond H.: **Classical and Modern Regression with Applications**, second edition, Washington, Duxbury, 1990.
- Netter, John, William Wasserman, Michael H. Kutner: **Applied Linear Regression Models**, Illinois, Richard D. Irwin, Inc., 1983.
- Ngo, S.H., S. Kemeny, A. Deak: “Performance of the Ridge Regression Method as Applied to Complex Linear and Nonlinear Models”, **Chemometrics and Intelligent Laboratory Systems**, Vol. LXVII, 2003, s. 69-78.
- Orhunbilge, Neyran: **Zaman Serileri Analizi Tahmin ve Fiyat İndeksleri**, İstanbul, Avcıol Basım, 1999.
- Orhunbilge, Neyran: **Örnekleme Yöntemleri ve Hipotez Testleri**, 2. bs., İstanbul, Avcıol Basım, 2000.

- Orhunbilge, Neyran: **Tanımsal İstatistik Olasılık ve Olasılık Dağılımları**, İstanbul, Avcıol Basım Yayın, 2000.
- Orhunbilge, Neyran: **Uygulamalı Regresyon ve Korelasyon Analizi**, 2. bs., İstanbul, İ.Ü. İşletme Fakültesi, 2002.
- Öztürk, Fikri, Fikri Akdeniz: “Ill-Conditioning and Multicollinearity”, **Linear Algebra and Its Applications**, Vol. CCCXXI, 2000, s. 295-305.
- Park R. E.: “Estimation with Heteroscedastic Error Terms”, **Econometrica**, Vol. XXXIV, 1966, p. 888.
- Pesaran, Hashem, L.W. Taylor: “Diagnostics for IV Regressions”, **Oxford Bulletin of Economics and Statistics**, Vol.LXI, 1999, s. 255-281.
- Prais, S., C Winston: “Trend Estimation and Serial Correlation”, **Discussion Paper 383**, Cowles Commission, Chicago, 1954.
- Rousseeuw, Peter J.: “Least Median of Squares”, **Journal of the American Statistical Association**, Vol.79, 1984, s. 871-880.
- Rousseeuw, Peter J., Annick M. Leroy: **Robust Regression and Outlier Detection**, New York, Wiley-Interscience, 2003.
- Rousseeuw, Peter J., B. C. van Zomeren: “Unmasking Multivariate Outliers and Leverage Points”, **Journal of American Statistical Association**, Vol.LXXXV, 1990a, s. 648-651.
- Royston, J. P.: “An extension of Shapiro and Wilk’s W Test for Normality to Large Samples”, **Applied Statistics**, Vol. XXXI, 1982, s. 115-124.
- Ruppert, David: “Computing S-estimators for Regression and Multivariate Location/Dispersion”, **Journal of Computational and Graphical Statistics**, Vol. I, 1992, s. 253-270.
- Ryan, Thomas P.: **Modern Regression Methods**, New York, John Wiley & Sons, Inc., 1997.
- Savin, Eugene, Kenneth White: “The Durbin-Watson Test for Serial Correlation with Extreme Sample Sizes or Many Regressors”, **Econometrica**, Vol.XXIX, 1977, s.1989-1996.
- Schildt, Herbert: **Java 2**, Çev. Taner Akbaş, İstanbul, Alfa, 2001.

- Sen, Ashish, Muni Srivastva: **Regression Analysis Theory, Methods, and Applications**, New York, Springer-Verlag, 1990.
- Sen, Pranab Kumar, Jana Jureckova, Jan Picek: “Goodness-of-Fit Test of Shapiro-Wilk Type with Nuisance Regression and Scale”, **Austrian Journal of Statistics**, Vol. XXXII, 2003, s. 163-177.
- Shapiro, S. S., M. B. Wilk: “An Analysis of Variance Test for Normality (Complete Samples)”, **Biometrika**, Vol. LII, 1965, s. 591-611.
- Shi, Lei, Xueren Wang: “Local Influence in Ridge Regression”, **Computational Statistics & Data Analysis**, Vol. XXXI, 1999, s. 341-353.
- Sibson, Robert E.: **Maximizing Employee Productivity: A Manager’s Guide**, New York, Amacom, 1994.
- Stefanski, Leonard A.: “A Note on High-Breakdown Estimators”, **Statistics and Probability Letters**, Vol.XI, 1991, s. 353-358.
- Şahinler, Suat: “En Küçük Kareler Yöntemi ile Doğrusal Regresyon modeli Oluşturmanın Temel Prensipleri”, **MKÜ Ziraat Fakültesi Dergisi**, C.V, 2000, s. 57-63.
- Theil, Henri: **Principles of Econometrics**, New York, John Wiley & Sons, 1971.
- Theil, Henri, A. Nagar: “Testing the Independence of Regression Disturbances” **Journal of the American Statistical Association**”, Vol.LII, 1961, s. 793-806.
- Thursby, Jerry: “Misspecification, Heteroscedasticity, and the Chow and Goldfeld-Quandt Tests”, **Review of Economics and Statistics**, Vol. LXIV, 1982, s. 314-321.
- Waldman, Donald: “A Note on the Algebraic Equivalence of White’s Test and a Variant of the Godfrey/Breusch-Pagan Test for Heteroscedasticity”, **Econometric Letters**, Vol.XIII, 1983, s. 197-200.
- Wallantin, Bo, Anders Agren: “Test of Heteroscedasticity in a Regression Model in the Presence of Measurement Errors”, **Economic Letters**, Vol.LXXVI, 2002, s. 205-211.

- Webster, Thomas J.: “A Principal Component Analysis of the U.S. News & World Report Tier Rankings of Colleges and Universities”, **Economics of Education Review**, Vol. XX, 2001, s. 235-244.
- White, Hal: “A Heteroscedasticity- Consistance Covariance Matrix estimator and a Direct Test of Heteroscedasticity”, **Econometrica**”, Vol.XLVIII, 1980, s. 817-838.
- Wonnacott, Ronald J.,  
Thomas H. Wonnacott: **Econometrics**, New York, Wiley, 1979.
- “The Java Tutorial, Object-Oriented Programming Concepts”, (Çevrimiçi)  
<http://java.sun.com/docs/books/tutorial/concepts/index.html>, 6 Haziran 2005.
- “The Java Tutorial, Getting Started”, (Çevrimiçi)  
<http://java.sun.com/docs/books/tutorial/getStarted/index.html>, 6 Haziran 2005.
- “Verimlilik ve Maliyet Göstergeleri”, (Çevrimiçi)  
<http://www.canaktan.org/yeni-trendler/yeni-rekabet/olcme.htm>, 27 Temmuz 2005.



## EKLER

### Ek 1: (Ek:2)'deki Basit Doğrusal Regresyon Modeli Verilerinin (txt) Uzantılı Bir Dosyada Saklanmış Hali

32  
2  
0.764 0.01805  
0.801 0.020193  
0.948 0.019031  
1.171 0.017244  
1.563 0.021968  
2.086 0.027947  
3.778 0.038242  
4.739 0.047452  
5.129 0.053684  
5.796 0.061264  
4.599 0.066828  
5.069 0.081696  
7.909 0.068391  
8.933 0.071504  
8.843 0.064209  
9.235 0.060492  
10.757 0.059098  
11.343 0.066891  
11.105 0.075173  
14.158 0.085979  
14.335 0.09046  
15.792 0.107544  
22.302 0.150758  
21.047 0.150168  
22.871 0.158122  
29.428 0.178715  
23.27 0.132302  
35.709 0.170081  
42.734 0.183579  
48.559 0.360258  
45.935 0.649752  
40.687 0.930296

## Ek 2: Basit Regresyon Modeli Örnek Verileri 1

Y	X
0,764	0,01805
0,801	0,020193
0,948	0,019031
1,171	0,017244
1,563	0,021968
2,086	0,027947
3,778	0,038242
4,739	0,047452
5,129	0,053684
5,796	0,061264
4,599	0,066828
5,069	0,081696
7,909	0,068391
8,933	0,071504
8,843	0,064209
9,235	0,060492
10,757	0,059098
11,343	0,066891
11,105	0,075173
14,158	0,085979
14,335	0,09046
15,792	0,107544
22,302	0,150758
21,047	0,150168
22,871	0,158122
29,428	0,178715
23,27	0,132302
35,709	0,170081
42,734	0,183579
48,559	0,360258
45,935	0,649752
40,687	0,930296

**Kaynak:** Selahattin Güriş, Ebru Çağlayan, **Ekonometri Temel Kavramlar**, İstanbul, Der Yayınevi Yayın No:282, 2000, s. 441.

### Ek 3: Basit Regresyon Modeli Örnek Verileri 2

Y	X
68,7	65,5
70,7	68
72,2	70,4
75	73,2
77,9	76,4
79,6	78,5
82,9	80,7
84,9	82,7
88,2	85,3
89,7	85,7
91,2	86,9
93	89,8
95,8	92,6
98	95
97	93,2
97,7	95,4
100,8	98,2
102,3	99,8
103,4	100,4
102	99,3
99,5	98,6
98,7	99,9
100	100
100,5	102
100,4	104,6
101,3	106,1
104,4	108,3
104,3	109,4
104,4	110,4
103	109,5
103,2	109,7
103,9	110,1

**Kaynak:** Damodar N. Gujarati, **Temel Ekonometri**, Çev. Ümit Şenesen, Gülay Günlük Şenesen, İstanbul, Literatür, 1999, s. 451.

#### Ek 4: Çoklu Regresyon Modeli Verileri

Y	X1	X2	X3	X4
2,301	6,7	62	81	2,59
2,0043	5,1	59	66	1,7
2,3096	7,4	57	83	2,16
2,0043	6,5	73	41	2,01
2,7067	7,8	65	115	4,3
1,9031	5,8	38	72	1,42
1,9031	5,7	46	63	1,91
2,1038	3,7	68	81	2,57
2,3054	6	67	93	2,5
2,3075	3,7	76	94	2,4
2,5172	6,3	84	83	4,13
1,8129	6,7	51	43	1,86
2,9191	5,8	96	114	3,95
2,5185	5,8	83	88	3,95
2,2253	7,7	62	67	3,4
2,3365	7,4	74	68	2,4
1,9395	6	85	28	2,98
1,5315	3,7	51	41	1,55
2,3324	7,3	68	74	3,56
2,2355	5,6	57	87	3,02
2,0374	5,2	52	76	2,85
2,1335	3,4	83	53	1,12
1,8451	6,7	26	68	2,1
2,3424	5,8	67	86	3,4
2,4409	6,3	59	100	2,95
2,1584	5,8	61	73	3,5
2,2577	5,2	52	86	2,45
2,7589	11,2	76	90	5,59
1,8573	5,2	54	56	2,71
2,2504	5,8	76	59	2,58
1,8513	3,2	64	65	0,74
1,7634	8,7	45	23	2,52
2,0645	5	59	73	3,5
2,4698	5,8	72	93	3,3
2,0607	5,4	58	70	2,64
2,2648	5,3	51	99	2,6
2,0719	2,6	74	86	2,05
2,0792	4,3	8	119	2,85
2,179	4,8	61	76	2,45
2,1703	5,4	52	88	1,81
1,9777	5,2	49	72	1,84
1,8751	3,6	28	99	1,3
2,684	8,8	86	88	6,4
2,1847	6,5	56	77	2,85

2,281	3,4	77	93	1,48
2,0899	6,5	40	84	3
2,4928	4,5	73	106	3,05
2,5999	4,8	86	101	4,1
2,1987	5,1	67	77	2,86
2,4914	3,9	82	103	4,55
2,0934	6,6	77	46	1,95
2,0969	6,4	85	40	1,21
2,2967	6,4	59	85	2,33
2,4955	8,8	78	72	3,2

**Kaynak:** John Netter, William Wassermann, Michael H. Kutner, **Applied Linear Regression Models**, Illinois, Richard D. Irwin, Inc., 1983, p.420.

