

FEN BİLİMLERİ ENSTİTÜSÜ

(YÜKSEK LİSANS TEZİ)

**PARAMETRE TAHMİNİNDE KULLANILAN
İTERASYON YÖNTEMLERİNİN
KARŞILAŞTIRILMASI**

Nalan MUTLU

Tez Danışmanı: Doç. Dr. Hakan Savaş SAZAK

İstatistik Anabilim Dalı


Sunum Tarihi: 09.07.2018


Bornova-İZMİR


2018

Nalan MUTLU tarafından yüksek lisans tezi olarak sunulan "Parametre Tahmininde Kullanılan İterasyon Yöntemlerinin Karşılaştırılması" başlıklı bu çalışma E.Ü. Lisansüstü Eğitim ve Öğretim Yönetmeliği ile E.Ü. Fen Bilimleri Enstitüsü Eğitim ve Öğretim Yönergesi'nin ilgili hükümleri uyarınca tarafımızdan değerlendirilerek savunmaya değer bulunmuş ve 09.07.2018 tarihinde yapılan tez savunma sınavında aday oybirliği/oyçokluğu ile başarılı bulunmuştur.

Jüri Üyeleri:İmza

Jüri Başkanı : Doç. Dr. Hakan Sarac SARAC 

Raportör Üye : Dr. Öğr. Ü. Burcu AYTAÇOĞLU 

Üye : Doç. Dr. A. Fırat ÖZLEMİR 

EGE ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ

ETİK KURALLARA UYGUNLUK BEYANI

E.Ü. Lisansüstü Eğitim ve Öğretim Yönetmeliğinin ilgili hükümleri uyarınca Yüksek Lisans Tezi olarak sunduğum "Parametre Tahmininde Kullanılan İterasyon Yöntemlerinin Karşılaştırılması" başlıklı bu tezin kendi çalışmam olduğunu, sunduğum tüm sonuç, doküman, bilgi ve belgeleri bizzat ve bu tez çalışması kapsamında elde ettiğimi, bu tez çalışmasıyla elde edilemeyen bütün bilgi ve yorumlara atıf yaptığımı ve bunları kaynaklar listesinde usulüne uygun olarak verdiğimi, tez çalışması ve yazımı sırasında patent ve telif haklarını ihlal edici bir davranışımın olmadığını, bu tezin herhangi bir bölümünü bu üniversite veya diğer bir üniversitede başka bir tez çalışması içinde sunmadığımı, bu tezin planlanmasından yazımına kadar bütün sayfalarında bilimsel etik kurallarına uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul edeceğimi beyan ederim.

09/07/2018



Nalan MUTLU

ÖZET

PARAMETRE TAHMİNİNDE KULLANILAN İTERASYON YÖNTEMLERİNİN KARŞILAŞTIRILMASI

MUTLU, Nalan

Yüksek Lisans Tezi, İstatistik Anabilim Dalı

Tez Danışmanı: Doç. Dr. Hakan Savaş SAZAK

Temmuz 2018, 36 sayfa

Bu tezde genel doğrusal regresyon modellerindeki parametreleri tahmin edebilmek için kullanılan iterasyon yöntemlerinin simülasyon yoluyla karşılaştırılması amaç edinilmiştir. Kullanılan yöntemlerin en küçük kareler (LS; Least Squares) yöntemine göre göreceli etkinlik değerleri hesaplanmış ve tablolar halinde düzenlenerek yorumlanmıştır. Çalışmaya Huber M, Tukey M, S ve MM yöntemleri dahil edilmiştir. Tüm yöntemlerin programları Matlab programı kullanılarak oluşturulmuştur. Simülasyon sonuçlarında beklenildiği gibi MM tahmin edicileri genel anlamda en yüksek etkinliğe sahip tahmin ediciler olarak bulunmuştur. S tahmin edicileri de yine beklenildiği gibi diğer iteratif yöntemlere göre daha düşük etkinliğe sahip çıkmıştır. LS yöntemi de doğal olarak normal dağılımda etkinken diğer modellerde düşük etkinliğe sahiptir. Çalışmada çıkan en önemli sonuçlardan biri de hata teriminin varyansını tahminlemede ortaya çıkmıştır. LS tahmin edicisi yansızken diğer yöntemlerin yan ürettiği görülmektedir. MM varyans tahmin edicisi diğerlerinden farklı olarak yan üretmekle beraber daha küçük miktarlarda yan değerlerine sahiptir ve asimptotik olarak yan azalması konusunda daha iyi cevap vermektedir. Çalışmanın sonunda tüm tahmin edicilerin değerlerinin verildiği iki gerçek hayat uygulaması verilmiştir. Her iki uygulamada da bir çok kriter ve grafikten istifade edilmiştir.

Anahtar kelimeler: Genel doğrusal regresyon modeli, iterasyon, M tahmin edicileri, MM tahmin edicileri, en küçük kareler, dayanıklılık.

ABSTRACT

THE COMPARISON OF THE ITERATION METHODS USED IN PARAMETER ESTIMATION

MUTLU, Nalan

MSc in Statistics

Supervisor: Assoc. Prof. Dr. Hakan Savaş SAZAK

July 2018, 36 pages

In this thesis the comparison of the iteration methods for the estimation of the parameters of the general linear regression models via simulation is aimed. The relative efficiencies of the used methods with respect to the least squares (LS) method are calculated and interpreted while being formatted as tables. The Huber M, the Tukey M, the S and the MM were included in the study. The programs of all of the methods were written by using Matlab. As expected, in general, the MM estimators are found to be the most efficient estimators depending on the simulation results. Again as expected, the S estimators are found to be less efficient than the other iterative methods. The LS estimators are while naturally being more efficient under normality than the other methods, they are less efficient under the other models. In this study, one of the most important results appeared in the estimation of the variance of the error term. While the LS estimator of the variance of the error term is unbiased, the other estimators produce bias. The MM estimator of the variance is, as a different feature from the others while being biased, less biased and asymptotically it responses better in reducing the bias. At the end of the study, two real life data examples were given with the estimated values of each method. For both examples, various criteria and graphs were utilized.

Keywords: General linear regression model, iteration, M estimators, MM estimators, least squares, robustness.



TEŞEKKÜR

Bu çalışmanın gerçekleştirilmesinde değerli katkıları olan ve tez çalışmamın konusunu öneren Doç. Dr. Hakan Savaş SAZAK'a, bu süreçte her zaman yanımda olan değerli aileme ve arkadaşlarıma teşekkürü bir borç bilirim.

İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET	v
ABSTRACT	vi
TEŞEKKÜR	vii
ŞEKİLLER DİZİNİ	x
TABLolar DİZİNİ.....	xi
1. GİRİŞ.....	1
2. GENEL BİLGİLER.....	3
2.1 Genel Doğrusal Regresyon Modeline Yönelik Varsayımlar	3
2.2 En Küçük Kareler (LS) Yöntemi.....	3
2.3 Ağırlıklı En Küçük Kareler (WLS)	4
2.4 Ardışık Ağırlıklı En Küçük Kareler (IRWLS)	4
2.5 En Küçük Medyan Kareler (LMS) Yöntemi	5
3. KARŞILAŞTIRILAN YÖNTEMLER.....	6
3.1 M Tahmin Edicileri	6
3.2 S Tahmin Edicileri.....	9

İÇİNDEKİLER (devam)

	<u>Sayfa</u>
3.3 MM Tahmin Edicileri.....	11
4. SİMÜLASYONLAR.....	13
5. GERÇEK HAYAT VERİSİ ÜZERİNE UYGULAMALAR.....	26
5.1 Uygulama 1.....	26
5.2 Uygulama 2.....	29
6. SONUÇ VE TARTIŞMA.....	33
KAYNAKLAR DİZİNİ.....	34
ÖZGEÇMİŞ.....	36

ŞEKİLLER DİZİNİ

<u>Şekil</u>	<u>Sayfa</u>
Şekil 5.1 Uygulama 1 için artık değerlerin kutu grafiği	27
Şekil 5.2 Uygulama 1 için q-q grafiği	27
Şekil 5.3 Uygulama 1 için studentize artık değer - kaldıraç grafiği	28
Şekil 5.4 Uygulama 1 için tutturulan Y – studentize artık değer grafiği.....	28
Şekil 5.5 Uygulama 2 için kutu grafiği.....	30
Şekil 5.6 Uygulama 2 için q-q grafiği.....	30
Şekil 5.7 Uygulama 2 için studentize artık değer – kaldıraç grafiği	31
Şekil 5.8 Uygulama 2 için tutturulan Y - studentize artık değer grafiği	31

TABLOLAR DİZİNİ

Tablo	Sayfa
Tablo 3.1 Amaç fonksiyonları özet tablo.....	8
Tablo 3.2 S tahmin edicisi için farklı sabitlere göre etkinlik ve BDP değerleri....	11
Tablo 4.1 $X \sim N(0,1)$, $\varepsilon \sim N(0,1)$, $n = 30$ için simülasyon sonuçları.....	14
Tablo 4.2 $X \sim N(0,1)$, $\varepsilon \sim N(0,1)$, $n = 50$ için simülasyon sonuçları.....	15
Tablo 4.3 $X \sim N(0,1)$, $\varepsilon \sim N(0,1)$, $n = 100$ için simülasyon sonuçları.....	16
Tablo 4.4 $X = 1, \dots, n$, $\varepsilon \sim N(0,1)$, $n = 30$ için simülasyon sonuçları.....	17
Tablo 4.5 $X = 1, \dots, n$, $\varepsilon \sim N(0,1)$, $n = 50$ için simülasyon sonuçları.....	18
Tablo 4.6 $X = 1, \dots, n$, $\varepsilon \sim N(0,1)$, $n = 100$ için simülasyon sonuçları.....	19
Tablo 4.7 $X = 1, \dots, n$, $\varepsilon \sim \pi N(0, 3^2) + (1 - \pi)N(0,1)$, $\pi = 0.1$, $n = 30$ için simülasyon sonuçları.....	20
Tablo 4.8 $X = 1, \dots, n$, $\varepsilon \sim \pi N(0, 3^2) + (1 - \pi)N(0,1)$, $\pi = 0.1$, $n = 50$ için simülasyon sonuçları.....	21
Tablo 4.9 $X = 1, \dots, n$, $\varepsilon \sim \pi N(0, 3^2) + 1 - \pi N(0,1)$, $\pi = 0.1$, $n = 100$ için simülasyon sonuçları	22
Tablo 4.10 $X = 1, \dots, n$, $\varepsilon \sim r \text{ tane } N(0, 3^2) + n - r \text{ tane } N(0,1)$, $n = 30$ için simülasyon sonuçları.....	23
Tablo 4.11 $X = 1, \dots, n$, $\varepsilon \sim r \text{ tane } N(0, 3^2) + n - r \text{ tane } N(0,1)$, $n = 50$ için simülasyon sonuçları.....	24

TABLOLAR DİZİNİ (DEVAM)

<u>Tablo</u>	<u>Sayfa</u>
Tablo 4.12 $X = 1, \dots, n$, $\varepsilon \sim r$ tane $N(0,3^2) + n - r$ tane $N(0,1)$, $n = 100$ için simülasyon sonuçları.....	25
Tablo 5.1 Uygulama 1 için tahmin değerleri.....	26
Tablo 5.2 Uygulama 2 için tahmin değerleri.....	29



1. GİRİŞ

Regresyon analizi hemen hemen tüm bilim alanlarında kullanılan popüler bir analiz yöntemidir. Regresyon analizinin amacı bağımlı ve bağımsız değişkenlerin arasındaki ilişkinin yapısının incelenmesi ve oluşturulan model kullanılarak yorum ve tahminlerde bulunulmasıdır.