**Ek 5: İşgücü Verimliliği Örnek Veri Seti**

500 Büyük Firma Sıra No 2003	Firma ve Müesseseler	Sektör	Satış Hasılatı (Net) [1.000.000 TL]	Brüt Katma Değer [1.000.000 TL]	Makine, Tesis ve Cihazlar
9	VESTEL ELEKTRONİK SAN. VE TİC. A.Ş.	elek-elek	1.712.574.277	261.733.781	286.702.551
23	BSH PROFİLO ELEKTRİKLİ GEREÇLER SANAYİİ A.Ş.	Elek- Elektro	986.818.613	264.293.850	194.202.233
48	GOODYEAR LASTİKLERİ T.A.Ş.	lastik	446.207.773	115.133.452	229.506.702
49	ÜLKER GIDA SAN. VE TİC. A.Ş.	Gıda	411.384.086	76.545.569	94.687.799
56	AKENERJİ ELEKTRİK ÜRETİMİ OTOPRODÜKTÖR GRUBU A.Ş.	Enerji- Petrol	351.645.209	73.273.374	268.622.634
69	KARDEMİR KARABÜK DEMİR ÇELİK SAN. VE TİC. A.Ş.	demir-çelik	317.586.274	162.289.628	164.798.493
70	SARKUYSAN ELEKTROLİTİK BAKIR SAN. VE TİC. A.Ş.	metal	318.079.023	28.143.261	75.134.735
72	İZMİR DEMİR ÇELİK SANAYİ A.Ş.	demir-çelik	313.975.980	2.476.693	98.847.680
79	ASELSAN ELEKTRONİK SANAYİ VE TİCARET A.Ş.	Elek- Elektro	306.154.847	123.704.022	169.414.245
103	ECZACIBAŞI YAPI GEREÇLERİ SAN. VE TİC. A.Ş.	İnsaat	227.795.730	89.564.893	152.039.818
111	MENDERES TEKSTİL SAN. VE TİC. A.Ş.	Tekstil	208.552.686	50.689.107	86.008.638
133	ALTINYILDIZ MENSUCAT VE KONFEKSİYON FABRİKALARI A.Ş.	Tekstil	168.808.498	36.936.699	40.140.856
141	GOLDAŞ KUYUMCULUK SANAYİ İTHALAT İHRACAT ANONİM ŞİRKETİ	kuyum	245.495.459	1.486.091	1.162.946
149	AKIN TEKSTİL A.Ş.	Tekstil	159.567.264	36.248.193	49.301.556
160	ADANA ÇİMENTO SANAYİİ T.A.Ş.	Çimento	149.164.258	40.753.051	117.763.910
164	BATIÇİM BATI ANADOLU ÇİMENTO SANAYİİ A.Ş.	Çimento	163.634.945	29.665.404	16.767.218
178	NUH ÇİMENTO SANAYİ A.Ş.	Çimento	137.399.644	58.764.381	148.911.298

Kısa Vadeli Borçlar	Uzun Vadeli Borçlar	Toplam Borçlar	Net Aktifler Tutarı [1.000.000 TL]	Dönem Karı [V.Ö.] Tutarı [1.000.000 TL]	İhracat Tutarı [1.000 \$]	Ücretle Çalışanlar Ortalaması	Y=BKD/Çalışan Ort.	X1=Satış
821.132.073	571.653.165	1.392.785.238	1.929.043.514	140.631.726	1.036.517	3.397	77048,5078	1712574277
258.880.434	135.820.356	394.700.790	549.571.390	75.294.801	182.283	2.334	113236,4396	986818613
116.873.153	26.639.985	143.513.138	233.439.646	13.822.521	182.165	1.492	77167,19303	446207773
160.812.795	26.816.802	187.629.597	426.352.291	36.858.207	78.029	1.888	40543,20392	411384086
68.875.061	22.057.592	90.932.653	448.686.377	35.269.025	0	284	258004,838	351645209
148.926.785	161.536.360	310.463.145	355.764.473	85.663.492	6.221	3.678	44124,42306	317586274
46.853.779	8.583.833	55.437.612	132.277.768	7.153.520	125.011	543	51829,20994	318079023
107.227.452	24.063.393	131.290.845	298.746.921	-16.217.916	176.449	676	3663,747041	313975980
155.882.506	340.258.866	496.141.372	678.850.580	45.614.636	32.795	2.923	42320,91071	306154847
67.841.316	28.554.941	96.396.257	249.029.870	5.487.922	123.369	3.531	14355,4537	208552686
126.321.039	7.272.637	133.593.676	180.892.053	2.020.299	26.156	1.566	23586,65326	168808498
29.942.161	3.302.183	33.244.344	140.055.223	8.322.663	17.931	257	5782,455253	245495459
33.114.686	22.588.799	55.703.485	112.997.754	14.274.438	77.000	1.688	21474,04799	159567264
28.580.683	5.725.745	34.306.428	241.539.930	32.658.399	33.318	364	111958,9313	149164258
26.277.632	4.418.194	30.695.826	126.410.876	25.950.791	33.514	352	84276,71591	163634945
24.308.599	20.731.714	45.040.313	187.713.386	26.626.291	30.853	433	135714,5058	137399644

<b>X2=Mak. Tes. ve Cih./Çalışan</b>	<b>X3=Kar/Satış</b>	<b>X4=Toplam Borçlar/ Aktifler</b>	<b>X5=İhracat/ Satış</b>	<b>X6=Yabancı Sermaye Oranı</b>	<b>X7=Sektör</b>
84398,74919	0,08212	0,72201	0,00085	51,59	1
83205,75536	0,07630	0,71820	0,00026	93,86	1
153824,8673	0,03098	0,61478	0,00057	74,6	1
50152,43591	0,08960	0,44008	0,00027	33	1
945854,3451	0,10030	0,20266	0,00000	0	0
44806,55057	0,26973	0,87266	0,00003	0	0
138369,6777	0,02249	0,41910	0,00055	0	0
146224,3787	-0,05165	0,43947	0,00079	5,15	0
57959,03011	0,14899	0,73086	0,00015	0	1
86879,896	0,02052	0,38995	0,00069	0,71	0
24358,15293	0,02631	0,38709	0,00083	0	1
25632,73052	0,01197	0,73853	0,00022	0	1
4525,081712	0,03390	0,23737	0,00010	0	1
29207,08294	0,08946	0,49296	0,00068	0	1
323527,2253	0,21894	0,14203	0,00031	0	0
47634,14205	0,15859	0,24283	0,00029	0	0
343906	0,19379	0,23994	0,00031	0	0



500 Büyük Firma Sıra No 2003	Firma ve Müesseseler	Sektör	Satış Hasılatı (Net) [1.000.000 TL]	Brüt Katma Değer [1.000.000 TL]	Makine, Tesis ve Cihazlar
192	AK-AL TEKSTİL SANAYİİ A.Ş.	Tekstil	129.516.524	18.463.839	52.921.049
198	TUKAŞ GIDA SAN. VE TİC. A.Ş.	Gıda	128.627.064	11.439.262	15.036.133
202	DEVA HOLDİNG A.Ş.	Kimya-İlaç	139.975.140	63.619.172	20.625.907
211	BAGFAŞ BANDIRMA GÜBRE FABRİKALARI A.Ş.	Kimya-İlaç	148.608.387	7.710.957	41.017.081
213	KÜTAHYA PORSELEN SAN. A.Ş.	cam- seramik	119.333.409	39.897.223	57.988.245
215	KARTONSAN KARTON SAN. VE TİC. A.Ş.	Ağaç- orman	115.509.528	42.357.813	90.179.253
218	TİRE KUTSAN OLUKLU MUKAVVA KUTU VE KAĞIT SANAYİ A.Ş.	Ağaç- orman	113.588.580	32.764.753	89.471.537
225	MUTLU AKÜ VE MALZEMELERİ SAN. A.Ş.	makine	112.061.317	22.932.094	51.639.462
233	MENSA MENSUCAT SAN. VE TİC. A.Ş.	Tekstil	112.621.035	6.808.307	112.315.345
236	LİO YAĞ SAN. VE TİC. A.Ş.	Gıda	119.097.679	6.639.476	12.087.147
238	EGE SERAMİK SANAYİ VE TİCARET A.Ş.	cam- seramik	113.966.857	31.561.317	38.194.028
245	BOLU ÇİMENTO SANAYİİ A.Ş.	Çimento	104.339.155	28.791.170	82.659.905
251	SİFAŞ SENTETİK İPLİK FABRİKALARI A.Ş.	Tekstil	104.032.493	-7.007.190	49.327.039
273	GÜBRE FABRİKALARI T.A.Ş.	Kimya-İlaç	253.961.729	25.360.200	20.340.173
287	BERDAN TEKSTİL SAN. VE TİC. A.Ş.	Tekstil	103.369.155	34.946.341	115.975.327
297	KARSU TEKSTİL SANAYİİ VE TİC. A.Ş.	Tekstil	88.470.262	33.340.142	61.486.793
300	ŞEKER PİLİÇ VE YEM SANAYİ TİCARET A.Ş.	Gıda	88.791.703	14.233.813	12.427.509
310	YATAŞ YATAK VE YORGAN SAN. TİC. A.Ş.	Tekstil	87.998.258	23.327.562	8.104.125
311	ALARKO CARRIER SANAYİ VE TİCARET ANONİM ŞİRKETİ	Elek-Elek	166.272.647	51.345.145	13.213.904
313	ARSAN TEKSTİL TİC. VE SAN. A.Ş.	Tekstil	82.436.245	19.971.814	100.123.168
314	UÇAK SERVİSİ A.Ş.		129.749.350	80.907.902	11.321.321
321	AKSU İPLİK DOKUMA VE BOYA APRE FAB. T.A.Ş.	Tekstil	81.942.835	28.459.735	29.459.209
346	BURSA ÇİMENTO FABRİKASI A.Ş.	Çimento	74.559.305	22.134.799	46.809.481

Kısa Vadeli Borçlar	Uzun Vadeli Borçlar	Toplam Borçlar	Net Aktifler Tutarı [1.000.000 TL]	Dönem Karı [V.Ö.] Tutarı [1.000.000 TL]	İhracat Tutarı [1.000 \$]	Ücretle Çalışanlar Ortalaması	Y=BKD/Çalışan Ort.	X1=Satış
49.150.321	6.833.558	55.983.879	127.345.128	101.832	34.106	1.363	13546,47029	129516524
74.992.047	11.331.945	86.323.992	104.603.087	-3.384.735	31.248	447	25591,19016	128627064
81.355.648	23.307.758	104.663.406	146.771.468	17.155.846	7.282	1.045	60879,59043	139975140
45.373.122	4.690.048	50.063.170	106.104.294	4.057.482	6.188	346	22286,00289	148608387
17.694.354	3.911.087	21.605.441	76.230.544	15.180.721	21.930	760	52496,34605	119333409
15.453.619	4.460.685	19.914.304	134.134.551	21.602.548	19.803	288	147075,7396	115509528
16.022.974	18.980.201	35.003.175	93.783.145	5.033.284	2.648	821	39908,34714	113588580
20.600.751	14.120.366	34.721.117	68.342.710	4.884.067	24.004	479	47874,93528	112061317
75.298.111	50.183.621	125.481.732	167.459.216	-1.343.583	45.708	1.100	6189,37	112621035
41.903.815	4.841.851	46.745.666	88.142.815	2.178.133	48.288	135	49181,3037	119097679
65.816.555	29.080.364	94.896.919	113.860.010	2.690.348	34.976	1.047	30144,52436	113966857
27.957.132	5.255.663	33.212.795	137.404.939	24.207.554	0	256	112465,5078	104339155
55.228.086	82.460.865	137.688.951	132.741.433	-4.688.643	8.717	1.298	-5398,451464	104032493
44.336.348	9.237.000	53.573.348	88.291.634	11.055.382	1.874	426	59530,98592	253961729
57.772.288	45.440.157	103.212.445	94.861.665	1.779.881	47.740	928	37657,69504	103369155
34.031.989	9.037.034	43.069.023	80.272.655	1.988.084	13.473	832	40072,28606	88470262
14.529.751	12.098.295	26.628.046	41.263.013	2.928.331	583	827	17211,3821	88791703
53.134.934	5.474.779	58.609.713	80.745.306	6.485.792	12.750	480	48599,0875	87998258
35.241.142	6.087.256	41.328.398	133.410.098	23.768.217	16.694	756	67916,85847	166272647
16.023.679	20.693.497	36.717.176	103.191.014	-4.741.253	15.060	1.117	17879,86929	82436245
18.762.580	7.230.278	25.992.858	67.300.625	26.768.423	0	1.134	71347,35626	129749350
25.160.767	8.435.950	33.596.717	65.880.645	-1.840.496	21.665	1.045	27234,19617	81942835
14.577.493	3.579.186	18.156.679	68.013.969	9.803.995	18.758	267	82901,86891	74559305

<b>X2=Mak. Tes. ve Cih./Çalışan</b>	<b>X3=Kar/Satış</b>	<b>X4=Toplam Borçlar/ Aktifler</b>	<b>X5=İhracat/ Satış</b>	<b>X6=Yaban cı Sermaye Oranı</b>	<b>X7=Sektör</b>
38826,88848	0,00079	0,43962	0,00037	0	1
33637,88143	-0,02631	0,82525	0,00034	0	1
19737,71005	0,12256	0,71310	0,00007	0	1
118546,4769	0,02730	0,47183	0,00006	0	1
76300,32237	0,12721	0,28342	0,00026	0	1
313122,4063	0,18702	0,14847	0,00024	0	1
108978,7296	0,04431	0,37324	0,00003	0	1
107806,81	0,04358	0,50804	0,00030	0	1
102104,8591	-0,01193	0,74933	0,00057	0	1
89534,42222	0,01829	0,53034	0,00057	0	1
36479,49188	0,02361	0,83345	0,00043	0	1
322890,2539	0,23201	0,24171	0,00000	0	0
38002,34129	-0,04507	1,03727	0,00012	0	1
47746,88498	0,04353	0,60678	0,00001	0	1
124973,4127	0,01722	1,08803	0,00065	0	1
73902,39543	0,02247	0,53653	0,00021	0	1
15027,21765	0,03298	0,64532	0,00001	0	1
16883,59375	0,07370	0,72586	0,00020	0	1
17478,70899	0,14295	0,30978	0,00014	42,03	1
89635,78156	-0,05751	0,35582	0,00026	0	1
9983,528219	0,20631	0,38622	0,00000	58,75	0
28190,63062	-0,02246	0,50996	0,00037	0	1
175316,4082	0,13149	0,26696	0,00035	0	0

500 Büyük Firma Sıra No 2003	Firma ve Müesseseler	Sektör	Satış Hasılatı (Net) [1.000.000 TL]	Brüt Katma Değer [1.000.000 TL]	Makine, Tesis ve Cihazlar
357	ESCORT COMPUTER ELEKTRONİK SAN. VE TİC. A.Ş.	Bilişim	72.845.171	2.051.374	11.198
360	ÜNYE ÇİMENTO SANAYİİ VE TİCARET A.Ş.	Çimento	71.804.201	6.981.457	186.876.648
369	SÖNMEZ FİLAMANT SENTETİK İPLİK VE ELYAF SANAYİİ A.Ş.	Tekstil	69.705.305	6.856.944	62.594.134
371	BAK AMBALAJ SAN. VE TİC. A.Ş.	Ambalaj	74.477.262	10.015.118	22.768.845
377	LAFARGE ASLAN ÇİMENTO A.Ş.	Çimento	70.667.137	28.206.483	43.780.836
391	SÖNMEZ PAMUKLU SANAYİİ A.Ş.	Tekstil	66.193.773	13.897.263	76.150.460
404	DARDANEL ÖNENTAŞ GIDA SAN. A.Ş.	Gıda	64.249.179	7.530.796	10.645.024
416	ERBOSAN ERCİYAS BORU SAN. VE TİC. A.Ş.	metal	62.080.053	3.357.640	12.065.913
430	MERKO GIDA SAN VE TİC A.Ş	Gıda	94.000.475	6.814.745	76.145.373
431	MARDİN ÇİMENTO SANAYİİ VE TİCARET A.Ş.	Çimento	60.682.341	22.302.492	22.400.895
436	DENTAŞ AMBALAJ VE KAĞIT SANAYİ A.Ş.	Ambalaj	60.216.551	15.324.591	50.043.388
441	EDİP İPLİK SAN. VE TİC. A.Ş.	Tekstil	65.695.801	31.313.711	56.241.459
444	ALKİM KAĞIT SAN. VE TİC. A.Ş.	Ağaç- orman	58.613.149	13.488.542	63.468.112
463	EGEPLAST EGE PLASTİK TİCARET VE SANAYİ A.Ş.	Plastik	61.698.487	22.200.832	10.134.267
464	EGE GÜBRE SANAYİİ A.Ş.	Kimya-ilaç	94.099.670	6.103.061	4.739.147
470	PENGUEN GIDA SANAYİ A.Ş.	Gıda	57.264.717	1.989.489	12.506.128
471	POLYLEN SENTETİK İPLİK SANAYİİ A.Ş.	Tekstil	58.419.563	3.652.434	50.762.695
475	BATISÖKE SÖKE ÇİMENTO SAN. T.A.Ş.	Çimento	54.697.308	14.365.823	69.636.262
483	KLİMASAN KLİMA SAN. VE TİC. A.Ş.	elekt-elektr	55.653.378	11.798.750	13.165.673
496	BOSCH FREN SİSTEMLERİ SAN. VE TİC. A.Ş.	Otomotiv	52.413.856	8.115.983	23.550.133

Kısa Vadeli Borçlar	Uzun Vadeli Borçlar	Toplam Borçlar	Net Aktifler Tutarı [1.000.000 TL]	Dönem Karı [V.Ö.] Tutarı [1.000.000 TL]	İhracat Tutarı [1.000 \$]	Ücretle Çalışanlar Ortalaması	Y=BKD/Çalışan Ort.	X1=Satış
14.058.083	136.729	14.194.812	26.901.403	885.613	20.438	78	26299,66667	72845171
22.628.668	59.598.986	82.227.654	158.232.623	-24.692.809	7.388	188	37135,40957	71804201
148.805.520	11.894.967	160.700.487	73.698.061	-5.915.045	2.954	607	11296,44811	69705305
17.345.425	1.268.846	18.614.271	37.145.749	87.462	25.902	300	33383,72667	74477262
22.461.781	16.409.664	38.871.445	72.703.404	10.729.019	12.490	248	113735,8185	70667137
7.342.669	2.663.597	10.006.266	77.249.372	2.228.560	6.421	789	17613,76806	66193773
39.121.197	154.629.902	193.751.099	51.179.232	27.592.348	23.550	867	8686,039216	64249179
13.123.215	893.847	14.017.062	29.468.106	3.214.865	20.129	346	9704,16185	62080053
89.882.852	4.735.651	94.618.503	147.465.118	2.124.343	45.787	249	27368,45382	94000475
10.347.103	2.469.777	12.816.880	70.886.250	17.497.334	20.146	251	88854,5498	60682341
5.940.138	15.782.615	21.722.753	61.321.208	2.552.693	1.118	435	35228,94483	60216551
18.118.182	6.152.337	24.270.519	77.069.096	2.541.472	9.785	1.107	28287,00181	65695801
16.690.203	5.203.269	21.893.472	91.705.445	4.744.764	13.308	226	59683,81416	58613149
26.879.450	49.943.737	76.823.187	79.437.237	8.840.943	3.342	344	64537,30233	61698487
41.577.243	1.647.123	43.224.366	57.173.196	3.724.425	3.369	174	35075,06322	94099670
45.269.580	9.621.314	54.890.894	53.191.613	-10.479.465	24.519	308	6459,37987	57264717
42.659.678	57.838.764	100.498.442	103.122.183	-4.962.265	0	791	4617,489254	58419563
11.350.224	2.179.534	13.529.758	102.405.331	4.455.629	11.216	273	52622,06227	54697308
15.989.981	569.226	16.559.207	31.401.271	1.953.465	15.302	463	25483,26134	55653378
8.233.473	3.077.967	11.311.440	29.490.314	559.356	11.349	184	44108,60326	52413856

<b>X2=Mak. Tes. ve Cih./Çalışan</b>	<b>X3=Kar/Satı ş</b>	<b>X4=Topla m Borçlar/ Aktifler</b>	<b>X5=İhracat/ Satış</b>	<b>X6=Yabancı Sermaye Oranı</b>	<b>X7=Sektör</b>
143,5641026	0,01216	0,52766	0,00039	0	0
994024,7234	-0,34389	0,51966	0,00014	0	0
103120,4843	-0,08486	2,18053	0,00006	0	1
75896,15	0,00117	0,50111	0,00049	0	1
176535,629	0,15182	0,53466	0,00025	97	0
96515,15843	0,03367	0,12953	0,00014	0	1
12277,99769	0,42946	3,78574	0,00051	0	1
34872,58092	0,05179	0,47567	0,00045	0	0
305804,7108	0,02260	0,64163	0,00068	40,75	1
89246,59363	0,28834	0,18081	0,00046	0	0
115042,2713	0,04239	0,35425	0,00003	0	1
50805,29268	0,03869	0,31492	0,00021	0	1
280832,354	0,08095	0,23874	0,00032	0	1
29460,07849	0,14329	0,96709	0,00008	0	0
27236,47701	0,03958	0,75603	0,00005	0	1
40604,31169	-0,18300	1,03195	0,00060	14,65	1
64175,34134	-0,08494	0,97456	0,00000	0	1
255077,8828	0,08146	0,13212	0,00029	0	0
28435,57883	0,03510	0,52734	0,00038	0	1
127989,8533	0,01067	0,38356	0,00030	84,49	1

**Ek 6: (Ek:2)'deki Basit Doğrusal Regresyon Modelinde Basit Korelasyon Katsayıları**

**Correlations**

		Y	X
Y	Pearson Correlation	1	,761**
	Sig. (2-tailed)		,000
	N	32	32
X	Pearson Correlation	,761**	1
	Sig. (2-tailed)	,000	
	N	32	32

\*\* . Correlation is significant at the 0.01 level

## Ek 7: Çoklu Regresyon Modelinde Basit Korelasyon Katsayıları

Correlations

		Y	X1	X2	X3	X4
Y	Pearson Correlation	1	,346*	,593**	,665**	,726**
	Sig. (2-tailed)		,010	,000	,000	,000
	N	54	54	54	54	54
X1	Pearson Correlation	,346*	1	,090	-,150	,502**
	Sig. (2-tailed)	,010		,517	,280	,000
	N	54	54	54	54	54
X2	Pearson Correlation	,593**	,090	1	-,024	,369**
	Sig. (2-tailed)	,000	,517		,865	,006
	N	54	54	54	54	54
X3	Pearson Correlation	,665**	-,150	-,024	1	,416**
	Sig. (2-tailed)	,000	,280	,865		,002
	N	54	54	54	54	54
X4	Pearson Correlation	,726**	,502**	,369**	,416**	1
	Sig. (2-tailed)	,000	,000	,006	,002	
	N	54	54	54	54	54

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\*. Correlation is significant at the 0.01 level (2-tailed).



## Ek 8: (Ek:2)'deki Veriler ile Basit Doğrusal Regresyon Analizi SPSS Sonuçları

### Regression

#### Variables Entered/Removed<sup>a</sup>

Model	Variables Entered	Variables Removed	Method
1	X <sup>a</sup>	.	Enter

a. All requested variables entered.

b. Dependent Variable: Y

#### Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,761 <sup>a</sup>	,578	,564	9,436567

a. Predictors: (Constant), X

#### ANOVA<sup>b</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3664,951	1	3664,951	41,157	,000 <sup>a</sup>
	Residual	2671,464	30	89,049		
	Total	6336,415	31			

a. Predictors: (Constant), X

b. Dependent Variable: Y

#### Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	7,337	2,056		3,569	,001
	X	57,517	8,966	,761	6,415	,000

a. Dependent Variable: Y

**Ek 9: (Ek:2)'deki Veriler ile İkinci Derece Eğrisel Regresyon Analizi SPSS Sonuçları**

**Regression**

**Variables Entered/Removed<sup>a</sup>**

Model	Variables Entered	Variables Removed	Method
1	sqr <sub>x</sub> , X <sup>a</sup>	.	Enter

- a. All requested variables entered.  
b. Dependent Variable: Y

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,956 <sup>a</sup>	,914	,908	4,337050

- a. Predictors: (Constant), sqr<sub>x</sub>, X

**ANOVA<sup>b</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5790,925	2	2895,463	153,932	,000 <sup>a</sup>
	Residual	545,490	29	18,810		
	Total	6336,415	31			

- a. Predictors: (Constant), sqr<sub>x</sub>, X  
b. Dependent Variable: Y

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-3,340	1,379		-2,422	,022
	X	206,193	14,579	2,726	14,143	,000
	sqr <sub>x</sub>	-175,804	16,537	-2,049	-10,631	,000

- a. Dependent Variable: Y

**Ek 10: (Ek:2)'deki Veriler için Bağımsız Değişkene Uygulanan Logaritmik Dönüşüm ile Regresyon Analizi SPSS Sonuçları**

**Regression**

**Variables Entered/Removed<sup>b</sup>**

Model	Variables Entered	Variables Removed	Method
1	lnx <sup>a</sup>	.	Enter

a. All requested variables entered.

b. Dependent Variable: Y

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,908 <sup>a</sup>	,824	,818	6,091529

a. Predictors: (Constant), lnx

**ANOVA<sup>b</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5223,213	1	5223,213	140,762	,000 <sup>a</sup>
	Residual	1113,202	30	37,107		
	Total	6336,415	31			

a. Predictors: (Constant), lnx

b. Dependent Variable: Y

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	49,096	3,066		16,016	,000
	lnx	13,449	1,134	,908	11,864	,000

a. Dependent Variable: Y

## Ek 11: Çoklu Regresyon Modelinin “Enter” Yöntemiyle SPSS Çıktısı

### Regression

#### Variables Entered/Removed<sup>b</sup>

Model	Variables Entered	Variables Removed	Method
1	X4, <sup>a</sup> X2, X3, X1	.	Enter

a. All requested variables entered.

b. Dependent Variable: Y

#### Model Summary<sup>b</sup>

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,986 <sup>a</sup>	,972	,970	,04733	2,146

a. Predictors: (Constant), X4, X2, X3, X1

b. Dependent Variable: Y

#### ANOVA<sup>b</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3,863	4	,966	431,097	,000 <sup>a</sup>
	Residual	,110	49	,002		
	Total	3,973	53			

a. Predictors: (Constant), X4, X2, X3, X1

b. Dependent Variable: Y

#### Coefficients<sup>b</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	,489	,050		9,730	,000		
	X1	,069	,005	,401	12,596	,000	,556	1,799
	X2	,009	,000	,571	21,189	,000	,776	1,289
	X3	,009	,000	,736	23,910	,000	,596	1,678
	X4	,002	,010	,008	,198	,844	,391	2,555

a. Dependent Variable: Y

## Ek 12: Çoklu Regresyon Modelinin “Stepwise” Yöntemiyle SPSS Çıktısı

### Regression

Variables Entered/Removed<sup>a</sup>

Model	Variables Entered	Variables Removed	Method
1	X4	.	Stepwise (Criteria: F-to-enter >= 4,000, F-to-remo ve <= 3,990).
2	X3	.	Stepwise (Criteria: F-to-enter >= 4,000, F-to-remo ve <= 3,990).
3	X2	.	Stepwise (Criteria: F-to-enter >= 4,000, F-to-remo ve <= 3,990).
4	X1	.	Stepwise (Criteria: F-to-enter >= 4,000, F-to-remo ve <= 3,990).
5	.	X4	Stepwise (Criteria: F-to-enter >= 4,000, F-to-remo ve <= 3,990).

a. Dependent Variable: Y

**Model Summary<sup>f</sup>**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,726 <sup>a</sup>	,527	,518	,19002	
2	,829 <sup>b</sup>	,687	,674	,15626	
3	,940 <sup>c</sup>	,883	,876	,09646	
4	,986 <sup>d</sup>	,972	,970	,04733	
5	,986 <sup>e</sup>	,972	,971	,04687	2,141