Modelin oluşturulması ve modelin doğrulanması için çeşitli yöntemler vardır. Verilerin kullanıldığı alanda daha önce önerilen modeller yoksa genelde genel doğrusal modeller kullanılarak modeldeki parametreler tahmin edilir ve modelin doğruluğu araştırılır. Bu çalışmada aşağıdaki genel doğrusal regresyon modeli kullanılmıştır.

$$Y_{nx1} = X_{nx(p+1)}\beta_{(p+1)x1} + \varepsilon_{nx1} \quad (1.1)$$

Denklem (1.1)'de verilen modelde Y bağımlı değişken vektörü, X bağımsız değişken matrisi, β parametre vektörü, ε ise hata terimi vektörüdür. Burada n örneklem hacmi, p ise eğim parametresi sayısıdır.

Parametre tahmini için bilinen en kullanışlı ve hesaplanması en kolay yöntem en küçük kareler (LS; Least Squares) yöntemidir. LS yöntemi hata kareler toplamını minimize eder ve Gauss tarafından 1809'da önerilmiştir. LS yöntemi kolay anlaşılabilir ve pratik bir yöntemdir ancak aykırı değerlere karşı aşırı hassastır. Bu sorunun çözümü için çeşitli yöntemler önerilmiştir. Bu yöntemlerden bir tanesi de dayanıklı regresyon yöntemidir.

Dayanıklı regresyon aykırı değerlere sahip bir veri topluluğunda daha dayanıklı tahmin ediciler elde etmemizi sağlar. Bir tahmin edicinin dayanıklılığını ölçmek için çeşitli yöntemler vardır. Bunlardan ikisi kırılma noktası (BDP; Breakdown Point) ve etki fonksiyonudur (Influence function). BDP, bir örneklemde tahmin ediciyi sonsuz (veya eksi sonsuz) yapmadan örneklemdeki sonsuza (veya eksi sonsuza) gönderilebilecek gözlem oranıdır. Donoho ve Huber (1983)'ün tanımına göre dayanıklı bir tahmin edici yüksek bir BDP değerine sahip olmalıdır ve bir tahmin edicinin sahip olabileceği en yüksek BDP değeri %50'dir. Bu değer üstünde bir değere sahip olan tahmin ediciler aykırı değerler ve normal gözlemler arasında ayırım yapamamaktadır.

LS yönteminin BDP değeri sıfırdır ve aykırı değerlerden çok kolay etkilendiğinden dayanıklı bir yöntem değildir. Literatürde dayanıklı ve etkin $\hat{\beta}$ tahmin edicileri olarak genellikle M tahmin edicileri kullanılır. Huber (1973) tarafından önerilmiş olan M tahmin edicileri her ne kadar daha sonra tanımını vereceğimiz amaç fonksiyonuna bağlı olsa da diğer dayanıklı tahmin edicilere göre daha düşük BDP değeri ve çok yüksek olmayan bir etkinlik değeri sunarlar. Literatürdeki gelişmeler dahilinde Rousseeuw (1984) tarafından tanıtilen en küçük medyan kareler (LMS; Least Median Square) yöntemi yüksek BDP değerine sahip dayanıklı tahmin yöntemi olarak karşımıza çıkar. Ancak bu yöntem hataların medyan değerini en küçük yapmayı amaçlarken geriye kalan $(n - 1)$ adet gözlemi dikkate almadığından asimptotik olarak iyi özellikler göstermez (Ryan, 1997). LMS yönteminin genelleştirilmiş hali olan S tahmin edicileri Rousseeuw ve Yohai (1984)

tarafından yüksek BDP'ye sahip dayanıklı tahmin ediciler olarak tanıtılmıştır. Ancak düşük etkinlik değerine sahip olduğundan yeni bir tahmin edici arayış sürecine her iki özelliği de içeren MM tahmin edicileri son vermiştir. MM tahmin edicileri Yohai (1987) tarafından literatüre kazandırılmış yüksek BDP ve yüksek etkinlik değerine sahip tahmin edicilerdir.

Bu çalışmada genel doğrusal regresyon modelindeki parametrelerin tahmini için kullanılan ve iterasyon içeren Huber M, Tukey M, S ve MM yöntemlerinin LS yöntemiyle çeşitli modeller altında simülasyon yoluyla karşılaştırılması amaç edinilmiştir. Bunu gerçekleştirmek için Matlab ortamında çeşitli programlar yazılmıştır. Çalışmanın sonunda iki gerçek hayat veri uygulaması verilmiştir.



2. GENEL BİLGİLER

2.1 Genel Doğrusal Regresyon Modeline Yönelik Varsayımlar

Denklem (1.1)'de verilen genel doğrusal regresyon modeline yönelik aşağıdaki varsayımlar yapılır.

1. $rank(X_{nx(p+1)}) = p + 1$ yani model tam ranklıdır.
2. $E(\varepsilon_{nx1}) = 0$
3. $Var(\varepsilon_{nx1}) = \sigma^2 I_{n \times n}$

Madde 3'te verilen varsayım tek bir tanımla varyans sabitliğini ve bağımsızlığı göstermektedir (normal dağılımda). Bazı durumlarda madde 3 aşağıdaki şekilde de ifade edilebilir. Bu durumda hata terimlerinin farklı varyanslara sahip oldukları ve birbirleriyle ilişkili olduğu kabul edilir fakat bu durum tahminleme aşamasını daha karmaşık hale getirecektir.

$$Var(\varepsilon_{nx1}) = \Sigma_{n \times n}$$

2.2 En Küçük Kareler (LS) Yöntemi

LS yöntemi, regresyonda parametre tahmininde kullanılan yaygın bir yöntemdir. Hata kareleri toplamını minimize eder ve Gauss tarafından 1809'da bulunmuştur.

$$\min \sum_{i=1}^n \varepsilon_i^2 = \min \sum_{i=1}^n (y_i - x_i^t \beta)^2 \quad (2.2.1)$$

2.2.1 denkleminin bilinmeyen parametrelere göre türevi alınıp sıfıra eşitlenir.

$$\frac{\partial}{\partial \beta_j} \sum_{i=1}^n (y_i - x_i^t \beta)^2 = 0 \quad , j = 1, 2, \dots, (p + 1) \quad (2.2.2)$$

2.2.2 denkleminin çözümü ile $\hat{\beta}$ tahmin edicileri elde edilir.

LS yöntemi anlaşılması ve uygulaması kolay bir yöntemdir ancak aykırı değerlere karşı hassasiyeti bilinmemektedir.

2.3 Ağırlıklı En Küçük Kareler (WLS)

Ağırlıklı en küçük kareler (WLS; Weighted Least Squares) yönteminde ağırlıklı hata kareler toplamı minimize edilir. Ağırlıklar bilinir ya da $\omega_i = 1/\sigma_i^2$ formülüyle tahmin edilir.

$$\min \sum_{i=1}^n \omega_i \varepsilon_i^2 \quad (2.3.1)$$

$W_{n \times n} = \text{diag}(\{\omega_i: i = 1, \dots, n\})$ olmak üzere, matris formatıyla aşağıdaki şekilde de ifade edilebilir.

$$\min \varepsilon' W \varepsilon \quad (2.3.2)$$

β parametre vektörüne göre türev alınmasıyla WLS tahmin edicileri denklem 2.3.3' te gösterildiği şekilde elde edilir.

$$\hat{\beta} = (X^t W X)^{-1} (X^t W Y) \quad (2.3.3)$$

2.4 Ardışık Ağırlıklı En Küçük Kareler (IRWLS)

Regresyonda M tahmin edicileri genellikle tek bir adımda elde edilemez ve bu nedenle denklem çözümleri için iteratif bir sürece ihtiyaç duyulur. İhtiyaç duyulan bu iteratif süreç ardışık ağırlıklı en küçük kareler (IRWLS; Iteratively Reweighted Least Squares) olarak adlandırılır ve Andersen (2008) tarafından aşağıdaki gibi tanımlanmıştır.

Adım 1: LS yöntemiyle başlangıç tahmini $\hat{\beta}^{(0)}$ elde edilir.

Adım 2: LS yöntemiyle elde edilen tahminden artık değerler $e_i^{(0)}$ elde edilir. Bu artık değerlerle $\hat{\sigma}^{(0)}$ başlangıç tahmini, medyan mutlak sapma (MAD; Median Absolute Deviation) kullanılarak hesaplanır.

$$\hat{\sigma}^{(0)} = 1.4826 * MAD$$

$$MAD(e) = \text{medyan}|e_i - \text{medyan}(e_i)|$$

Adım 3: Artık değerler $u = \frac{e_i^{(0)}}{\hat{\sigma}^{(0)}}$ olarak standardize edilir ve ağırlıklar u kullanılarak $\omega_i^{(0)}$ olarak hesaplanır.

Adım 4: Ağırlıklı en küçük kareler yöntemiyle $\sum_{i=1}^n \omega_i^{(0)} (\varepsilon_i^{(0)})^2$ ifadesi minimize edilerek $\hat{\beta}^1$ elde edilir. Matris formatında ise

$$\hat{\beta}^1 = (X^t W^{(0)} X)^{-1} (X^t W^{(0)} Y)$$

şeklinde gösterilir. Burada

$$W_{n \times n} = \text{diag}(\{\omega_i: i = 1, \dots, n\})$$

Adım 5: $\hat{\beta}^{(1)}$ kullanılarak yeni artık değerler $e_i^{(1)}$ ve sigma değeri $\hat{\sigma}^{(1)}$ olarak elde edilir. Bunlara bağlı olarak yeni ağırlıklar $W^{(1)}$ olarak elde edilir.

Adım 6: Yeni ağırlıklar kullanılarak ağırlıklı en küçük kareler yöntemiyle $\hat{\beta}^{(2)}$ tahmini elde edilir.

Adım 7: Yeni tahmin ve bir önceki tahmin arasında istenilen yakınsama sağlanıncaya kadar adım 4 ve 6 arasında iterasyona devam edilir.

Yakınsamanın sağlanıp sağlanmadığını görmek için

$$\frac{\|\hat{\beta}^{(q+1)} - \hat{\beta}^{(q)}\|}{\|\hat{\beta}^{(q+1)}\|} < \delta \quad (2.4.1)$$

kriteri kontrol edilir.

2.5 En Küçük Medyan Kareler (LMS) Yöntemi

LMS yöntemi ilk olarak Rousseeuw (1984) tarafından bulunmuştur. Parametre tahminlerini hata karelerinin medyanını minimize ederek gerçekleştirir.

$$\min \text{median} (y_i - x_i^t \beta)^2 \quad (2.5.1)$$

LMS yöntemi dayanıklı bir tahmin yöntemidir. Yüksek BDP değerine sahiptir ancak iyi asimptotik özellikler göstermez.

3. KARŞILAŞTIRILAN YÖNTEMLER

3.1 M Tahmin Edicileri

Dayanıklı regresyonda kullanılan M tahmin edicileri ilk olarak Huber (1973) tarafından bulunmuştur. En küçük kareler yönteminde kullanılan hata kareler toplamını minimize etme ilkesi yerine uygun bir ρ fonksiyonunun toplamını minimize etme ilkesine dayanır. Bu değişimi yapmaktaki amaç aykırı değerlere karşı duyarlılığı azaltmaktır. En çok olabilirlik (ML; Maximum Likelihood) tahmin yönteminde de $\hat{\beta}$ yi tahmin edebilmek için en çok olabilirlik fonksiyonu maksimize edilir.

$$\prod_{i=1}^n f(e_i) = \prod_{i=1}^n f(y_i - x_i^t \beta) \quad (3.1.1)$$

Matematiksel olarak daha kolay olduğundan log olabilirlik fonksiyonunu maksimize eder.

$$\sum_{i=1}^n \ln f(e_i) = \sum_{i=1}^n \ln f(y_i - x_i^t \beta) \quad (3.1.2)$$

(3.1.2) denklemini maksimize etmek

$$\sum_{i=1}^n -\ln f(y_i - x_i^t \beta) \quad (3.1.3)$$

(3.1.3) denklemini minimize etmek anlamına gelir.

Yani ML yöntemi, M tahmin yönteminin özel bir durumu olduğundan maksimizenin baş harfi olan m harfi kullanılarak M tahmin edicilerinin ismi oluşturulmuştur. Burada

$$\rho(y_i - x_i^t \beta) = -\ln f(y_i - x_i^t \beta) \text{ dir.}$$

ρ , sıfır noktasında tek bir minimum noktaya sahip, sürekli, türevlenebilir ve simetrik bir fonksiyondur. (Leroy, 1987; Andersen, 2008)

$\hat{\beta}$, M tahmin edicisi

$$\min \sum_{i=1}^n \rho(e_i) = \min \sum_{i=1}^n \rho(y_i - x_i^t \beta) \quad (3.1.4)$$

ifadesini sağlayan en uygun β değeri olarak ya da $\psi = \rho'$ olmak üzere

$$\sum_{i=1}^n \psi(y_i - x_i^t \beta) x_i = 0 \quad (3.1.5)$$

denkleminin çözümünden elde edilir.

(3.1.5) denkleminde denk olarak

$$\sum_{i=1}^n \omega_i (y_i - x_i^t \beta) x_i = 0 \quad (3.1.6)$$

denklemini yazılabilir.

(3.1.6) denkleminde ω_i ağırlık fonksiyonudur ve

$$\omega_i = \frac{\psi(y_i - x_i^t \beta)}{(y_i - x_i^t \beta)} \quad (3.1.7)$$

şeklinde ifade edilir.

M tahmin edicilerinin ölçek değişmezlik özelliğini korumak amacıyla bir σ ölçeği ile β 'nin

$$\sum_{i=1}^n \psi \left(\frac{y_i - x_i^t \beta}{\hat{\sigma}} \right) x_i = 0 \quad (3.1.8)$$

denklemler sistemi eş zamanlı tahmin edilir.

M tahmin edicilerinde kullanılacak birden fazla ρ fonksiyonu vardır. Bunlar Tablo 3.1'de özetlenmiştir (Türkay, 2004).

Tablo 3.1 Amaç fonksiyonları özet tablo

TAHMİN EDİCİLER	AMAÇ FONKSİYONU $\rho(r)$	ETKİ FONKSİYONU $\Psi(r)$	AĞIRLIK FONKSİYONU	TANIM ARALIĞI
HUBER $c > 0$	$(1/2)r^2$	r	1	$ r \leq c$
	$c(r - (1/2)c)$	$c \operatorname{sgn}(r)$	$c/ r $	$ r > c$
HAMPEL $a, b, c > 0$	$(1/2)r^2$	r	1	$ r < a$
	$a(r - (1/2)a)$	$a \operatorname{sgn}(r)$	$a/ r $	$a \leq r < b$
	$a \frac{c r - (1/2)r^2}{c-b} - \frac{7a^2}{6}$	$a \frac{c \operatorname{sgn}(r) - r}{c-b}$	$a \frac{c/ r - 1}{c-b}$	$b \leq r \leq c$
	$a(b+c-a)$	0	0	Diğer durumlarda
TUKEY BISQUARE $c > 0$	$\frac{c^2}{6} \left(1 - \left(1 - \left(\frac{r}{c} \right)^2 \right)^3 \right)$	$r \left(1 - \left(\frac{r}{c} \right)^2 \right)^2$	$\left(1 - \left(\frac{r}{c} \right)^2 \right)^2$	$ r \leq c$
	$\frac{1}{6}c^2$	0	0	$ r > c$
EN KÜÇÜK KARELER	$\frac{1}{2}r^2$	r	1	$-\infty \leq r \leq \infty$
FAIR	$c^2 \left[\frac{ r }{c} - \log \left(1 + \frac{ r }{c} \right) \right]$	$\frac{r}{1 + \frac{ r }{c}}$	$\frac{1}{1 + \frac{ r }{c}}$	\mathbb{R}
CAUCHY	$\frac{c^2}{2} \log \left[1 + \left(\frac{r}{c} \right)^2 \right]$	$\frac{r}{1 + \left(\frac{r}{c} \right)^2}$	$\frac{1}{1 + \left(\frac{r}{c} \right)^2}$	\mathbb{R}
ANDREWS	$\frac{1}{\pi^2} (1 - \cos(\pi r))$	$\frac{1}{\pi} \sin(\pi r)$	$\frac{1}{\pi r} \sin(\pi r)$	$ r \leq 1$
	$\frac{2}{\pi^2}$	0	0	$ r > 1$
WELSCH	$\frac{c^2}{2} \log \left[1 + \left(\frac{ r }{c} \right)^2 \right]$	$r \exp \left[- \left(\frac{r}{c} \right)^2 \right]$	$\exp \left[- \left(\frac{r}{c} \right)^2 \right]$	\mathbb{R}

M tahmin edicileri LS yöntemine göre aykırı değerlere karşı daha dayanıklıdır, etkinliği de düşük değildir ancak düşük BDP değerine sahiptir.

Amaç fonksiyonu olan ρ genelde doğrusal bir fonksiyon değildir. Bu nedenle iteratif yöntemlerle çözülür. M tahmin edicilerini elde etmek için IRWLS yöntemi kullanılarak oluşturulan algoritma aşağıdaki gibidir (Susanti et al., 2014).

Algoritma 1:

Adım 1: LS yöntemi kullanılarak $\hat{\beta}^0$ başlangıç tahmini elde edilir.

Adım 2: Yeni tahmine göre artık değerler hesaplanır.

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

Adım 3: Sigma değerleri tahmin edilir.

$$\hat{\sigma}_i = 1.4826 * MAD$$

Adım 4: Artık değerler standardize edilir.

$$u_i = \frac{e_i}{\hat{\sigma}_i}$$

Adım 5: Ağırlıklar hesaplanır.

$$\omega_i = \begin{cases} 1.345/|u_i| & |u_i| > 1.345 \\ 1 & |u_i| \leq 1.345 \end{cases}$$

Adım 6: IRWLS yöntemiyle $\hat{\beta}_M$ elde edilir.

Adım 7: $\hat{\beta}_M$ için istenen yakınsama sağlanıncaya kadar adım 2 ve 7 arasında iterasyona devam edilir.

Bu çalışmada algoritma 1'de Tablo 3.1'de gösterilen ρ fonksiyonlarından Huber'ın ρ fonksiyonu kullanılmıştır ve ağırlıklar bu fonksiyona göre hesaplanmıştır.

3.2 S Tahmin Edicileri

Dayanıklı regresyonda yüksek BDP değerine sahip olan S tahmin edicileri Rousseeuw ve Yohai (1984) tarafından bulunmuştur. Ölçek tahmini temel alınarak hesaplandığından S tahmin edicileri olarak adlandırılır. LMS yönteminin geliştirilmiş halidir (Çetin ve Toka, 2011). Hataların saçılımının dayanıklı ölçümünü minimize eder.

$$\min s(e_i)$$

Rousseeuw ve Leroy (1987) hataların saçılımını

$$\frac{1}{n} \sum_{i=1}^n \rho\left(\frac{e_i}{s}\right) = K \quad , \quad K = \text{sabit}$$

ifadesinin çözümü olarak tanımlamışlardır.

Salibian ve Yohai (2006), ρ fonksiyonunu minimize ederken $\hat{\sigma}$ tahminini

$$\hat{\sigma}_s = \sqrt{\frac{1}{nK} \sum_{i=1}^n \omega_i e_i^2}$$

denklemleri ile ifade etmiştir.

S tahmin edicileri, M tahmin edicileri ile aynı asimptotik özellikleri gösterirler. M tahmin edicisinden farklı olarak yüksek BDP değerine ve düşük etkinliğe sahiptirler. S tahmin edicilerini IRWLS yöntemini kullanarak hesaplamak için kullanılan algoritma aşağıdaki gibidir. (Susanti et al., 2014)

Algoritma 2:

Adım 1: LS yöntemi kullanılarak $\hat{\beta}^0$ başlangıç tahmini elde edilir.

Adım 2: Yeni tahmine göre artık değerleri hesaplanır.

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

Adım 3: Sigma değerleri tahmin edilir.

$$\hat{\sigma}_i = \begin{cases} 1.4826 * MAD & \text{iterasyon} = 1 \\ \sqrt{\frac{1}{nK} \sum_{i=1}^n \omega_i e_i^2} & \text{iterasyon} > 1 \end{cases}$$

Adım 4: Artık değerler standardize edilir.

$$u_i = \frac{e_i}{\hat{\sigma}_i}$$

Adım 5: Ağırlıklar hesaplanır.

$$\omega_i = \begin{cases} \left[1 - \left(\frac{u_i}{1.547} \right)^2 \right]^2 & |u_i| \leq 1.547 & \text{iterasyon} = 1 \\ 0 & |u_i| > 1.547 \\ \frac{\rho(u_i)}{u_i^2} & & \text{iterasyon} > 1 \end{cases}$$

Adım 6: IRWLS yöntemiyle $\hat{\beta}_S$ elde edilir.

Adım 7: $\hat{\beta}_S$ için istenen yakınsama sağlanıncaya kadar adım 2 ve 7 arasında iterasyona devam edilir.

Bu çalışmada algoritma 2'de Tablo 3.1'de gösterilen ρ fonksiyonlarından Tukey'in ρ fonksiyonu kullanılmıştır ve ağırlıklar bu fonksiyona göre hesaplanmıştır.

S tahmin edicileri için farklı sabitlere göre etkinlik ve BDP değerleri Tablo 3.2’de gösterilmiştir.

Tablo 3.2 S tahmin edicisi için farklı sabitlere göre etkinlik ve BDP değerleri

Bdp	etkinlik	c	K
%50	%28.7	1.547	0.1995
%45	%37.0	1.756	0.2312
%40	%46.2	1.988	0.2634
%35	%56.0	2.251	0.2957
%30	%66.1	2.560	0.3278
%25	%75.9	2.973	0.3593
%20	%84.7	3.420	0.3899
%15	%91.7	4.096	0.4194
%12	%95.0	4.685	0.4368
%10	%96.6	5.182	0.4475

3.3 MM Tahmin Edicileri

Dayanıklı regresyonda yüksek BDP değeri ve yüksek etkinliğe sahip olan MM tahmin edicileri Yohai (1987) tarafından bulunmuştur. Bu tahmin edicilerin etkinliğini yüksek tutabilmek amacıyla sabitlerin seçimi uygun bir şekilde yapılır (Maronna and Yohai, 2010). MM tahmin edicileri hesaplanırken başlangıç olarak yüksek BDP değerine sahip bir tahmin yapılır, daha sonra uygun sabit seçimiyle yüksek etkinlik elde edilir. Genelde başlangıç tahmini olarak yüksek BDP değerinden dolayı S tahmin edicileri tercih edilir. Yüksek etkinliğe ulaşabilmek amacıyla da M tahmin edicileriyle devam edilir (Copt and Heritier, 2006). MM tahmin edicilerini IRWLS yöntemini kullanarak hesaplamak için kullanılan algoritma aşağıdaki gibidir. (Susanti et al., 2014)

Algoritma 3:

Adım 1: S tahmin yöntemiyle $\hat{\beta}_0$ başlangıç tahmini elde edilir.