- a. Predictors: (Constant), X4
- b. Predictors: (Constant), X4, X3
- c. Predictors: (Constant), X4, X3, X2
- d. Predictors: (Constant), X4, X3, X2, X1
- e. Predictors: (Constant), X3, X2, X1
- f. Dependent Variable: Y

**ANOVA<sup>f</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2,095	1	2,095	58,024	,000 <sup>a</sup>
	Residual	1,878	52	,036		
	Total	3,973	53			
2	Regression	2,727	2	1,364	55,849	,000 <sup>b</sup>
	Residual	1,245	51	,024		
	Total	3,973	53			
3	Regression	3,508	3	1,169	125,663	,000 <sup>c</sup>
	Residual	,465	50	,009		
	Total	3,973	53			
4	Regression	3,863	4	,966	431,097	,000 <sup>d</sup>
	Residual	,110	49	,002		
	Total	3,973	53			
5	Regression	3,863	3	1,288	586,043	,000 <sup>e</sup>
	Residual	,110	50	,002		
	Total	3,973	53			

- a. Predictors: (Constant), X4
- b. Predictors: (Constant), X4, X3
- c. Predictors: (Constant), X4, X3, X2
- d. Predictors: (Constant), X4, X3, X2, X1
- e. Predictors: (Constant), X3, X2, X1
- f. Dependent Variable: Y

**Coefficients**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	1,696	,072		23,645	,000		
	X4	,186	,024	,726	7,617	,000	1,000	1,000
2	(Constant)	1,389	,084		16,441	,000		
	X4	,139	,022	,543	6,303	,000	,827	1,210
	X3	,006	,001	,439	5,089	,000	,827	1,210
3	(Constant)	,942	,071		13,199	,000		
	X4	,082	,015	,320	5,464	,000	,683	1,464
	X3	,007	,001	,543	9,981	,000	,790	1,265
	X2	,008	,001	,488	9,157	,000	,826	1,211
4	(Constant)	,489	,050		9,730	,000		
	X4	,002	,010	,008	,198	,844	,391	2,555
	X3	,009	,000	,736	23,910	,000	,596	1,678
	X2	,009	,000	,571	21,189	,000	,776	1,289
	X1	,069	,005	,401	12,596	,000	,556	1,799
5	(Constant)	,484	,043		11,345	,000		
	X3	,010	,000	,739	31,082	,000	,978	1,023
	X2	,009	,000	,574	24,299	,000	,992	1,008
	X1	,069	,004	,405	16,975	,000	,970	1,031

a. Dependent Variable: Y

**Ek 13: Çoklu Regresyon Modeli için SPSS’de “Ridge Regression” Modülünün Çalıştırılması Sonucunda Ulaşılan Rapor**

▽  
**Report**

R-SQUARE AND BETA COEFFICIENTS FOR ESTIMATED VALUES OF K

K	RSQ	X1	X2	X3	X4
,00000	,97237	,401192	,571318	,735505	,007527
,01000	,97213	,386325	,559720	,718009	,026300
,02000	,97147	,372779	,548931	,701823	,043225
,03000	,97049	,360387	,538854	,686787	,058540
,04000	,96925	,349007	,529409	,672766	,072440
,05000	,96781	,338523	,520525	,659647	,085092
,06000	,96622	,328831	,512145	,647333	,096638
,07000	,96449	,319845	,504217	,635741	,107200
,08000	,96266	,311491	,496698	,624800	,116881
,09000	,96075	,303705	,489550	,614447	,125773
,10000	,95877	,296429	,482741	,604628	,133954
,11000	,95674	,289616	,476240	,595296	,141493
,12000	,95467	,283222	,470022	,586408	,148451
,13000	,95256	,277209	,464065	,577929	,154882
,14000	,95043	,271543	,458349	,569824	,160831
,15000	,94827	,266195	,452855	,562065	,166341
,16000	,94610	,261139	,447567	,554626	,171450
,17000	,94392	,256351	,442471	,547483	,176190
,18000	,94172	,251809	,437554	,540616	,180592
,19000	,93952	,247494	,432805	,534004	,184683
,20000	,93732	,243390	,428211	,527632	,188486
,21000	,93511	,239480	,423765	,521484	,192024
,22000	,93290	,235751	,419456	,515545	,195316
,23000	,93068	,232190	,415277	,509803	,198380
,24000	,92847	,228786	,411221	,504246	,201233
,25000	,92626	,225528	,407280	,498862	,203890
,26000	,92405	,222406	,403448	,493643	,206364
,27000	,92184	,219411	,399721	,488578	,208667
,28000	,91963	,216536	,396091	,483661	,210812
,29000	,91743	,213772	,392555	,478882	,212808
,30000	,91523	,211114	,389108	,474234	,214665
,31000	,91303	,208555	,385746	,469712	,216393
,32000	,91084	,206088	,382464	,465309	,217999
,33000	,90865	,203710	,379260	,461019	,219491
,34000	,90646	,201414	,376129	,456836	,220876
,35000	,90428	,199196	,373068	,452757	,222160
,36000	,90211	,197052	,370075	,448776	,223349
,37000	,89993	,194978	,367147	,444888	,224450
,38000	,89777	,192971	,364281	,441091	,225467
,39000	,89560	,191026	,361474	,437380	,226405
,40000	,89344	,189141	,358725	,433751	,227268
,41000	,89129	,187312	,356030	,430201	,228061
,42000	,88914	,185537	,353389	,426728	,228788
,43000	,88699	,183814	,350799	,423327	,229453
,44000	,88485	,182139	,348258	,419996	,230058



,45000	,88272	,180511	,345764	,416734	,230607
,46000	,88059	,178927	,343317	,413536	,231103
,47000	,87846	,177386	,340914	,410401	,231549
,48000	,87634	,175885	,338553	,407326	,231948
,49000	,87423	,174422	,336234	,404310	,232302
,50000	,87212	,172997	,333955	,401350	,232613
,51000	,87001	,171607	,331715	,398445	,232884
,52000	,86791	,170251	,329512	,395592	,233116
,53000	,86582	,168927	,327346	,392790	,233312
,54000	,86373	,167635	,325215	,390038	,233474
,55000	,86165	,166373	,323118	,387333	,233602
,56000	,85957	,165139	,321055	,384675	,233700
,57000	,85750	,163933	,319024	,382061	,233768
,58000	,85543	,162753	,317024	,379491	,233808
,59000	,85337	,161599	,315055	,376962	,233821
,60000	,85131	,160469	,313116	,374475	,233809
,61000	,84926	,159364	,311205	,372028	,233773
,62000	,84721	,158280	,309323	,369619	,233714
,63000	,84517	,157219	,307468	,367248	,233633
,64000	,84314	,156179	,305640	,364913	,233531
,65000	,84111	,155160	,303837	,362614	,233410
,66000	,83909	,154160	,302060	,360350	,233269
,67000	,83707	,153179	,300308	,358119	,233111
,68000	,83506	,152216	,298579	,355920	,232936
,69000	,83305	,151272	,296875	,353754	,232745
,70000	,83105	,150344	,295193	,351619	,232538
,71000	,82906	,149433	,293533	,349514	,232316
,72000	,82707	,148539	,291896	,347439	,232081
,73000	,82509	,147659	,290279	,345392	,231832
,74000	,82311	,146795	,288684	,343374	,231570
,75000	,82114	,145946	,287109	,341383	,231297
,76000	,81917	,145111	,285554	,339419	,231012
,77000	,81722	,144289	,284018	,337481	,230716
,78000	,81526	,143481	,282501	,335568	,230410
,79000	,81332	,142686	,281003	,333681	,230094
,80000	,81137	,141903	,279524	,331818	,229768
,81000	,80944	,141133	,278062	,329979	,229434
,82000	,80751	,140374	,276618	,328163	,229091
,83000	,80559	,139627	,275191	,326370	,228740
,84000	,80367	,138891	,273780	,324600	,228381
,85000	,80176	,138166	,272386	,322851	,228015
,86000	,79985	,137452	,271009	,321124	,227642
,87000	,79796	,136748	,269647	,319418	,227263
,88000	,79606	,136054	,268300	,317732	,226877
,89000	,79418	,135370	,266969	,316066	,226486
,90000	,79230	,134695	,265653	,314420	,226088
,91000	,79042	,134030	,264351	,312794	,225686
,92000	,78856	,133374	,263063	,311186	,225278
,93000	,78669	,132726	,261790	,309597	,224866
,94000	,78484	,132088	,260530	,308026	,224449
,95000	,78299	,131457	,259284	,306473	,224028
,96000	,78114	,130835	,258052	,304937	,223603
,97000	,77931	,130220	,256832	,303419	,223174
,98000	,77748	,129614	,255625	,301917	,222741
,99000	,77565	,129015	,254431	,300432	,222306
1,0000	,77383	,128424	,253249	,298963	,221866

▽

**Ek 14: (Ek 3)'deki Basit Doğrusal Regresyon Modelinin "Enter" yöntemiyle SPSS Çıktısı**

**Regression**

**Variables Entered/Removed<sup>b</sup>**

Model	Variables Entered	Variables Removed	Method
1	X <sup>a</sup>	.	Enter

- a. All requested variables entered.  
 b. Dependent Variable: Y

**Model Summary<sup>b</sup>**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,973 <sup>a</sup>	,946	,944	2,64146	,135

- a. Predictors: (Constant), X  
 b. Dependent Variable: Y

**ANOVA<sup>b</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3658,169	1	3658,169	524,294	,000 <sup>a</sup>
	Residual	209,320	30	6,977		
	Total	3867,489	31			

- a. Predictors: (Constant), X  
 b. Dependent Variable: Y

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	17,674	3,348		5,279	,000
	X	,814	,036	,973	22,897	,000

- a. Dependent Variable: Y

**Ek 15:(Ek 3)'deki Basit Doğrusal Regresyon Modeli için Genelleştirilmiş Farklar Yöntemiyle Elde Edilen Analiz Sonuçları**

**Regression**

**Variables Entered/Removed<sup>b</sup>**

Model	Variables Entered	Variables Removed	Method
1	XTransformed <sup>a</sup>	.	Enter

- a. All requested variables entered.  
b. Dependent Variable: YTransformed

**Model Summary<sup>b</sup>**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,653 <sup>a</sup>	,426	,406	,96969	1,565

- a. Predictors: (Constant), XTransformed  
b. Dependent Variable: YTransformed

**ANOVA<sup>b</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	20,227	1	20,227	21,511	,000 <sup>a</sup>
	Residual	27,269	29	,940		
	Total	47,495	30			

- a. Predictors: (Constant), XTransformed  
b. Dependent Variable: YTransformed

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1,630	1,113		1,464	,154
	XTransformed	,728	,157	,653	4,638	,000

- a. Dependent Variable: YTransformed

**Ek 16: Çoklu Regresyon Modelinde Spearman Sıra Korelasyon Katsayılarının SPSS ile Hesaplanması**

**Nonparametric Correlations**

**Correlations**

			X1	mutstepe
Spearman's rho	X1	Correlation Coefficient	1,000	-,354**
		Sig. (2-tailed)	.	,009
		N	54	54
	mutstepe	Correlation Coefficient	-,354**	1,000
		Sig. (2-tailed)	,009	.
		N	54	54

\*\* . Correlation is significant at the 0.01 level (2-tailed).

**Correlations**

			mutstepe	X2
Spearman's rho	mutstepe	Correlation Coefficient	1,000	,085
		Sig. (2-tailed)	.	,542
		N	54	54
	X2	Correlation Coefficient	,085	1,000
		Sig. (2-tailed)	,542	.
		N	54	54

**Correlations**

			mutstepe	X3
Spearman's rho	mutstepe	Correlation Coefficient	1,000	,234
		Sig. (2-tailed)	.	,089
		N	54	54
	X3	Correlation Coefficient	,234	1,000
		Sig. (2-tailed)	,089	.
		N	54	54

**Ek 17: Çoklu Regresyon Modelinde X1 Ağırlık Değişkeni Olarak Seçilerek Uygulanan AEKK Yönteminin SPSS Çıktısı**

**Weighted Least Squares**

MODEL: MOD\_1.