Adım 2: Yeni tahmine göre artık değerler hesaplanır.

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

Adım 3: S tahmin yöntemiyle sigma değerleri tahmin edilir.

$$\hat{\sigma}_i = \hat{\sigma}_s$$

Adım 4: Artık değerler standardize edilir.

$$u_i = \frac{e_i}{\hat{\sigma}_i}$$

Adım 5: Ağırlıklar hesaplanır.

$$\omega_i = \begin{cases} \left[1 - \left(\frac{u_i}{4.685}\right)^2\right]^2 & |u_i| \leq 4.685 \\ 0 & |u_i| > 4.685 \end{cases}$$

Adım 6: IRWLS yöntemiyle $\hat{\beta}_{MM}$ elde edilir.

Adım 7: $\hat{\beta}_{MM}$ için istenen yakınsama sağlanıncaya kadar adım 2 ve 7 arasında iterasyona devam edilir.

Bu çalışmada algoritma 3'te, Tablo 3.1'de gösterilen ρ fonksiyonlarından Tukey'in ρ fonksiyonu kullanılmıştır ve ağırlıklar bu fonksiyona göre hesaplanmıştır.



4. SİMÜLASYONLAR

Bu bölümde simülasyonlar için Matlab programı kullanılarak Huber M, Tukey M, S ve MM tahmin edicilerinin programları yazılmıştır. Programlar tezin başında tanımlanan denklem (1.1)'e uygun yazılmıştır ancak simülasyonlarda yorum kolaylığı sağlamak amacıyla aşağıda verilen basit doğrusal regresyon modeli kullanılmıştır.

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

Modelde β_0 doğrusal denklemin Y eksenini kestiği noktayı, β_1 ise doğrunun eğim parametresini göstermektedir.

Simülasyonlar genel yoruma engel olmayacak bir şekilde $\beta_0 = 0, \beta_1 = 1,$

$n = 30, 50$ ve $100,$ ve $nn = [300000/n]$ için aşağıdaki dağılımlar altında yapılmıştır.

1. İki değişkenli normal dağılım;

$$X \sim N(0,1), \varepsilon \sim N(0,1)$$

2. Tasarımsal $X,$ hata terimi normal dağılım;

$$X = 1, 2, \dots, n, \varepsilon \sim N(0,1)$$

3. Karma tipi aykırı modeli;

$$X = 1, \dots, n, \varepsilon \sim \pi N(0, 3^2) + (1 - \pi)N(0,1), \pi = 0.1$$

4. Dixon aykırı değer modeli;

$$X = 1, \dots, n, \varepsilon \sim r \text{ tane } N(0, 3^2) + (n - r) \text{ tane } N(0,1)$$

Tezde bahsedilen tüm yöntemlerin β_0, β_1 ve σ^2 parametreleri için üretilen tahminlerinin ortalamaları (mean), yanları (bias), varyansları (var) ve hata kareler ortalamaları (mse) bulunmuş ve LS tahmin edicilerine göre göreceli etkinlik değerleri (REFF; Relative Efficiency)

$$\text{REFF}(a|b) = 100 \times \frac{\text{MSE}(b)}{\text{MSE}(a)} \quad (4.1)$$

formülüyle hesaplanmıştır. Tüm sonuçlar Tablo 4.1 - 4.12 ile verilmiş ve yorumlanmıştır.

Tüm iterasyonlarda denklem 2.4.1 ile verilen eşitsizlikte yakınsamanın sağlanıp sağlanmadığını kontrol etmek amacıyla $\delta = 0.00001$ olarak kullanılmıştır. Yorumlarda kolaylık sağlamak amacıyla hata terimi model 3'te $\sqrt{1 - \pi + \pi k^2}$

ifadesine model 4'te ise $\sqrt{1 - \left(\frac{r}{n}\right) + (r/n)k^2}$ ifadesine bölünerek standardize edilmiştir. Bu modeller için $k = 3$ olarak kullanılmıştır.

İki değişkenli normal dağılıma ait simülasyonlar Tablo 4.1 - 4.3 ile verilmiştir.

Tablo 4.1 $X \sim N(0,1)$, $\varepsilon \sim N(0,1)$, $n = 30$ için simülasyon sonuçları

		β_0	β_1	σ^2
mean	LS	0.002225	0.997943	0.995908
	Huber M	0.001901	0.997285	0.941969
	Tukey M	0.002047	0.997202	0.938624
	S	-0.000225	0.995847	0.866131
	MM	0.001760	0.997231	0.935249
bias	LS	0.002225	-0.002057	-0.004092
	Huber M	0.001901	-0.002715	-0.058031
	Tukey M	0.002047	-0.002798	-0.061376
	S	-0.000225	-0.004153	-0.133869
	MM	0.001760	-0.002769	-0.064751
nxvar	LS	1.038991	1.107597	2.142773
	Huber M	1.100392	1.153855	4.957198
	Tukey M	1.130091	1.182679	4.941174
	S	1.456670	1.535583	3.418712
	MM	1.088666	1.146471	1.924377
nxmse	LS	1.039140	1.107724	2.143276
	Huber M	1.100500	1.154076	5.058225
	Tukey M	1.130217	1.182913	5.054183
	S	1.456671	1.536100	3.956337
	MM	1.088758	1.146701	2.050157
REff	LS	100.0000	100.0000	100.0000
	Huber M	94.4243	95.9836	42.3721
	Tukey M	91.9416	93.6437	42.4060
	S	71.3366	72.1128	54.1732
	MM	95.4426	96.6010	104.5421

Tablo 4.1'e göre Huber M, Tukey M, S ve MM tahmin edicilerinin REFF değerleri LS tahmin edicisine göre daha düşüktür. β_0 ve β_1 parametreleri için REFF değerleri ayrıntılı olarak incelendiğinde MM tahmin edicisinin diğer tahmin edicilere göre daha yüksek bir değere sahip olduğu görülmektedir. MM tahmin edicisini sırasıyla Huber M, Tukey M ve S tahmin edicileri takip etmektedir. σ^2 için incelendiğinde ise belli miktarda yan üretmenin verdiği avantajla MM tahmin edicisinin REFF değeri LS tahmin edicisinden daha yüksek çıkmıştır. Bu her ne kadar kulağa garip gelse de ridge regresyon gibi bazı yöntemlerin varyans ve mse değerlerinde azalma sağlamak için özellikle yan ürettiği bilinmektedir. bazı LS

tahmin edicisinden sonra sırasıyla S, Tukey M, Huber M tahmin edicileri gelmektedir.

Tablo 4.2 $X \sim N(0,1)$, $\varepsilon \sim N(0,1)$, $n = 50$ için simülasyon sonuçları

		β_0	β_1	σ^2
mean	LS	0.001981	0.998069	0.996005
	Huber M	0.001661	0.997618	0.965083
	Tukey M	0.001604	0.997849	0.963442
	S	0.000101	0.996793	0.909776
	MM	0.001685	0.997668	0.959763
bias	LS	0.001981	-0.001931	-0.003995
	Huber M	0.001661	-0.002382	-0.034917
	Tukey M	0.001604	-0.002151	-0.036558
	S	0.000101	-0.003207	-0.090224
	MM	0.001685	-0.002332	-0.040237
nxvar	LS	1.025216	1.059594	2.141526
	Huber M	1.091179	1.110004	5.221268
	Tukey M	1.102804	1.123161	5.208683
	S	1.446793	1.460474	3.643037
	MM	1.080345	1.106108	2.024964
nxmse	LS	1.025412	1.059780	2.142324
	Huber M	1.091317	1.110288	5.282229
	Tukey M	1.102932	1.123392	5.275508
	S	1.446794	1.460988	4.050055
	MM	1.080487	1.106380	2.105914
REff	LS	100.0000	100.0000	100.0000
	Huber M	93.9610	95.4509	40.5572
	Tukey M	92.9715	94.3375	40.6089
	S	70.8748	72.5386	52.8962
	MM	94.9028	95.7881	101.7290

Genel bir değerlendirme yapılırsa MM tahmin edicilerinin iki değişkenli normal dağılımda bile çok yüksek bir etkinliğe sahip oldukları söylenebilir.

Tablo 4.1 – 4.3 düşünüldüğünde β_0 tahmin edicilerinin Tukey M dışında örneklem hacmi artıkça LS tahmin edicilerine göre bir düzelme kaydetmediği, durağan bir şekilde REFF değerlerini sürdürdüğü görülmektedir. β_1 için bakıldığında ise çok durağan olmasa da aynı seviyelerde bir seyir görülmektedir. Bunun dışında hem β_0 hem β_1 için tüm tahmin edicilerin yansız tahminler ürettiği görülmektedir.

Tablo 4.3 $X \sim N(0,1)$, $\varepsilon \sim N(0,1)$, $n = 100$ için simülasyon sonuçları

		β_0	β_1	σ^2
mean	LS	0.002002	0.998351	0.995692
	Huber M	0.001622	0.997816	0.978427
	Tukey M	0.001594	0.997910	0.977641
	S	0.000233	0.998239	0.935892
	MM	0.001645	0.998097	0.977486
bias	LS	0.002002	-0.001649	-0.004308
	Huber M	0.001622	-0.002184	-0.021573
	Tukey M	0.001594	-0.002090	-0.022359
	S	0.000233	-0.001761	-0.064108
	MM	0.001645	-0.001903	-0.022514
nxvar	LS	1.011624	1.052775	2.056568
	Huber M	1.078633	1.096791	5.161643
	Tukey M	1.085742	1.103081	5.145270
	S	1.437698	1.434008	3.725803
	MM	1.075380	1.096170	2.018996
nxmse	LS	1.012024	1.053047	2.058424
	Huber M	1.078896	1.097268	5.208183
	Tukey M	1.085997	1.103517	5.195261
	S	1.437704	1.434318	4.136787
	MM	1.075651	1.096532	2.069684
REff	LS	100.0000	100.0000	100.0000
	Huber M	93.8019	95.9699	39.5229
	Tukey M	93.1885	95.4264	39.6212
	S	70.3917	73.4179	49.7590
	MM	94.0848	96.0343	99.4560

σ^2 için tablolar incelendiğinde LS tahmin edicileri dışında tüm tahmin edicilerin yanlış tahmin değerleri ürettiği görülüyor. Genelde örneklem hacmindeki artışın yanlış miktarını azaltması beklenir fakat S tahmin edicisinde bu durum diğerlerine göre daha yavaş gerçekleşmiştir.

Genel olarak iki değişkenli normal dağılım için zaten beklenildiği üzere en iyi tahmin edicilerin LS tahmin edicileri olduğu görülmüştür. Sadece σ^2 tahmin edilirken düşük örneklem hacimlerinde MM tahmin edicilerinin yanlış tahmin edici üretmesinin avantajıyla öne geçtiği görülmektedir. Yüksek örneklem hacimlerinde ($n = 100$) ise MM tahmin edicisinin yanlış miktarının sıfıra yaklaştığı görülmektedir. Yanlış miktarının sıfıra yaklaşmasından dolayı MM tahmin edicisinin etkinliği doğal olarak LS tahmin edicisinin altına düşmüştür. Ancak yine de REFF değerinin %99.46 gibi olağanüstü bir değere sahip olduğu görülüyor.

Tüm parametreler ve tahmin ediciler değerlendirildiğinde MM tahmin edicilerinin en iyi ikinci tahmin ediciler olduğu görülmüştür. S ise en az etkinliği olan tahmin edicilere sahiptir. S tahmin edicileri genelde iterasyonlarda ilk

aşamalarda kullanılan ve etkinliği düşük tahmin ediciler olduklarından dolayı bu şartırtıcı bir sonuç değildir.

X 'in tasarımsal deęişken ve hata teriminin normal daęıldığı durumlara yönelik simülasyonlar Tablo 4.4 – 4.6 ile verilmiştir.

Tablo 4.4 $X = 1, \dots, n$, $\varepsilon \sim N(0,1)$, $n = 30$ için simülasyon sonuçları

		β_0	β_1	σ^2
mean	LS	0.001866	1.000044	1.002119
	Huber M	0.002559	1.000008	0.953909
	Tukey M	0.002966	0.999993	0.950310
	S	0.001276	1.000036	0.864782
	MM	0.002527	1.000002	0.940795
bias	LS	0.001866	0.000044	0.002119
	Huber M	0.002559	0.000008	-0.046091
	Tukey M	0.002966	-0.000007	-0.049690
	S	0.001276	0.000036	-0.135218
	MM	0.002527	0.000002	-0.059205
nxvar	LS	4.192345	0.013263	2.128409
	Huber M	4.437811	0.013938	5.050304
	Tukey M	4.552124	0.014305	5.012450
	S	5.926115	0.018535	3.569422
	MM	4.390596	0.013810	1.920482
nxmse	LS	4.192449	0.013263	2.128544
	Huber M	4.438008	0.013938	5.114037
	Tukey M	4.552388	0.014305	5.086523
	S	5.926164	0.018535	4.117941
	MM	4.390788	0.013810	2.025639
REff	LS	100.0000	100.0000	100.0000
	Huber M	94.4669	95.1553	41.6216
	Tukey M	92.0934	92.7159	41.8467
	S	70.7447	71.5551	51.6895
	MM	95.4829	96.0346	105.0801

Tablo 4.5 $X = 1, \dots, n$, $\varepsilon \sim N(0,1)$, $n = 50$ için simülasyon sonuçları

		β_0	β_1	σ^2
mean	LS	-0.002469	1.000197	1.002095
	Huber M	-0.003139	1.000216	0.975017
	Tukey M	-0.002960	1.000211	0.972576
	S	-0.004288	1.000225	0.904064
	MM	-0.002754	1.000202	0.965391
bias	LS	-0.002469	0.000197	0.002095
	Huber M	-0.003139	0.000216	-0.024983
	Tukey M	-0.002960	0.000211	-0.027424
	S	-0.004288	0.000225	-0.095936
	MM	-0.002754	0.000202	-0.034609
nxvar	LS	4.154817	0.004889	2.097306
	Huber M	4.397870	0.005134	5.272050
	Tukey M	4.451806	0.005207	5.262098
	S	5.868321	0.006838	3.823240
	MM	4.363263	0.005109	1.984214
nxmse	LS	4.155122	0.004891	2.097526
	Huber M	4.398363	0.005137	5.303259
	Tukey M	4.452244	0.005210	5.299701
	S	5.869240	0.006841	4.283425
	MM	4.363643	0.005111	2.044104
REff	LS	100.0000	100.0000	100.0000
	Huber M	94.4697	95.2102	39.5516
	Tukey M	93.3265	93.8735	39.5782
	S	70.7949	71.4890	48.9684
	MM	95.2214	95.6850	102.6134

Genel olarak bakıldığında Tablo 4.4 – 4.6'daki sonuçlar, Tablo 4.1 – 4.3'teki sonuçlar ile çok benzerdir. Sadece β_0 için üretilen değerlerde Tukey M tahmin edicisinin de diğer tahmin ediciler gibi örneklem hacmine göre durağan bir seyir izlediği görülmektedir.

Tablo 4.6 $X = 1, \dots, n$, $\varepsilon \sim N(0,1)$, $n = 100$ için simülasyon sonuçları

		β_0	β_1	σ^2
mean	LS	0.001009	1.000031	1.002265
	Huber M	0.001317	1.000022	0.990256
	Tukey M	0.001494	1.000018	0.989431
	S	0.001532	0.999999	0.930861
	MM	0.001563	1.000015	0.983856
bias	LS	0.001009	0.000031	0.002265
	Huber M	0.001317	0.000022	-0.009744
	Tukey M	0.001494	0.000018	-0.010569
	S	0.001532	-0.000001	-0.069139
	MM	0.001563	0.000015	-0.016144
nxvar	LS	4.049889	0.001178	2.049823
	Huber M	4.289180	0.001255	5.365072
	Tukey M	4.285941	0.001256	5.358429
	S	5.590845	0.001643	4.229279
	MM	4.252650	0.001242	2.024754
nxmse	LS	4.049991	0.001178	2.050336
	Huber M	4.289353	0.001255	5.374566
	Tukey M	4.286164	0.001256	5.369599
	S	5.591080	0.001643	4.707293
	MM	4.252894	0.001243	2.050817
REff	LS	100.0000	100.0000	100.0000
	Huber M	94.4196	93.8828	38.1489
	Tukey M	94.4899	93.7975	38.1842
	S	72.4367	71.7036	43.5566
	MM	95.2291	94.8184	99.9765

Tablo 4.7 $X = 1, \dots, n$, $\varepsilon \sim \pi N(0, 3^2) + (1 - \pi)N(0, 1)$, $\pi = 0.1$, $n = 30$ için simülasyon sonuçları

		β_0	β_1	σ^2
mean	LS	0.000414	1.000063	0.997363
	Huber M	0.001655	1.000007	0.631309
	Tukey M	0.002783	0.999954	0.627278
	S	0.001099	1.000005	0.575998
	MM	0.001578	1.000022	0.793914
bias	LS	0.000414	0.000063	-0.002637
	Huber M	0.001655	0.000007	-0.368691
	Tukey M	0.002783	-0.000046	-0.372722
	S	0.001099	0.000005	-0.424002
	MM	0.001578	0.000022	-0.206086
nxvar	LS	4.072526	0.012924	7.370363
	Huber M	3.090792	0.009763	2.394504
	Tukey M	3.085856	0.009774	2.384998
	S	3.690600	0.011594	1.713031
	MM	3.073310	0.009735	3.121696
nxmse	LS	4.072531	0.012924	7.370571
	Huber M	3.090874	0.009763	6.472494
	Tukey M	3.086088	0.009775	6.552649
	S	3.690636	0.011594	7.106358
	MM	3.073385	0.009735	4.395840
REff	LS	100.0000	100.0000	100.0000
	Huber M	131.7599	132.3835	113.8753
	Tukey M	131.9642	132.2246	112.4823
	S	110.3477	111.4734	103.7180
	MM	132.5096	132.7642	167.6715

Karma modele yönelik simülasyonlar Tablo 4.7 – 4.9 ile verilmiştir.

Tablo 4.7 – 4.9'a bakıldığında β_0 tahmin edilirken M ve MM tahmin edicileri arasında önemli bir fark bulunmamakla beraber genel kullanım için MM tahmin edicilerinin kullanımının daha uygun olduğu söylenebilir. β_1 için düşünüldüğünde ise M ve MM tahmin edicileri arasında yine önemli bir fark görülmemektedir ancak düşük örneklem hacminde ($n = 30$) Huber M'nin, orta ve yüksek örneklem hacimlerinde ($n = 50$ ve 100) ise Tukey M'in öne çıktığı görülmektedir. Hem β_0 hem β_1 için hassas yapılarından dolayı LS tahmin edicilerinin performansı tüm tahmin edicilerin altında kalmıştır. İterasyonla üretilen tahmin ediciler içinde de en az etkinliğe sahip olanın S tahmin edicisi olduğu görülmüştür.

Tablo 4.8 $X = 1, \dots, n$, $\varepsilon \sim \pi N(0, 3^2) + (1 - \pi)N(0, 1)$, $\pi = 0.1$, $n = 50$ için simülasyon sonuçları

		β_0	β_1	σ^2
mean	LS	-0.004006	1.000212	0.996651
	Huber M	-0.002840	1.000175	0.639458
	Tukey M	-0.002073	1.000149	0.637134
	S	-0.002793	1.000151	0.598342
	MM	-0.002971	1.000176	0.798454
bias	LS	-0.004006	0.000212	-0.003349
	Huber M	-0.002840	0.000175	-0.360542
	Tukey M	-0.002073	0.000149	-0.362866
	S	-0.002793	0.000151	-0.401658
	MM	-0.002971	0.000176	-0.201546
nxvar	LS	4.207796	0.004829	7.159571
	Huber M	3.057760	0.003568	2.415616
	Tukey M	3.001018	0.003542	2.377917
	S	3.596012	0.004226	1.776129
	MM	3.020584	0.003536	2.924275
nxmse	LS	4.208598	0.004831	7.160132
	Huber M	3.058163	0.003570	8.915136
	Tukey M	3.001233	0.003543	8.961505
	S	3.596402	0.004227	9.842572
	MM	3.021026	0.003538	4.955315
REff	LS	100.0000	100.0000	100.0000
	Huber M	137.6185	135.3313	80.3143
	Tukey M	140.2289	136.3689	79.8988
	S	117.0224	114.2875	72.7466
	MM	139.3102	136.5615	144.4940

σ^2 düşünülürken tek yansız tahmin edici LS tahmin edicisi olduğundan, asimptotik olarak en yüksek etkinliğe sahip tahmin edici de doğal olarak LS tahmin edicisidir. Dayanıklı tahmin edicilere bakıldığında eksi yana sahip oldukları görülmektedir ve bu eksi yan miktarı örneklem hacmi arttıkça azalmamaktadır. LS tahmin ediciyle karşılaştırıldıklarında dayanıklı tahmin ediciler düşük örneklem hacimlerinde ($n = 30$) yan miktarlarından dolayı avantaja sahipken, örneklem hacmi arttıkça etkinlik değerleri hızlı bir şekilde düşmektedir.

Tablo 4.9 $X = 1, \dots, n$, $\varepsilon \sim \pi N(0, 3^2) + (1 - \pi)N(0, 1)$, $\pi = 0.1$, $n = 100$ için simülasyon sonuçları

		β_0	β_1	σ^2
mean	LS	-0.000523	1.000038	0.996826
	Huber M	0.000370	1.000024	0.645820
	Tukey M	0.000623	1.000020	0.644534
	S	0.001074	0.999996	0.614155
	MM	-0.000011	1.000030	0.802166
bias	LS	-0.000523	0.000038	-0.003174
	Huber M	0.000370	0.000024	-0.354180
	Tukey M	0.000623	0.000020	-0.355466
	S	0.001074	-0.000004	-0.385845
	MM	-0.000011	0.000030	-0.197834
nxvar	LS	4.141320	0.001187	6.948147
	Huber M	2.957752	0.000867	2.382351
	Tukey M	2.862243	0.000848	2.370896
	S	3.404990	0.001005	1.923170
	MM	2.901395	0.000854	2.780392
nxmse	LS	4.141347	0.001187	6.949155
	Huber M	2.957765	0.000867	14.926714
	Tukey M	2.862281	0.000848	15.006510
	S	3.405105	0.001005	16.810773
	MM	2.901395	0.000854	6.694227
REff	LS	100.0000	100.0000	100.0000
	Huber M	140.0161	136.9629	46.5552
	Tukey M	144.6869	139.9995	46.3076
	S	121.6217	118.1639	41.3375
	MM	142.7364	139.0428	103.8082

Dayanıklı tahmin edicilerin çoğunun $n = 30$ örneklem hacminden sonra σ^2 tahmin edicileri LS tahmin edicisinden daha kötü sonuçlar vermiştir. LS σ^2 tahmin edicisi yansız olduğundan örneklem hacmi arttıkça diğerlerine göre daha avantajlı duruma geçmesine rağmen MM σ^2 tahmin edicisi $n = 100$ örneklem hacminde bile etkinliğini %103.81 değerinde tutturmayı başarmıştır.