Source variable.. X1

Log-likelihood Function = 98,333924    POWER value = -2,000  
 Log-likelihood Function = 97,049251    POWER value = -1,500  
 Log-likelihood Function = 95,335041    POWER value = -1,000  
 Log-likelihood Function = 93,194227    POWER value = -,500  
 Log-likelihood Function = 90,633086    POWER value = ,000  
 Log-likelihood Function = 87,657792    POWER value = ,500  
 Log-likelihood Function = 84,275811    POWER value = 1,000  
 Log-likelihood Function = 80,501566    POWER value = 1,500  
 Log-likelihood Function = 76,362388    POWER value = 2,000

The Value of POWER Maximizing Log-likelihood Function = -2,000



Source variable.. X1                      POWER value = -2,000

Dependent variable.. Y

Listwise Deletion of Missing Data

Multiple R            ,99238  
 R Square            ,98481  
 Adjusted R Square   ,98390  
 Standard Error      ,22625

Analysis of Variance:

	DF	Sum of Squares	Mean Square
Regression	3	165,94716	55,315720
Residuals	50	2,55935	,051187

F = 1080,66153    Signif F = ,0000

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
X1	,061970	,002972	,373742	20,852	,0000
X2	,009222	,000333	,497700	27,660	,0000
X3	,009292	,000238	,684753	39,020	,0000
(Constant)	,552323	,031149		17,732	,0000

Log-likelihood Function = 98,333924

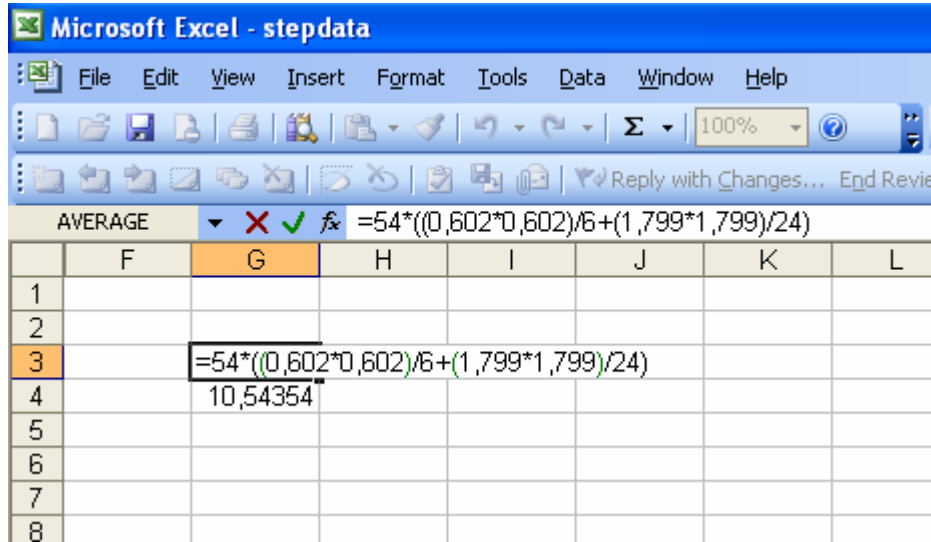
**Ek 18a: Çoklu Regresyon Modelinde Asimetri(S) ve Basıklık(K) Ölçülerinin SPSS ile Hesaplanması**

**Descriptives**

**Descriptive Statistics**

	N	Mean	Std.	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Unstandardized Res	54	0000000	4552806	,602	,325	1,799	,639
Valid N (listwise)	54						

**Ek 18b: Çoklu Regresyon Modelinde JB Test İstatistik Değerinin Excel 2003 ile Hesaplanması**



The screenshot shows the Microsoft Excel 2003 interface. The formula bar displays the formula:  $=54*((0,602*0,602)/6+(1,799*1,799)/24)$ . The spreadsheet grid shows the result of this formula, 10,54354, in cell G3.

	F	G	H	I	J	K	L
1							
2							
3		=54*((0,602*0,602)/6+(1,799*1,799)/24)					
4		10,54354					
5							
6							
7							
8							

**Ek 19: İşgücü Verimliliği Verileri için SPSS ile Korelasyon Tablosu ve “Enter” Yöntemiyle Regresyon Analizi Sonuçları**

**Correlations**

		y	x1	x2	x3	x4	x5	x6	x7
y	Pearson Correlation	1	,243	,577**	,437**	-,301*	-,189	,241	-,429**
	Sig. (2-tailed)		,061	,000	,000	,019	,149	,064	,001
	N	60	60	60	60	60	60	60	60
x1	Pearson Correlation	,243	1	-,001	,070	,010	,276*	,430**	,068
	Sig. (2-tailed)	,061		,996	,594	,938	,033	,001	,604
	N	60	60	60	60	60	60	60	60
x2	Pearson Correlation	,577**	-,001	1	-,178	-,225	-,123	-,023	-,387**
	Sig. (2-tailed)	,000	,996		,175	,084	,350	,859	,002
	N	60	60	60	60	60	60	60	60
x3	Pearson Correlation	,437**	,070	-,178	1	,212	-,096	,077	-,260*
	Sig. (2-tailed)	,000	,594	,175		,104	,467	,560	,045
	N	60	60	60	60	60	60	60	60
x4	Pearson Correlation	-,301*	,010	-,225	,212	1	,133	-,017	,230
	Sig. (2-tailed)	,019	,938	,084	,104		,313	,900	,077
	N	60	60	60	60	60	60	60	60
x5	Pearson Correlation	-,189	,276*	-,123	-,096	,133	1	,102	,017
	Sig. (2-tailed)	,149	,033	,350	,467	,313		,439	,898
	N	60	60	60	60	60	60	60	60
x6	Pearson Correlation	,241	,430**	-,023	,077	-,017	,102	1	,026
	Sig. (2-tailed)	,064	,001	,859	,560	,900	,439		,845
	N	60	60	60	60	60	60	60	60
x7	Pearson Correlation	-,429**	,068	-,387**	-,260*	,230	,017	,026	1
	Sig. (2-tailed)	,001	,604	,002	,045	,077	,898	,845	
	N	60	60	60	60	60	60	60	60

\*\* . Correlation is significant at the 0.01 level (2-tailed).

\* . Correlation is significant at the 0.05 level (2-tailed).

**Regression**

**Variables Entered/Removed<sup>a</sup>**

Model	Variables Entered	Variables Removed	Method
1	x7, x5, x6, x4 <sub>a</sub> , x3, x1, x2	.	Enter

a. All requested variables entered.

b. Dependent Variable: y

**Model Summary<sup>b</sup>**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,881 <sup>a</sup>	,777	,747	22068,6446	1,939

a. Predictors: (Constant), x7, x5, x6, x4, x3, x1, x2

b. Dependent Variable: y

**ANOVA<sup>b</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	8,8E+010	7	1,258E+010	25,829	,000 <sup>a</sup>
	Residual	2,5E+010	52	487025073,1		
	Total	1,1E+011	59			

a. Predictors: (Constant), x7, x5, x6, x4, x3, x1, x2

b. Dependent Variable: y

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	27787,266	9282,556		2,993	,004		
	x1	2,96E-005	,000	,167	2,186	,033	,740	1,352
	x2	,148	,018	,614	8,041	,000	,736	1,358
	x3	223990,5	29668,209	,577	7,550	,000	,737	1,358
	x4	-24708,9	6426,167	-,276	-3,845	,000	,835	1,197
	x5	-2E+007	1E+007	-,082	-1,155	,253	,849	1,178
	x6	253,793	129,387	,143	1,961	,055	,811	1,233
	x7	795,308	7605,574	,008	,105	,917	,668	1,497

a. Dependent Variable: y



**Ek 20: İşgücü Verimliliği Verileri için “Stepwise” Yöntemiyle Elde Edilen SPSS Çıktısı**

**Regression**

**Variables Entered/Removed<sup>f</sup>**

Model	Variables Entered	Variables Removed	Method
1	x2	.	Stepwise (Criteria: F-to-enter >= 4,000, F-to-remo ve <= 3,990).
2	x3	.	Stepwise (Criteria: F-to-enter >= 4,000, F-to-remo ve <= 3,990).
3	x4	.	Stepwise (Criteria: F-to-enter >= 4,000, F-to-remo ve <= 3,990).
4	x6	.	Stepwise (Criteria: F-to-enter >= 4,000, F-to-remo ve <= 3,990).

a. Dependent Variable: y

**Model Summary<sup>f</sup>**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,577 <sup>a</sup>	,333	,322	36103,8216	
2	,796 <sup>b</sup>	,634	,621	26973,2369	
3	,844 <sup>c</sup>	,712	,697	24145,5672	
4	,868 <sup>d</sup>	,754	,736	22528,3904	2,134

a. Predictors: (Constant), x2

b. Predictors: (Constant), x2, x3

c. Predictors: (Constant), x2, x3, x4

d. Predictors: (Constant), x2, x3, x4, x6

e. Dependent Variable: y

**ANOVA<sup>e</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3,8E+010	1	3,778E+010	28,982	,000 <sup>a</sup>
	Residual	7,6E+010	58	1303485931		
	Total	1,1E+011	59			
2	Regression	7,2E+010	2	3,595E+010	49,419	,000 <sup>b</sup>
	Residual	4,1E+010	57	727555510,1		
	Total	1,1E+011	59			
3	Regression	8,1E+010	3	2,691E+010	46,158	,000 <sup>c</sup>
	Residual	3,3E+010	56	583008415,6		
	Total	1,1E+011	59			
4	Regression	8,5E+010	4	2,137E+010	42,099	,000 <sup>d</sup>
	Residual	2,8E+010	55	507528375,1		
	Total	1,1E+011	59			

- a. Predictors: (Constant), x2
- b. Predictors: (Constant), x2, x3
- c. Predictors: (Constant), x2, x3, x4
- d. Predictors: (Constant), x2, x3, x4, x6
- e. Dependent Variable: y

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	31879,485	5675,973		5,617	,000		
	x2	,139	,026	,577	5,384	,000	1,000	1,000
2	(Constant)	15892,188	4840,495		3,283	,002		
	x2	,163	,020	,676	8,308	,000	,968	1,033
	x3	216600,6	31623,857	,558	6,849	,000	,968	1,033
3	(Constant)	31238,393	5859,915		5,331	,000		
	x2	,150	,018	,620	8,345	,000	,932	1,073
	x3	236650,2	28774,033	,609	8,224	,000	,937	1,067
	x4	-26074,1	6702,824	-,291	-3,890	,000	,919	1,088
4	(Constant)	27498,858	5602,857		4,908	,000		
	x2	,151	,017	,624	8,995	,000	,931	1,074
	x3	230141,5	26931,305	,592	8,546	,000	,932	1,074
	x4	-25378,6	6258,038	-,283	-4,055	,000	,918	1,090
	x6	364,680	119,401	,205	3,054	,003	,993	1,007

- a. Dependent Variable: y

**Ek 21: İşgücü Verimliliği Verileri için Spearman Sıra Korelasyon Katsayılarının SPSS ile Hesaplanması**

**Nonparametric Correlations**

**Correlations**

			x2	abserror
Spearman's rho	x2	Correlation Coefficient	1,000	,126
		Sig. (2-tailed)	.	,336
		N	60	60
	abserror	Correlation Coefficient	,126	1,000
		Sig. (2-tailed)	,336	.
		N	60	60

**Correlations**

			abserror	x3
Spearman's rho	abserror	Correlation Coefficient	1,000	,146
		Sig. (2-tailed)	.	,267
		N	60	60
	x3	Correlation Coefficient	,146	1,000
		Sig. (2-tailed)	,267	.
		N	60	60

**Correlations**

			abserror	x4
Spearman's rho	abserror	Correlation Coefficient	1,000	-,011
		Sig. (2-tailed)	.	,934
		N	60	60
	x4	Correlation Coefficient	-,011	1,000
		Sig. (2-tailed)	,934	.
		N	60	60

**Correlations**

			abserror	x6
Spearman's rho	abserror	Correlation Coefficient	1,000	,135
		Sig. (2-tailed)	.	,305
		N	60	60
	x6	Correlation Coefficient	,135	1,000
		Sig. (2-tailed)	,305	.
		N	60	60

## Ek 22a: İşgücü Verimliliği Verileri için Asimetri ve Basıklık Ölçülerinin SPSS ile Hesaplanması

### Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std.	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Unstandardized Residual	60	-49412,9	70099,93	,0000000	21751,31	,291	,309	,853	,608
Valid N (listwise)	60								

## Ek 22b: İşgücü Verimliliği Verileri için JB Test İstatistik Değerinin Excel 2003 ile Hesaplanması

	F	G	H	I	J
1			0,291	0,853	
2					
3		=60*(\$H\$1*\$H\$1)/6+(\$I\$1*\$I\$1)/24			
4		2,665833			
5					

**Ek 23: t-Dağılım Tablosu**

SD	Tek Taraflı Test					
	.10	.05	0.25	.01	.005	.0005
	Çift Taraflı Test					
	.20	.10	.05	.02	.01	.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.965	31.598
3	1.638	2.353	3.182	4.541	5.841	12.941
4	1.533	2.132	2.776	3.747	4.601	8.610
5	1.476	2.015	2.571	3.365	4.032	6.859
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.698	3.499	5.405
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587
11	1.363	1.796	2.201	2.718	3.106	4.437
12	1.356	1.782	2.179	2.681	3.055	4.318
13	1.350	1.771	2.160	2.650	3.012	4.221
14	1.345	1.761	2.145	2.624	2.977	4.140
15	1.341	1.753	2.131	2.602	2.947	4.073
16	1.337	1.746	2.120	2.583	2.921	4.015
17	1.333	1.740	2.110	2.567	2.898	3.965
18	1.330	1.734	2.101	2.552	2.878	3.922
19	1.328	1.729	2.093	2.539	2.861	3.883
20	1.325	1.725	2.086	2.528	2.845	3.850
21	1.323	1.721	2.080	2.518	2.831	3.819
22	1.321	1.717	2.074	2.508	2.819	3.792
23	1.319	1.714	2.069	2.500	2.807	3.767
24	1.318	1.711	2.064	2.492	2.797	3.745
25	1.316	1.708	2.060	2.485	2.787	3.725
26	1.315	1.706	2.056	2.479	2.779	3.707
27	1.314	1.703	2.052	2.473	2.771	3.690
28	1.313	1.701	2.048	2.467	2.763	3.647
29	1.311	1.499	2.045	2.462	2.756	3.659
30	1.310	1.697	2.042	2.457	2.750	3.646
40	1.303	1.684	2.021	2.423	2.704	3.551
60	1.296	1.671	2.000	2.390	2.660	3.460
120	1.289	1.658	1.980	2.358	2.617	3.373
-	1.282	1.645	1.960	2.326	2.576	3.291

**Kaynak:** Selahattin Güriş, Ebru Çağlayan, **Ekonometri Temel Kavramlar**, İstanbul, Der Yayınevi Yayın No:282, 2000, s. 788.