Tablo 4.10 $X = 1, \dots, n$, $\varepsilon \sim r$ tane $N(0, 3^2) + (n - r)$ tane $N(0, 1)$, $n = 30$ için simülasyon sonuçları

		β_0	β_1	σ^2
mean	LS	0.009515	0.999648	0.978323
	Huber M	0.003800	0.999922	0.630582
	Tukey M	0.001713	1.000027	0.625324
	S	0.001459	1.000022	0.573566
	MM	0.002212	0.999994	0.778391
bias	LS	0.009515	-0.000352	-0.021677
	Huber M	0.003800	-0.000078	-0.369418
	Tukey M	0.001713	0.000027	-0.374676
	S	0.001459	0.000022	-0.426434
	MM	0.002212	-0.000006	-0.221609
nxvar	LS	8.690849	0.021681	5.237186
	Huber M	5.194764	0.013861	2.229288
	Tukey M	5.127244	0.013778	2.241700
	S	5.457278	0.015014	1.522470
	MM	5.317370	0.014092	1.924192
nxmse	LS	8.693565	0.021684	5.251283
	Huber M	5.195198	0.013861	6.323376
	Tukey M	5.127332	0.013778	6.453166
	S	5.457342	0.015014	6.977846
	MM	5.317517	0.014092	3.397511
REff	LS	100.0000	100.0000	100.0000
	Huber M	167.3385	156.4406	83.0456
	Tukey M	169.5534	157.3867	81.3753
	S	159.3004	144.4289	75.2565
	MM	163.4892	153.8754	154.5627

Tablo 4.11 $X = 1, \dots, n$, $\varepsilon \sim r$ tane $N(0, 3^2) + (n - r)$ tane $N(0, 1)$, $n = 50$ için simülasyon sonuçları

		β_0	β_1	σ^2
mean	LS	0.002857	1.000013	0.988317
	Huber M	-0.000002	1.000093	0.640347
	Tukey M	0.000255	1.000085	0.636678
	S	-0.004731	1.000213	0.601879
	MM	0.000535	1.000073	0.789466
bias	LS	0.002857	0.000013	-0.011683
	Huber M	-0.000002	0.000093	-0.359653
	Tukey M	0.000255	0.000085	-0.363322
	S	-0.004731	0.000213	-0.398121
	MM	0.000535	0.000073	-0.210534
nxvar	LS	8.520271	0.007887	5.511604
	Huber M	5.134799	0.005111	2.301757
	Tukey M	5.150906	0.005145	2.270587
	S	5.445058	0.005579	1.613242
	MM	5.257893	0.005201	1.950379
nxmse	LS	8.520679	0.007887	5.518429
	Huber M	5.134799	0.005111	8.769282
	Tukey M	5.150909	0.005145	8.870733
	S	5.446177	0.005581	9.538243
	MM	5.257907	0.005201	4.166615
REff	LS	100.0000	100.0000	100.0000
	Huber M	165.9399	154.2996	62.9291
	Tukey M	165.4209	153.2925	62.2094
	S	156.4525	141.3241	57.8558
	MM	162.0546	151.6371	132.4439

Dixon aykırı modele göre gerçekleştirilen simülasyon sonuçları Tablo 4.10 – 4.12 ile verilmiştir.

Dixon aykırı değer modelinde karma modele göre en büyük farklılık aykırı değerlerin LS β_0 ve β_1 tahmin edicilerindeki etkisinin daha yıkıcı olmasıdır. β_0 ve β_1 tahmin edilirken M tahmin edicileri MM tahmin edicilerine göre daha iyi sonuçlar vermiştir. M tahmin edicileri kendi aralarında karşılaştırıldığında çok büyük farklar olmamasına rağmen düşük ve yüksek örneklem hacimlerinde ($n = 30$ ve 100) Tukey M tahmin edicisi daha iyi sonuçlar gösterirken orta örneklem hacimlerinde ($n = 50$) Huber M tahmin edicisi daha iyi sonuçlar göstermiştir. S tahmin edicisi ise yine en az etkinliğe sahip tahmin edici olmuştur.

Tablo 4.12 $X = 1, \dots, n$, $\varepsilon \sim r$ tane $N(0, 3^2) + (n - r)$ tane $N(0, 1)$, $n = 100$ için simülasyon sonuçları

		β_0	β_1	σ^2
mean	LS	0.001340	1.000013	0.991007
	Huber M	0.002569	0.999994	0.646048
	Tukey M	0.003927	0.999972	0.644191
	S	0.000285	1.000015	0.620929
	MM	0.004091	0.999970	0.798507
bias	LS	0.001340	0.000013	-0.008993
	Huber M	0.002569	-0.000006	-0.353952
	Tukey M	0.003927	-0.000028	-0.355809
	S	0.000285	0.000015	-0.379071
	MM	0.004091	-0.000030	-0.201493
nxvar	LS	8.223095	0.001912	5.158126
	Huber M	4.865420	0.001203	2.329095
	Tukey M	4.840970	0.001188	2.309807
	S	5.163694	0.001317	1.740649
	MM	4.931098	0.001208	1.847073
nxmse	LS	8.223274	0.001912	5.166214
	Huber M	4.866080	0.001203	14.857304
	Tukey M	4.842513	0.001188	14.969820
	S	5.163702	0.001317	16.110107
	MM	4.932771	0.001208	5.907030
REff	LS	100.0000	100.0000	100.0000
	Huber M	168.9917	159.0280	34.7722
	Tukey M	169.8142	160.9123	34.5109
	S	159.2515	145.2175	32.0682
	MM	166.7070	158.2935	87.4587

σ^2 tahmin edilirken yine karma modeldeki gibi tek yansız tahmin edici LS tahmin edicisi oluştur. Karma modeline göre farklılığı ise tahmin edicilerin yanlarının etkinliklerine olan olumsuz etkisinin çok daha yıkıcı olmasıdır. $n = 30$ örneklem hacminde bile etkinlik değeri LS tahmin edicisinin etkinliğinin altına düşmeyen tek tahmin edici MM tahmin edicisidir. Ancak $n = 100$ örneklem hacminde MM tahmin edicisinin etkinlik değeri bile yan değeri düşmediği için LS tahmin edicisinin etkinlik değerinin altında kalmıştır.

5. GERÇEK HAYAT VERİSİ ÜZERİNE UYGULAMALAR

5.1 Uygulama 1

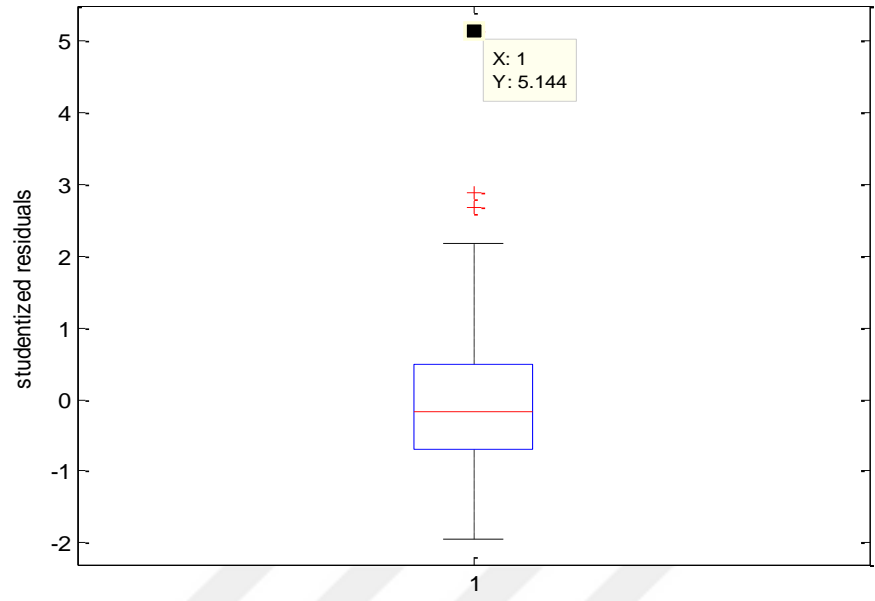
Bu bölümde R programının New York'taki günlük hava ölçümlerini içeren airquality veri seti kullanılmıştır. Bu veri seti 1 Mayıs 1973'ten 30 Eylül 1973'e kadar olan 111 veriyi içermektedir (Stuart, 2011). Bağımlı değişken ortalama ozon yoğunluğu, bağımsız değişkenler ise günlük en yüksek sıcaklık, ortalama rüzgar hızı ve güneş ışığı verileri olarak belirlenmiştir. LS, Tukey M, Huber M, S ve MM tahmin edicileriyle yapılan parametre tahmin değerleri Tablo 5.1'de gösterilmiştir.

Tablo 5.1 Uygulama 1 için tahmin değerleri

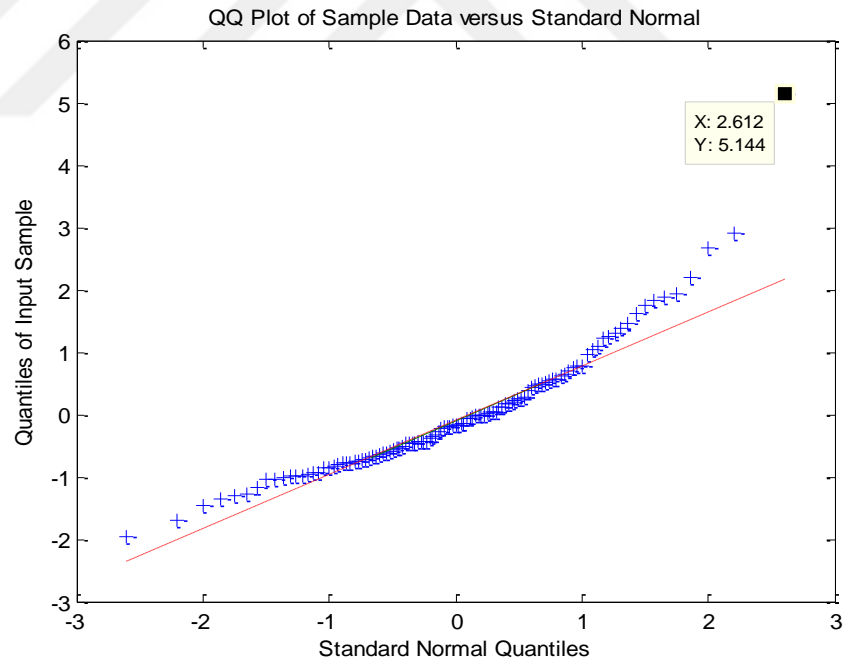
	ls	Tukey m	Huber m	s	mm
b0	-64.3421	-84.7042	-78.8615	-83.985	-84.107
b1	1.6521	1.7815	1.7499	1.802	1.7802
b2	-3.3336	-2.2977	-2.6565	-2.783	-2.3439
b3	0.0598	0.0455	0.0494	0.0487	0.0462
sigma^2	448.6242	324.661	314.227	278.592	361.789
mse	448.624	236.205	291.179	57.895	246.935
R^2	0.606	0.679	0.657	0.766	0.675

Tablo 5.1 incelendiğinde düşük etkinliğe sahip olarak bilinen S tahmin edicisinin, bilinenin aksine diğer tahmin edicilerle karşılaştırıldığında daha küçük MSE değerine ve daha yüksek R^2 'ye sahip olduğu görülmektedir. Tukey M, Huber M ve MM tahmin edicileri yakın sonuçlar vermelerine rağmen Huber M tahmin edicisi içlerinde en kötü sonucu veren tahmin edici olmuştur. LS tahmin edisi ise yüksek MSE değeri ve düşük R^2 değeriyle en zayıf performansı göstermiştir.

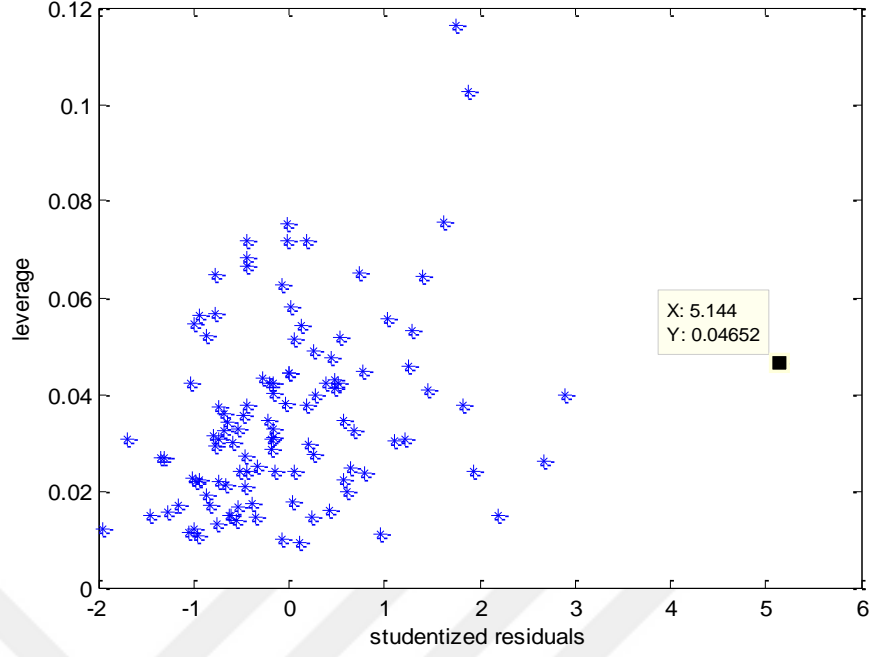
Uygulama 1 verisi için LS yöntemine dayanarak hazırlanan artık değer kutu grafiği, q-q grafiği, studentize artık değer, kaldıraç ve tutturulan Y-studentize artık değer grafiği sırasıyla Şekil 5.1 – 5.4'te verilmiştir.



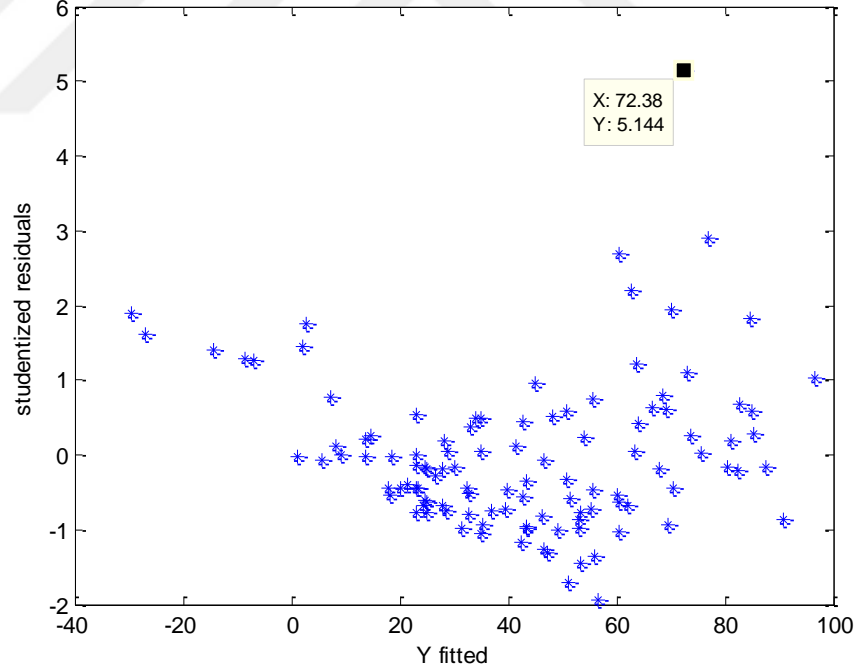
Şekil 5.1 Uygulama 1 için artık değerlerin kutu grafiği



Şekil 5.2 Uygulama 1 için q-q grafiği



Şekil 5.3 Uygulama 1 için studentize artık değer - kaldıraç grafiği



Şekil 5.4 Uygulama 1 için tutturulan Y - studentize artık değer grafiği

Şekil 5.4 incelendiğinde sistematik bir yapı görülmediği ve bu sebeple modelin doğru olduğu görülmektedir. Şekil 5.4, 5.1 ve 5.2 beraber incelendiğinde biri daha ekstrem olmak üzere üç aykırı değer (5.144, 2.8963 ve 2,6813) olduğu

söylenbilir. Bu aykırı değerlerin olası etkilerini gösteren Şekil 5.3'e bakıldığında ekstrem olan aykırı değer (5.144) kaldıraç etkisinin ortalamanın üstünde olduğu dikkat çekmektedir. 3 değerine yakın olan aykırı değer de kaldıraç değeri yüksektir. Bu değerlerin muhtemel olumsuz etkisi yüksek olacaktır. Genel olarak düşünüldüğünde veri setinde bu aykırı değerlerin bulunması S tahmin edicisinin iyi sonuçlar vermesinin sebebi olabilir.

5.2 Uygulama 2

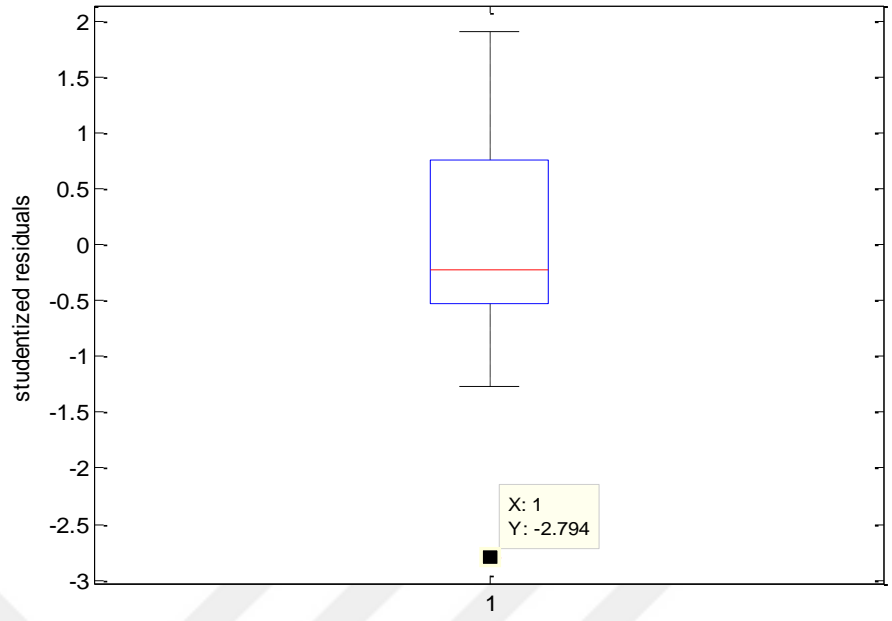
Bu bölümde Çetin ve Orsoy (2001)'de kullanılan Hacettepe Üniversitesi Hastanesi'ne başvuran 22 hastaya ait bilgileri içeren veri seti kullanılmıştır. Bağımlı değişken kemik mineral yoğunluğu, bağımsız değişkenler ise osteocalcin miktarı, paratiroid hormonu ve yaşıdır. LS, Tukey M, Huber M, S ve MM tahmin edicileriyle yapılan parametre tahmini Tablo 5.2'de gösterilmiştir.

Tablo 5.2 Uygulama 2 için tahmin değerleri

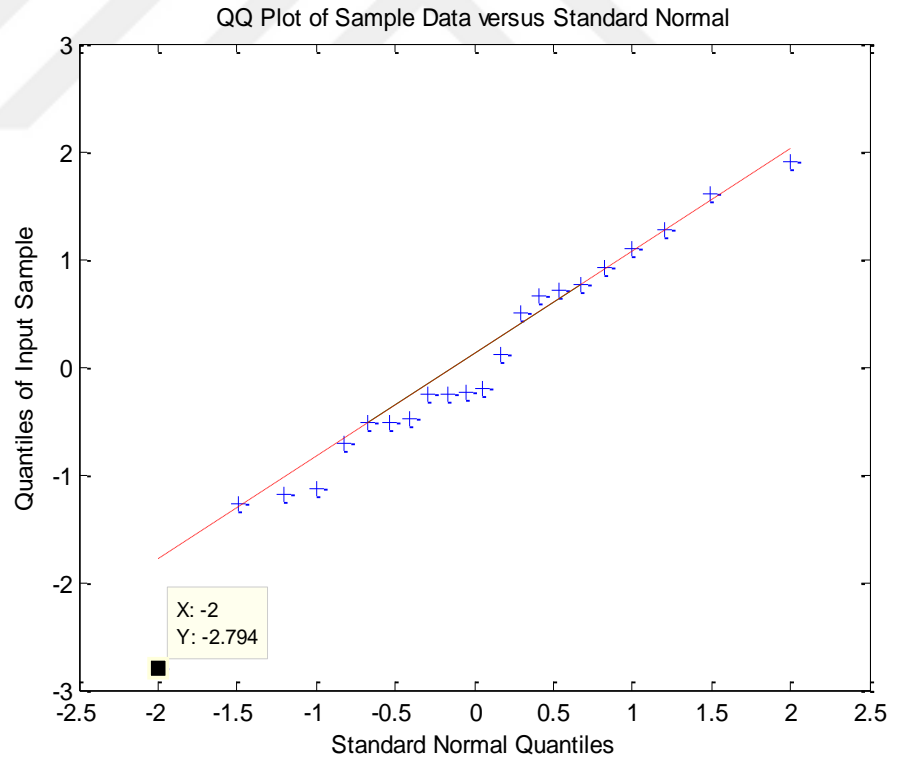
	ls	Tukey m	Huber m	s	mm
b0	1.0874	1.0856	1.0772	1.0432	1.0853
b1	0.0288	0.0340	0.0335	0.0395	0.0336
b2	-0.0022	-0.0024	-0.0024	-0.0025	-0.0024
b3	-0.0060	-0.0062	-0.0061	-0.0058	-0.0062
Sigma^2	0.0039	0.0030	0.0031	0.0027	0.0032
mse	0.0039	0.0029	0.0031	0.0006	0.0029
R^2	0.4029	0.5047	0.4842	0.7076	0.4970

Tablo 5.2 incelendiğinde düşük etkinliğe sahip olarak bilinen S tahmin edicisinin, bilinenin aksine diğer tahmin edicilerle karşılaştırıldığında daha küçük MSE değerine ve daha yüksek R^2 'ye sahip olduğu görülmektedir. Tukey M, Huber M ve MM tahmin edicileri incelendiğinde bu tahmin edicilerin birbirlerine oldukça yakın sonuçlar verdikleri görülmüştür. LS tahmin edicisi ise yüksek MSE değeri ve düşük R^2 değeriyle en zayıf performansı göstermiştir.