Ek 24: F Dağılım Tablosu,  $\alpha=0,05$

$v_2 \backslash v_1$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.4	199.5	251.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	218.0	249.1	250.1	251.1	252.2	253.3	251.3
2	18.51	9.16	9.16	19.25	19.20	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.45	19.47	19.43	19.49	19.50
3	10.18	9.28	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.04	8.04	8.59	8.57	8.55	8.53
4	7.71	6.59	6.59	9.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.77	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.61	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	2.15	2.12	3.08	3.04	3.01	3.97	3.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	3.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.59	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.35	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.43	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.20	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.10	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.16	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	3.00	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	3.00	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.73	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.71	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.70	1.65
29	4.18	3.33	2.92	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.65
30	4.17	3.22	2.92	2.69	2.53	2.43	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.18	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Kaynak: Selahattin Gürış, Ebru Çağlayan, Ekonometri Temel Kavramlar, İstanbul, Der Yayinevi Yayın No:282, 2000, s. 790.

Ek 25: F Dağılım Tablosu,  $\alpha=0,01$

$v_1 \backslash v_2$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	00
1	4052	4999.5	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	29.69	26.60	26.50	26.41	23.32	26.22	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.08	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.47	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	4.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.09
15	8.68	6.36	5.42	4.80	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.85	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.57	2.45	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.08	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
00	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

Kaynak: Selahattin Gürış, Ebru Çağlayan, **Ekonometri Temel Kavramlar**, İstanbul, Der YayıneviYayın No:282, 2000, s. 791.

Ek 26a: Durbin-Watson d İstatistiği Tablosu,  $\alpha=0,05$

n	k' = 1		k' = 2		k' = 3		k' = 4		k' = 5		k' = 6		k' = 7		k' = 8		k' = 9		k' = 10	
	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU
6	0.610	1.400	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
7	0.700	1.356	0.467	1.896	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
8	0.763	1.332	0.559	1.777	0.368	2.287	—	—	—	—	—	—	—	—	—	—	—	—	—	—
9	0.824	1.320	0.629	1.699	0.455	2.128	0.296	2.588	—	—	—	—	—	—	—	—	—	—	—	—
10	0.879	1.320	0.697	1.641	0.525	2.016	0.376	2.414	0.243	2.822	—	—	—	—	—	—	—	—	—	—
11	0.927	1.324	0.758	1.604	0.595	1.928	0.444	2.283	0.316	2.645	0.203	3.005	—	—	—	—	—	—	—	—
12	0.971	1.331	0.812	1.579	0.658	1.864	0.512	2.177	0.379	2.506	0.268	2.832	0.171	3.149	—	—	—	—	—	—
13	1.010	1.340	0.861	1.562	0.715	1.816	0.574	2.094	0.445	2.390	0.328	2.692	0.230	2.985	0.147	3.266	—	—	—	—
14	1.045	1.350	0.905	1.551	0.767	1.779	0.632	2.030	0.505	2.296	0.389	2.572	0.286	2.848	0.200	3.111	0.127	3.360	—	—
15	1.077	1.361	0.946	1.543	0.814	1.750	0.685	1.977	0.562	2.220	0.447	2.472	0.343	2.727	0.251	2.979	0.175	3.216	0.111	3.438
16	1.106	1.371	0.982	1.539	0.857	1.728	0.734	1.935	0.615	2.157	0.502	2.388	0.398	2.624	0.304	2.860	0.222	3.090	0.155	3.304
17	1.133	1.381	1.015	1.536	0.897	1.710	0.779	1.900	0.664	2.104	0.554	2.318	0.451	2.537	0.356	2.757	0.272	2.975	0.198	3.184
18	1.158	1.391	1.046	1.535	0.933	1.696	0.820	1.872	0.710	2.060	0.603	2.257	0.502	2.461	0.407	2.667	0.321	2.873	0.244	3.073
19	1.180	1.401	1.074	1.536	0.967	1.685	0.859	1.848	0.752	2.023	0.649	2.206	0.459	2.396	0.456	2.589	0.369	2.783	0.290	2.974
20	1.201	1.411	1.100	1.537	0.998	1.676	0.894	1.828	0.792	1.991	0.692	2.162	0.595	2.339	0.502	2.521	0.416	2.704	0.336	2.885
21	1.221	1.420	1.125	1.538	1.026	1.669	0.927	1.812	0.829	1.964	0.732	2.124	0.637	2.290	0.547	2.460	0.461	2.633	0.380	2.806
22	1.239	1.429	1.147	1.541	1.053	1.664	0.958	1.797	0.863	1.940	0.769	2.090	0.677	2.246	0.588	2.407	0.504	2.571	0.424	2.734
23	1.257	1.437	1.168	1.543	1.078	1.660	0.986	1.785	0.895	1.920	0.804	2.061	0.715	2.208	0.628	2.360	0.545	2.514	0.465	2.670
24	1.273	1.446	1.188	1.546	1.101	1.656	1.013	1.775	0.925	1.902	0.837	2.035	0.751	2.174	0.666	2.318	0.584	2.464	0.506	2.613
25	1.288	1.454	1.206	1.550	1.123	1.654	1.038	1.767	0.953	1.886	0.868	2.012	0.784	2.144	0.702	2.280	0.621	2.419	0.544	2.560
26	1.302	1.461	1.224	1.553	1.143	1.652	1.062	1.759	0.979	1.873	0.897	1.992	0.816	2.117	0.735	2.246	0.657	2.379	0.581	2.513
27	1.316	1.469	1.240	1.556	1.162	1.651	1.084	1.753	1.004	1.861	0.925	1.974	0.845	2.093	0.767	2.216	0.691	2.342	0.616	2.470
28	1.328	1.476	1.255	1.560	1.181	1.650	1.104	1.747	1.028	1.850	0.951	1.958	0.874	2.071	0.798	2.188	0.723	2.309	0.650	2.431
29	1.341	1.483	1.270	1.563	1.198	1.650	1.124	1.743	1.050	1.841	0.975	1.944	0.900	2.052	0.826	2.164	0.753	2.278	0.682	2.396
30	1.352	1.489	1.284	1.567	1.214	1.650	1.143	1.739	1.071	1.833	0.998	1.931	0.926	2.034	0.854	2.141	0.782	2.251	0.712	2.363
31	1.363	1.496	1.297	1.570	1.229	1.650	1.160	1.735	1.090	1.825	1.020	1.920	0.950	2.018	0.879	2.120	0.810	2.226	0.741	2.333
32	1.373	1.502	1.309	1.574	1.244	1.650	1.177	1.732	1.109	1.819	1.041	1.909	0.972	2.004	0.904	2.102	0.836	2.203	0.769	2.306
33	1.383	1.508	1.321	1.577	1.258	1.651	1.193	1.730	1.127	1.813	1.061	1.900	0.994	1.991	0.927	2.085	0.861	2.181	0.795	2.281
34	1.393	1.514	1.333	1.580	1.271	1.652	1.208	1.728	1.144	1.808	1.080	1.891	1.015	1.979	0.950	2.069	0.885	2.162	0.821	2.257
35	1.402	1.519	1.343	1.584	1.283	1.653	1.222	1.726	1.160	1.803	1.097	1.884	1.034	1.967	0.971	2.054	0.908	2.144	0.845	2.236
36	1.411	1.525	1.354	1.587	1.295	1.654	1.236	1.724	1.175	1.799	1.114	1.877	1.053	1.957	0.991	2.041	0.930	2.127	0.868	2.216
37	1.419	1.530	1.364	1.590	1.307	1.655	1.249	1.723	1.190	1.795	1.131	1.870	1.071	1.948	1.011	2.029	0.951	2.112	0.891	2.198



<i>n</i>	<i>k'</i> = 1		<i>k'</i> = 2		<i>k'</i> = 3		<i>k'</i> = 4		<i>k'</i> = 5		<i>k'</i> = 6		<i>k'</i> = 7		<i>k'</i> = 8		<i>k'</i> = 9		<i>k'</i> = 10	
	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>
38	1.427	1.535	1.373	1.594	1.318	1.656	1.261	1.722	1.204	1.792	1.146	1.864	1.088	1.939	1.029	2.017	0.970	2.098	0.912	2.180
39	1.435	1.540	1.382	1.597	1.328	1.658	1.273	1.722	1.218	1.789	1.161	1.859	1.104	1.932	1.047	2.007	0.990	2.085	0.932	2.164
40	1.442	1.544	1.391	1.600	1.338	1.659	1.285	1.721	1.230	1.786	1.175	1.854	1.120	1.924	1.064	1.997	1.008	2.072	0.945	2.149
45	1.475	1.566	1.430	1.615	1.383	1.666	1.336	1.720	1.287	1.776	1.238	1.835	1.189	1.895	1.139	1.958	1.089	2.002	1.038	2.088
50	1.503	1.585	1.462	1.628	1.421	1.674	1.378	1.721	1.355	1.771	1.291	1.822	1.246	1.875	1.201	1.930	1.156	1.986	1.110	2.044
55	1.528	1.601	1.490	1.641	1.452	1.681	1.414	1.724	1.374	1.768	1.334	1.814	1.294	1.861	1.253	1.909	1.212	1.959	1.170	2.010
60	1.549	1.616	1.514	1.652	1.480	1.689	1.444	1.727	1.408	1.767	1.372	1.808	1.335	1.850	1.298	1.894	1.260	1.939	1.222	1.984
65	1.567	1.629	1.536	1.662	1.503	1.696	1.471	1.731	1.438	1.767	1.404	1.805	1.370	1.843	1.336	1.882	1.301	1.923	1.266	1.964
70	1.583	1.641	1.554	1.672	1.525	1.703	1.494	1.735	1.464	1.768	1.433	1.802	1.401	1.837	1.369	1.873	1.337	1.910	1.305	1.948
75	1.598	1.652	1.571	1.680	1.543	1.709	1.515	1.739	1.487	1.770	1.458	1.801	1.428	1.834	1.399	1.867	1.369	1.901	1.339	1.935
80	1.611	1.662	1.586	1.688	1.560	1.715	1.534	1.743	1.507	1.772	1.480	1.801	1.453	1.831	1.425	1.861	1.397	1.893	1.369	1.925
85	1.624	1.671	1.600	1.696	1.575	1.721	1.550	1.747	1.525	1.774	1.500	1.801	1.474	1.829	1.448	1.857	1.422	1.886	1.396	1.916
90	1.635	1.679	1.612	1.703	1.589	1.726	1.566	1.751	1.542	1.776	1.518	1.801	1.494	1.827	1.469	1.854	1.445	1.881	1.420	1.909
95	1.645	1.687	1.623	1.709	1.602	1.732	1.579	1.755	1.557	1.778	1.535	1.802	1.512	1.827	1.489	1.852	1.465	1.877	1.442	1.903
100	1.654	1.694	1.634	1.715	1.613	1.736	1.592	1.758	1.571	1.780	1.550	1.803	1.528	1.826	1.506	1.850	1.484	1.874	1.462	1.898
150	1.720	1.746	1.706	1.760	1.693	1.774	1.679	1.788	1.665	1.802	1.651	1.817	1.637	1.832	1.622	1.847	1.608	1.862	1.594	1.877
200	1.758	1.778	1.748	1.789	1.738	1.799	1.728	1.810	1.718	1.820	1.707	1.831	1.697	1.841	1.686	1.852	1.675	1.863	1.665	1.874
<i>n</i>	<i>k'</i> = 11		<i>k'</i> = 12		<i>k'</i> = 13		<i>k'</i> = 14		<i>k'</i> = 15		<i>k'</i> = 16		<i>k'</i> = 17		<i>k'</i> = 18		<i>k'</i> = 19		<i>k'</i> = 20	
	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>
16	0.098	3.503	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
17	0.138	3.378	0.087	3.557	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
18	0.177	3.265	0.123	3.441	0.078	3.603	—	—	—	—	—	—	—	—	—	—	—	—	—	—
19	0.220	3.159	0.160	3.335	0.111	3.496	0.070	3.642	—	—	—	—	—	—	—	—	—	—	—	—
20	0.263	3.063	0.200	3.234	0.145	3.395	0.100	3.542	0.063	3.676	—	—	—	—	—	—	—	—	—	—
21	0.307	2.976	0.240	3.141	0.182	3.300	0.132	3.448	0.091	3.583	0.058	3.705	—	—	—	—	—	—	—	—
22	0.349	2.897	0.281	3.057	0.220	3.211	0.166	3.358	0.120	3.495	0.083	3.619	0.052	3.731	—	—	—	—	—	—
23	0.391	2.826	0.322	2.979	0.259	3.128	0.202	3.272	0.153	3.409	0.110	3.535	0.076	3.650	0.048	3.753	—	—	—	—
24	0.431	2.761	0.362	2.908	0.297	3.053	0.239	3.193	0.186	3.327	0.141	3.454	0.101	3.572	0.070	3.678	0.044	3.773	—	—

$n$	$k' = 11$		$k' = 12$		$k' = 13$		$k' = 14$		$k' = 15$		$k' = 16$		$k' = 17$		$k' = 18$		$k' = 19$		$k' = 20$	
	$dL$	$dU$	$dL$	$dU$	$dL$	$dU$	$dL$	$dU$	$dL$	$dU$	$dL$	$dU$	$dL$	$dU$	$dL$	$dU$	$dL$	$dU$	$dL$	$dU$
25	0.470	2.702	0.400	2.844	0.335	2.983	0.275	3.119	0.221	3.251	0.172	3.376	0.130	3.494	0.094	3.604	0.065	3.702	0.041	3.790
26	0.508	2.649	0.438	2.784	0.373	2.919	0.312	3.051	0.256	3.179	0.205	3.303	0.160	3.420	0.120	3.531	0.087	3.632	0.060	3.724
27	0.544	2.600	0.475	2.730	0.409	2.859	0.348	2.987	0.291	3.112	0.238	3.233	0.191	3.349	0.149	3.460	0.112	3.563	0.081	3.658
28	0.578	2.555	0.510	2.680	0.445	2.805	0.383	2.928	0.325	3.050	0.271	3.168	0.222	3.283	0.178	3.392	0.138	3.495	0.104	3.592
29	0.612	2.515	0.544	2.634	0.479	2.755	0.418	2.874	0.359	2.992	0.305	3.107	0.254	3.219	0.208	3.327	0.166	3.431	0.129	3.528
30	0.643	2.477	0.577	2.592	0.512	2.708	0.451	2.823	0.392	2.937	0.337	3.050	0.286	3.160	0.238	3.266	0.195	3.368	0.156	3.465
31	0.674	2.443	0.608	2.553	0.545	2.665	0.484	2.776	0.425	2.987	0.370	2.996	0.317	3.103	0.269	3.208	0.224	3.309	0.183	3.406
32	0.703	2.411	0.638	2.517	0.576	2.625	0.515	2.733	0.457	2.840	0.401	2.946	0.349	3.050	0.299	3.153	0.253	3.252	0.211	3.348
33	0.731	2.382	0.668	2.484	0.606	2.588	0.546	2.692	0.488	2.796	0.432	2.899	0.379	3.000	0.329	3.100	0.283	3.198	0.239	3.293
34	0.758	2.355	0.695	2.454	0.634	2.554	0.575	2.654	0.518	2.754	0.462	2.854	0.409	2.954	0.359	3.051	0.312	3.147	0.267	3.240
35	0.783	2.330	0.722	2.425	0.662	2.521	0.604	2.619	0.547	2.716	0.492	2.813	0.439	2.910	0.388	3.005	0.340	3.099	0.295	3.190
36	0.808	2.306	0.748	2.398	0.689	2.492	0.631	2.586	0.575	2.680	0.520	2.774	0.467	2.868	0.417	2.961	0.369	3.053	0.323	3.142
37	0.831	2.285	0.772	2.374	0.714	2.464	0.657	2.555	0.602	2.646	0.548	2.738	0.495	2.829	0.445	2.920	0.397	3.009	0.351	3.097
38	0.854	2.265	0.796	2.351	0.739	2.438	0.683	2.526	0.628	2.614	0.575	2.703	0.522	2.792	0.472	2.880	0.424	2.968	0.378	3.054
39	0.875	2.246	0.819	2.329	0.763	2.413	0.707	2.499	0.653	2.585	0.600	2.671	0.549	2.757	0.499	2.843	0.451	2.929	0.404	3.013
40	0.896	2.228	0.840	2.309	0.785	2.391	0.731	2.473	0.678	2.557	0.626	2.641	0.575	2.724	0.525	2.808	0.477	2.892	0.430	2.974
45	0.988	2.156	0.938	2.225	0.887	2.296	0.838	2.367	0.788	2.439	0.740	2.512	0.692	2.586	0.644	2.659	0.598	2.733	0.553	2.807
50	1.064	2.103	1.019	2.163	0.973	2.225	0.927	2.287	0.882	2.350	0.836	2.414	0.792	2.479	0.747	2.544	0.703	2.610	0.660	2.675
55	1.129	2.062	1.087	2.116	1.045	2.170	1.003	2.225	0.961	2.281	0.919	2.338	0.877	2.396	0.836	2.454	0.795	2.512	0.754	2.571
60	1.184	2.031	1.145	2.079	1.106	2.127	1.068	2.177	1.029	2.227	0.990	2.278	0.951	2.330	0.913	2.382	0.874	2.434	0.836	2.487
65	1.231	2.006	1.195	2.049	1.160	2.093	1.124	2.138	1.088	2.183	1.052	2.229	1.016	2.276	0.980	2.323	0.944	2.371	0.908	2.419
70	1.272	1.986	1.239	2.026	1.206	2.066	1.172	2.106	1.139	2.148	1.105	2.189	1.072	2.232	1.038	2.275	1.005	2.318	0.971	2.362
75	1.308	1.970	1.277	2.006	1.247	2.043	1.215	2.080	1.184	2.118	1.153	2.156	1.121	2.195	1.090	2.235	1.058	2.275	1.027	2.315
80	1.340	1.957	1.311	1.991	1.283	2.024	1.253	2.059	1.224	2.093	1.195	2.129	1.165	2.165	1.136	2.201	1.106	2.238	1.076	2.275
85	1.369	1.946	1.342	1.977	1.315	2.009	1.287	2.040	1.260	2.073	1.232	2.105	1.205	2.139	1.177	2.172	1.149	2.206	1.121	2.241
90	1.395	1.937	1.369	1.966	1.344	1.995	1.318	2.025	1.292	2.055	1.266	2.085	1.240	2.116	1.213	2.148	1.187	2.179	1.160	2.211
95	1.418	1.929	1.394	1.956	1.370	1.984	1.345	2.012	1.321	2.040	1.296	2.068	1.271	2.097	1.247	2.126	1.222	2.156	1.197	2.186
00	1.434	1.923	1.416	1.948	1.393	1.974	1.371	2.000	1.347	2.026	1.324	2.053	1.301	2.080	1.277	2.108	1.253	2.135	1.229	2.164
50	1.579	1.892	1.564	1.908	1.550	1.924	1.535	1.940	1.519	1.956	1.504	1.972	1.489	1.989	1.474	2.006	1.458	2.023	1.443	2.040
00	1.654	1.885	1.643	1.896	1.632	1.908	1.621	1.919	1.610	1.931	1.599	1.943	1.588	1.955	1.576	1.967	1.565	1.979	1.554	1.991

**Ek 26b: Durbin-Watson d İstatistiği Tablosu,  $\alpha=0,01$**

<i>n</i>	<i>k'</i> = 1		<i>k'</i> = 2		<i>k'</i> = 3		<i>k'</i> = 4		<i>k'</i> = 5		<i>k'</i> = 6		<i>k'</i> = 7		<i>k'</i> = 8		<i>k'</i> = 9		<i>k'</i> = 10	
	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>
6	0.390	1.142	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
7	0.435	1.036	0.294	1.676	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
8	0.497	1.003	0.345	1.489	0.229	2.102	—	—	—	—	—	—	—	—	—	—	—	—	—	—
9	0.554	0.998	0.408	1.389	0.279	1.875	0.183	2.433	—	—	—	—	—	—	—	—	—	—	—	—
10	0.604	1.001	0.466	1.333	0.340	1.733	0.230	2.193	0.150	2.690	—	—	—	—	—	—	—	—	—	—
11	0.653	1.010	0.519	1.297	0.396	1.640	0.286	2.030	0.193	2.453	0.124	2.892	—	—	—	—	—	—	—	—
12	0.697	1.023	0.569	1.274	0.449	1.575	0.339	1.913	0.244	2.280	0.164	2.665	0.105	3.053	—	—	—	—	—	—
13	0.738	1.038	0.616	1.261	0.499	1.526	0.391	1.826	0.294	2.150	0.211	2.490	0.140	2.838	0.090	3.182	—	—	—	—
14	0.776	1.054	0.660	1.254	0.547	1.490	0.441	1.757	0.343	2.049	0.257	2.354	0.183	2.667	0.122	2.981	0.078	3.287	—	—
15	0.811	1.070	0.700	1.252	0.591	1.464	0.488	1.704	0.391	1.967	0.303	2.244	0.226	2.530	0.161	2.817	0.107	3.101	0.068	3.374
16	0.844	1.086	0.737	1.252	0.633	1.446	0.532	1.663	0.437	1.900	0.349	2.153	0.269	2.416	0.200	2.681	0.142	2.944	0.094	3.201
17	0.874	1.102	0.772	1.255	0.672	1.432	0.574	1.630	0.480	1.847	0.393	2.078	0.313	2.319	0.241	2.566	0.179	2.811	0.127	3.053
18	0.902	1.118	0.805	1.259	0.708	1.422	0.613	1.604	0.522	1.803	0.435	2.015	0.355	2.238	0.282	2.467	0.216	2.697	0.160	2.925
19	0.928	1.132	0.835	1.265	0.742	1.415	0.650	1.584	0.561	1.767	0.476	1.963	0.396	2.169	0.322	2.381	0.255	2.597	0.196	2.813
20	0.952	1.147	0.863	1.271	0.773	1.411	0.685	1.567	0.598	1.737	0.515	1.918	0.436	2.110	0.362	2.308	0.294	2.510	0.232	2.714
21	0.975	1.161	0.890	1.277	0.803	1.408	0.718	1.554	0.633	1.712	0.552	1.881	0.474	2.059	0.400	2.244	0.331	2.434	0.268	2.625
22	0.997	1.174	0.914	1.284	0.831	1.407	0.748	1.543	0.667	1.691	0.587	1.849	0.510	2.015	0.437	2.188	0.368	2.367	0.304	2.548
23	1.018	1.187	0.938	1.291	0.858	1.407	0.777	1.534	0.698	1.673	0.620	1.821	0.545	1.977	0.473	2.140	0.404	2.308	0.340	2.479
24	1.037	1.199	0.960	1.298	0.882	1.407	0.805	1.528	0.728	1.658	0.652	1.797	0.578	1.944	0.507	2.097	0.439	2.255	0.375	2.417
25	1.055	1.211	0.981	1.305	0.906	1.409	0.831	1.523	0.756	1.645	0.682	1.766	0.610	1.915	0.540	2.059	0.473	2.209	0.409	2.362
26	1.072	1.222	1.001	1.312	0.928	1.411	0.855	1.518	0.783	1.635	0.711	1.759	0.640	1.889	0.572	2.026	0.505	2.168	0.441	2.313
27	1.089	1.233	1.019	1.319	0.949	1.413	0.878	1.515	0.808	1.626	0.738	1.743	0.669	1.867	0.602	1.997	0.536	2.131	0.473	2.269
28	1.104	1.244	1.037	1.325	0.969	1.415	0.900	1.513	0.832	1.618	0.764	1.729	0.696	1.847	0.630	1.970	0.566	2.098	0.504	2.229
29	1.119	1.254	1.054	1.332	0.988	1.418	0.921	1.512	0.855	1.611	0.788	1.718	0.723	1.830	0.658	1.947	0.595	2.068	0.533	2.193
30	1.133	1.263	1.070	1.339	1.006	1.421	0.941	1.511	0.877	1.606	0.812	1.707	0.748	1.814	0.684	1.925	0.622	2.041	0.562	2.160
31	1.147	1.273	1.085	1.345	1.023	1.425	0.960	1.510	0.897	1.601	0.834	1.698	0.772	1.800	0.710	1.906	0.649	2.017	0.589	2.131
32	1.160	1.282	1.100	1.352	1.040	1.428	0.979	1.510	0.917	1.597	0.856	1.690	0.794	1.788	0.734	1.889	0.674	1.995	0.615	2.104
33	1.172	1.291	1.114	1.358	1.055	1.432	0.996	1.510	0.936	1.594	0.876	1.683	0.816	1.776	0.757	1.874	0.698	1.975	0.641	2.080
34	1.184	1.299	1.128	1.364	1.070	1.435	1.012	1.511	0.954	1.591	0.896	1.677	0.837	1.766	0.779	1.860	0.722	1.957	0.665	2.057
35	1.195	1.307	1.140	1.370	1.085	1.439	1.028	1.512	0.971	1.589	0.914	1.671	0.857	1.757	0.800	1.847	0.744	1.940	0.689	2.037
36	1.206	1.315	1.153	1.376	1.098	1.442	1.043	1.513	0.988	1.588	0.932	1.666	0.877	1.749	0.821	1.836	0.766	1.925	0.711	2.018
37	1.217	1.323	1.165	1.382	1.112	1.446	1.058	1.514	1.004	1.586	0.950	1.662	0.895	1.742	0.841	1.825	0.787	1.911	0.733	2.001