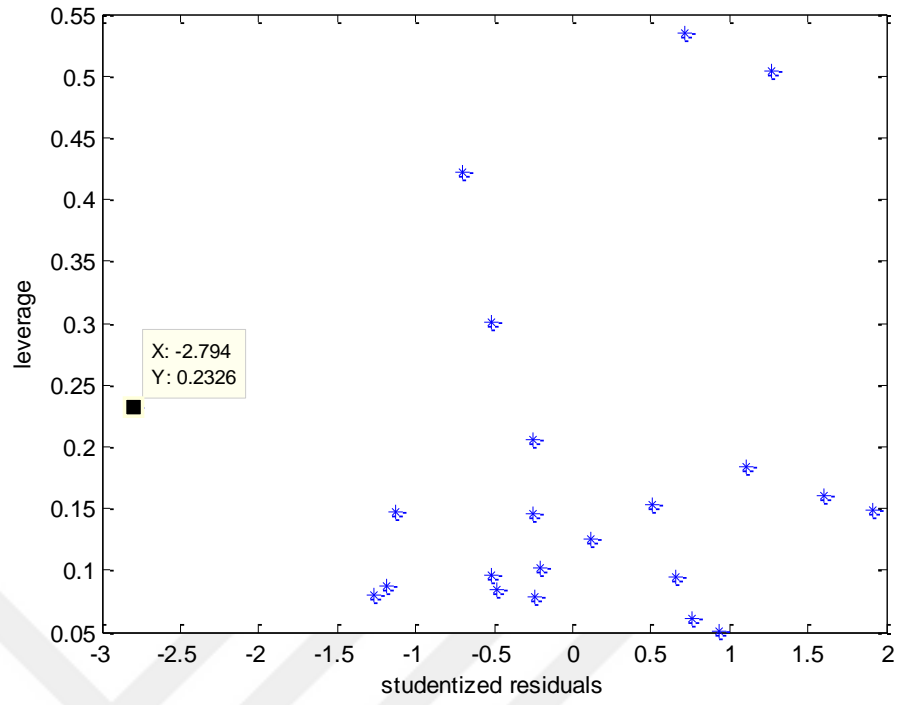
Uygulama 2 verisi için LS yöntemine dayanarak hazırlanan artık değer kutu grafiği, q-q grafiği, studentize artık değer, kaldıraç ve tutturulan Y-studentize artık değer grafiği sırasıyla Şekil 5.5 – 5.8'de verilmiştir.



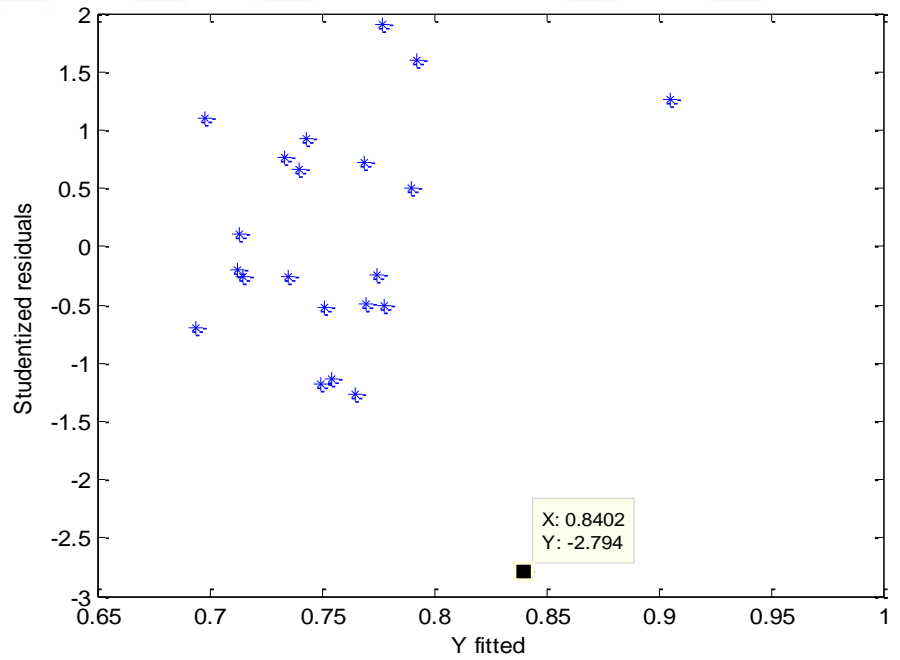
Şekil 5.5 Uygulama 2 için kutu grafiği



Şekil 5.6 Uygulama 2 için q-q grafiği



Şekil 5.7 Uygulama 2 için studentize artık değer – kaldıraç grafiği



Şekil 5.8 Uygulama 2 için tutturulan Y - studentize artık değer grafiği

Şekil 5.8 incelendiğinde sistematik bir yapı görülmediğinden modelin uygun olduğu söylenebilir. Şekil 5.5 ve 5.7'ye göre bir adet aykırı değer görülmektedir (-2.794). Şekil 5.7'ye baktığımızda aykırı değer olası olumsuz etkisinin de fazla olduğu ve parametre tahminini etkileyebileceği söylenebilir. Uygulama 1'de olduğu gibi aykırı değer varlığı S tahmin edicisinin iyi sonuçlar vermesinin sebebi olabilir.



6. SONUÇ VE TARTIŞMA

Parametre tahmininde yaygın olarak kullanılan yöntemlerden biri LS yöntemidir. Bu çalışmada LS tahmin edicisiyle beraber iterasyona dayalı M Huber, M Tukey, S ve MM tahmin edicileri Matlab'ta yazılan programlarla yürütülen simülasyonlar aracılığıyla karşılaştırılmıştır.

Literatürde bahsedildiği gibi MM tahmin edicileri yapılan simülasyonlar sonucunda etkinliği en yüksek tahmin ediciler olarak karşımıza çıkmaktadır. S tahmin edicileri ise dayanıklı tahmin ediciler arasında en düşük etkinliğe sahip tahmin ediciler olarak elde edilmiştir. Aykırı değer içeren durumlarda hassas yapılarından dolayı en çok etkilenen ise LS tahmin edicileri olmuştur.

Tahmin edicilerin σ^2 tahminlerine bakıldığında LS tahmin edicisinin yansız tahminler üreten tek tahmin edici olduğu görülmektedir. Diğer dayanıklı tahmin ediciler artan örneklem hacmine rağmen yan miktarlarını azaltamamışlardır.

Tahmin ediciler gerçek hayat verisi üzerinde karşılaştırıldıklarında şaşırtıcı bir şekilde S tahmin edicilerinin en iyi sonuçları verdiği gözlemlenmiştir. Gerçek hayat verilerinin içerdiği aykırı değerler bu duruma neden olarak düşünülmüştür.

KAYNAKLAR DİZİNİ

- Andersen, R.**, 2008, Modern Methods For Robust Regression, Thousand Oaks: SAGE Publications.
- Copt, S. and Heritier, S.**, 2006, Robust MM Estimation and Inference in Mixed Linear Models.
- Çetin, M. ve Orsoy, A.**, 2001, Doğrusal Regresyonda Sağlam Tahmin Ediciler ve Bir Uygulama , *Anadolu Üniversitesi Bilim ve Teknoloji Dergisi*, 2(2), 265-270ss.
- Çetin, M. and Toka, O.**, 2011, The Comparing of S Estimator and M Estimators in Linear Regression, *Gazi University Journal of Science*, 24, 747-752pp.
- Donoho, D.L. and Huber, P.J.**, 1983. The notion of breakdown point, *In A Festschrift for Erich L. Lehmann* (P. J. Bickel, K. Doksum and J. L. Hodges, Jr., eds.), Wadsworth, Belmont, CA, 157–184pp.
- Huber, P.J.** 1973, Robust regression: Asymptotics, Conjectures and Monte Carlo, *The Annal of Statistics*, 1(5), 799-821pp.
- Maronna, R. and Yohai, V.**, 2010, Correcting MM Estimates for Fat Data Sets, *Computational Statistics and Data Analysis*, 54, 3168-3173pp.
- Rousseeuw, P.J.**, 1984, Least Median of Squares Regression, *Journal of the American Statistical Association*, 79, 871-880pp.
- Rousseeuw, P.J. and Leroy, M.**, 1987, Robust Regression and Outlier Detection, New york: Wiley.
- Rousseeuw, P.J. and Yohai, V.**, 1984, Robust Regression by Means of S Estimators, *Robust and Nonlinear Time series Analysis*, edited by J. Franke, W. Hardle, and D. Martin, Lecture Notes in Statistics, 26, Springer Verlag (Berlin / New York), 256-272pp.
- Ryan, T.P.**, 1997, Modern Regression Methods', John Wiley Sons, New York.
- Salibian, M., Yohai, V.** , 2006, A Fast Algorithm for S Regression Estimates, *Journal of Computational and Graphical Statistics*, 15(2), 414-427pp.
- Stuart, C.**, 2011, Robust Regression, Department of Mathematical Sciences, Durham University.
- Susanti, Y., Pratiwi, H., Sulistijowati, S. and Liana, T.**, 2014, M estimation, S estimation, and MM estimation in robust regression, *International Journal of Pure and Applied Mathematics*, 91(3), 349-360pp.
- Türkay, H.**, Doğu Anadolu Bölgesi Araştırmaları, 2004, *Doğrusal Regresyon Analizinde M Tahminciler Ve Ekonometrik Bir Uygulama*, 106-115ss.

KAYNAKLAR DİZİNİ (devam)

Yohai, V.J., 1987, High Breakdown Point and High Efficiency Robust Estimates for Regression, *The Annals of Statistics*, 15(2), 642-656pp.



ÖZGEÇMİŞ

6 Nisan 1992 doğumlu olan Nalan MUTLU Bulgaristan'ın Kırcaali şehrinde doğdu. İlkokul eğitimini Bursa Adnan Menderes İlköğretim Okulu'nda, lise eğitimini ise Bursa Şükrü Şankaya Anadolu Lisesi'nde tamamladı. 2010 yılında Ege Üniversitesi Matematik Bölümünü kazandı. 2012 yılında ise Ege Üniversitesi İstatistik Bölümünde çift anadal programına başladı. 2015 yılında bu programları tamamlayıp istatistik bölümünde yüksek lisans eğitimine başladı. Yüksek lisans süreci devam ederken 2017 yılında aynı bölümde akademisyen olarak bulunan Doç. Dr. Sevcan DEMİR ATALAY ile birlikte inşaat mühendisliği ile ortak bir çalışma olan 'Forecasting Occupancy for Demand Driven HVAC Operations Using Time Series Analysis' başlıklı makalesi Journal of Asian Architecture and Building Engineering dergisinde yayımlanmıştır.