<i>n</i>	<i>k'</i> = 1		<i>k'</i> = 2		<i>k'</i> = 3		<i>k'</i> = 4		<i>k'</i> = 5		<i>k'</i> = 6		<i>k'</i> = 7		<i>k'</i> = 8		<i>k'</i> = 9		<i>k'</i> = 10	
	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>
38	1.227	1.330	1.176	1.388	1.124	1.449	1.072	1.515	1.019	1.585	0.966	1.658	0.913	1.735	0.860	1.816	0.807	1.899	0.754	1.985
39	1.237	1.337	1.187	1.393	1.137	1.453	1.085	1.517	1.034	1.584	0.982	1.655	0.930	1.729	0.878	1.807	0.826	1.887	0.774	1.970
40	1.246	1.344	1.198	1.398	1.148	1.457	1.098	1.518	1.048	1.584	0.997	1.652	0.946	1.724	0.895	1.799	0.844	1.876	0.789	1.956
45	1.288	1.376	1.245	1.423	1.201	1.474	1.156	1.528	1.111	1.584	1.065	1.643	1.019	1.704	0.974	1.768	0.927	1.834	0.881	1.902
50	1.324	1.403	1.285	1.446	1.245	1.491	1.205	1.538	1.164	1.587	1.123	1.639	1.081	1.692	1.039	1.748	0.997	1.805	0.955	1.864
55	1.356	1.427	1.320	1.466	1.284	1.506	1.247	1.548	1.209	1.592	1.172	1.638	1.134	1.685	1.095	1.734	1.057	1.785	1.018	1.837
60	1.383	1.449	1.350	1.484	1.317	1.520	1.283	1.558	1.249	1.598	1.214	1.639	1.179	1.682	1.144	1.726	1.108	1.771	1.072	1.817
65	1.407	1.468	1.377	1.500	1.346	1.534	1.315	1.568	1.283	1.604	1.251	1.642	1.218	1.680	1.186	1.720	1.153	1.761	1.120	1.802
70	1.429	1.485	1.400	1.515	1.372	1.546	1.343	1.578	1.313	1.611	1.283	1.645	1.253	1.680	1.223	1.716	1.192	1.754	1.162	1.792
75	1.448	1.501	1.422	1.529	1.395	1.557	1.368	1.587	1.340	1.617	1.313	1.646	1.284	1.682	1.256	1.716	1.227	1.746	1.199	1.785
80	1.466	1.515	1.441	1.541	1.416	1.568	1.390	1.595	1.364	1.624	1.338	1.653	1.312	1.683	1.285	1.714	1.259	1.745	1.232	1.777
85	1.482	1.528	1.458	1.553	1.435	1.578	1.411	1.603	1.386	1.630	1.362	1.657	1.337	1.685	1.312	1.714	1.287	1.743	1.262	1.773
90	1.496	1.540	1.474	1.563	1.452	1.587	1.429	1.611	1.406	1.636	1.383	1.661	1.360	1.687	1.336	1.714	1.312	1.741	1.288	1.769
95	1.510	1.552	1.489	1.573	1.468	1.596	1.446	1.618	1.425	1.642	1.403	1.666	1.381	1.690	1.358	1.715	1.336	1.741	1.313	1.767
100	1.522	1.562	1.503	1.583	1.482	1.604	1.462	1.625	1.441	1.647	1.421	1.670	1.400	1.693	1.378	1.717	1.357	1.741	1.335	1.765
150	1.611	1.637	1.598	1.651	1.584	1.665	1.571	1.679	1.557	1.693	1.543	1.708	1.530	1.722	1.515	1.737	1.501	1.752	1.486	1.767
200	1.664	1.684	1.653	1.693	1.643	1.704	1.633	1.715	1.623	1.725	1.613	1.735	1.603	1.746	1.592	1.755	1.582	1.768	1.571	1.779
<i>n</i>	<i>k'</i> = 11		<i>k'</i> = 12		<i>k'</i> = 13		<i>k'</i> = 14		<i>k'</i> = 15		<i>k'</i> = 16		<i>k'</i> = 17		<i>k'</i> = 18		<i>k'</i> = 19		<i>k'</i> = 20	
	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>
16	0.060	3.446	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
17	0.084	3.286	0.053	3.506	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
18	0.113	3.146	0.075	3.358	0.047	3.557	—	—	—	—	—	—	—	—	—	—	—	—	—	—
19	0.145	3.023	0.102	3.227	0.067	3.420	0.043	3.601	—	—	—	—	—	—	—	—	—	—	—	—
20	0.178	2.914	0.131	3.109	0.092	3.297	0.061	3.474	0.038	3.639	—	—	—	—	—	—	—	—	—	—
21	0.212	2.817	0.162	3.004	0.119	3.185	0.084	3.358	0.055	3.521	0.035	3.671	—	—	—	—	—	—	—	—
22	0.246	2.729	0.194	2.909	0.148	3.084	0.109	3.252	0.077	3.412	0.050	3.562	0.032	3.700	—	—	—	—	—	—
23	0.281	2.651	0.227	2.822	0.178	2.991	0.136	3.155	0.100	3.311	0.070	3.459	0.046	3.597	0.029	3.725	—	—	—	—
24	0.315	2.580	0.260	2.744	0.209	2.906	0.165	3.065	0.125	3.218	0.092	3.363	0.065	3.501	0.043	3.629	0.027	3.747	—	—

<i>n</i>	<i>k'</i> = 11		<i>k'</i> = 12		<i>k'</i> = 13		<i>k'</i> = 14		<i>k'</i> = 15		<i>k'</i> = 16		<i>k'</i> = 17		<i>k'</i> = 18		<i>k'</i> = 19		<i>k'</i> = 20	
	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>	<i>dL</i>	<i>dU</i>
25	0.348	2.517	0.292	2.674	0.240	2.829	0.194	2.982	0.152	3.131	0.116	3.274	0.085	3.410	0.060	3.538	0.039	3.657	0.025	3.766
26	0.381	2.460	0.324	2.610	0.272	2.758	0.224	2.906	0.180	3.050	0.141	3.191	0.107	3.325	0.079	3.452	0.055	3.572	0.036	3.682
27	0.413	2.409	0.356	2.552	0.303	2.694	0.253	2.836	0.208	2.976	0.167	3.113	0.131	3.245	0.100	3.371	0.073	3.490	0.051	3.602
28	0.444	2.363	0.387	2.499	0.333	2.635	0.283	2.772	0.237	2.907	0.194	3.040	0.156	3.169	0.122	3.294	0.093	3.412	0.068	3.524
29	0.474	2.321	0.417	2.451	0.363	2.582	0.313	2.713	0.266	2.843	0.222	2.972	0.182	3.098	0.146	3.220	0.114	3.338	0.087	3.450
30	0.503	2.283	0.447	2.407	0.393	2.533	0.342	2.659	0.294	2.785	0.249	2.909	0.208	3.032	0.171	3.152	0.137	3.267	0.107	3.379
31	0.531	2.248	0.475	2.367	0.422	2.487	0.371	2.609	0.322	2.730	0.277	2.851	0.234	2.970	0.196	3.087	0.160	3.201	0.128	3.311
32	0.558	2.216	0.503	2.330	0.450	2.446	0.399	2.563	0.350	2.680	0.304	2.797	0.261	2.912	0.221	3.026	0.184	3.137	0.151	3.246
33	0.585	2.187	0.530	2.296	0.477	2.408	0.426	2.520	0.377	2.633	0.331	2.746	0.287	2.858	0.246	2.969	0.209	3.078	0.174	3.184
34	0.610	2.160	0.556	2.266	0.503	2.373	0.452	2.481	0.404	2.590	0.357	2.699	0.313	2.808	0.272	2.915	0.233	3.022	0.197	3.126
35	0.634	2.136	0.581	2.237	0.529	2.340	0.478	2.444	0.430	2.550	0.383	2.655	0.339	2.761	0.297	2.865	0.257	2.969	0.221	3.071
36	0.658	2.113	0.605	2.210	0.554	2.310	0.504	2.410	0.455	2.512	0.409	2.614	0.364	2.717	0.322	2.818	0.282	2.919	0.244	3.019
37	0.680	2.092	0.628	2.186	0.578	2.282	0.528	2.379	0.480	2.477	0.434	2.576	0.389	2.675	0.347	2.774	0.306	2.872	0.268	2.969
38	0.702	2.073	0.651	2.164	0.601	2.256	0.552	2.350	0.504	2.445	0.458	2.540	0.414	2.637	0.371	2.733	0.330	2.828	0.291	2.923
39	0.723	2.055	0.673	2.143	0.623	2.232	0.575	2.323	0.528	2.414	0.482	2.507	0.438	2.600	0.395	2.694	0.354	2.787	0.315	2.879
40	0.744	2.039	0.694	2.123	0.645	2.210	0.597	2.297	0.551	2.386	0.505	2.476	0.461	2.566	0.418	2.657	0.377	2.748	0.338	2.838
45	0.835	1.972	0.790	2.044	0.744	2.118	0.700	2.193	0.655	2.269	0.612	2.346	0.570	2.424	0.528	2.503	0.488	2.582	0.448	2.661
50	0.913	1.925	0.871	1.987	0.829	2.051	0.787	2.116	0.746	2.182	0.705	2.250	0.665	2.318	0.625	2.387	0.586	2.456	0.548	2.526
55	0.979	1.891	0.940	1.945	0.902	2.002	0.863	2.059	0.825	2.117	0.786	2.176	0.748	2.237	0.711	2.298	0.674	2.359	0.637	2.421
60	1.037	1.865	1.001	1.914	0.965	1.964	0.929	2.015	0.893	2.067	0.857	2.120	0.822	2.173	0.786	2.227	0.751	2.283	0.716	2.338
65	1.087	1.845	1.053	1.889	1.020	1.934	0.986	1.980	0.953	2.027	0.919	2.075	0.886	2.123	0.852	2.172	0.819	2.221	0.786	2.272
70	1.131	1.831	1.099	1.870	1.068	1.911	1.037	1.953	1.005	1.995	0.974	2.038	0.943	2.082	0.911	2.127	0.880	2.172	0.849	2.217
75	1.170	1.819	1.141	1.856	1.111	1.893	1.082	1.931	1.052	1.970	1.023	2.009	0.993	2.049	0.964	2.090	0.934	2.131	0.905	2.172
80	1.205	1.810	1.777	1.844	1.150	1.878	1.122	1.913	1.094	1.949	1.066	1.984	1.039	2.022	1.011	2.057	0.983	2.097	0.955	2.135
85	1.236	1.803	1.210	1.834	1.184	1.866	1.158	1.898	1.132	1.931	1.106	1.965	1.080	1.999	1.053	2.033	1.027	2.068	1.000	2.104
90	1.264	1.798	1.240	1.827	1.215	1.856	1.191	1.886	1.166	1.917	1.141	1.948	1.116	1.979	1.091	2.012	1.066	2.044	1.041	2.077
95	1.290	1.793	1.267	1.821	1.244	1.848	1.221	1.876	1.197	1.905	1.174	1.934	1.150	1.963	1.126	1.993	1.102	2.023	1.079	2.054
100	1.314	1.790	1.292	1.816	1.270	1.841	1.248	1.868	1.225	1.895	1.203	1.922	1.181	1.949	1.158	1.977	1.136	2.006	1.113	2.034
150	1.473	1.783	1.458	1.799	1.444	1.814	1.429	1.830	1.414	1.847	1.400	1.863	1.385	1.880	1.370	1.897	1.355	1.913	1.340	1.931
200	1.561	1.791	1.550	1.801	1.539	1.813	1.528	1.824	1.518	1.836	1.507	1.847	1.495	1.860	1.484	1.871	1.474	1.883	1.462	1.896

**Kaynak:** N. E. Savin, K. J. White, "The Durbin-Watson Test for Serial Correlation with Extreme Sample Sizes of Many Regressors", *Econometrica*, Vol. XLV, 1977, s. 1992-1995

## ÖZGEÇMİŞ

**SONA MARDİKYAN**

[mardikya@boun.edu.tr](mailto:mardikya@boun.edu.tr)

### KİŞİSEL BİLGİLER:

**Doğum Yeri:** İstanbul

**Doğum Tarihi:** 5/5/1966

**Medeni Durum:** Evli

**Milliyeti:** T.C.

### EĞİTİM:

**Lise:** Özel Esayan Ermeni Kız Lisesi 1980-1983

**Lisans:** İTÜ-Kontrol ve Bilgisayar Mühendisliği 1983-1987

**Yüksek Lisans:** Boğaziçi Üniversitesi-Endüstri Mühendisliği 1987-1991

**YABANCI DİL:** İngilizce

### AKADEMİK GÖREVLER:

- Boğaziçi Üniversitesi, Uygulamalı Bilimler Yüksekokulu, Yönetim Bilişim Sistemleri Bölümü  
Tam Zamanlı Öğretim Görevlisi  
1997-...
- Boğaziçi Üniversitesi, Meslek Yüksekokulu, Bilgisayar Programcılığı Bölümü  
Tam Zamanlı Öğretim Görevlisi  
1995-1997

### İDARİ GÖREVLER:

- Boğaziçi Üniversitesi, Bilgisayar Programcılığı Sertifika Programı  
Proje Koordinatörü  
1998-2003

### **VERDİĞİ DERSLER:**

- UBYO, Yönetim Bilişim Sistemleri
  - MIS 123      Introductory Computing for Business
  - MIS 234      Introduction to Database
  - MIS 313      Quantitative Analysis for Decision Making
  - INTT 123     Introductory Computing for Business
  - TRM 154      Introductory Computing for Business
  
- MYO, Bilgisayar Programcılığı
  - COMP 171    Introduction to EDP
  - COMP 173    Programming Languages I (COBOL)
  - COMP 178    Computer Architecture
  - COMP 272    Operations Research
  - COMP 278    Numerical Methods
  - COMP 280    Term Project
  
- Bilgisayar Programcılığı Sertifika Programı
  - BS 171      Bilgisayara Giriş
  - BS 173      COBOL Programlama Dili
  - BS 175      Veri Tabanı Programlama (ACCESS)
  - BS 270      C Programlama Dili
  - BS 280      Proje
  
- İşletmelerde Bilişim Sistemleri Sertifika Programı
  - İBS 514      Veri Analizi
  
- Topluma Girişimci Bireyler Kazandırılması Programı
  - Veritabanı

### **PROFESYONEL TECRÜBE:**

- NCR Bilgi İşlem Sistemleri,  
Sistem Destek Elemanı,  
Ağustos 1991- Mart 1993