DOĞRUSAL REGRESYON MODELİ İÇİN

M-TAHMİNCİLERİN İNCELENMESİ

Vural YILDIRIM

Yüksek Lisans Tezi

İstatistik Anabilim Dalı

Temmuz – 2012

JÜRİ VE ENSTİTÜ ONAYI

Vural YILDIRIM'ın **"Doğrusal Regresyon Modeli İçin M-Tahmincilerin İncelenmesi"** başlıklı **İstatistik** anabilim dalındaki, Yüksek Lisans Tezi 23.07.2012 tarihinde, aşağıdaki jüri tarafından Anadolu Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin ilgili maddeleri uyarınca değerlendirilerek kabul edilmiştir.

	Adı-Soyadı	**İmza**
Üye (Tez Danışmanı)	**: Doç. Dr. Yeliz MERT KANTAR**	**.....**
Üye	**: Prof. Dr. Embiya AĞAOĞLU**	**.....**
Üye	**: Doç. Dr. Ali DENİZ**	**.....**

Anadolu Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu'nun …… tarih ve …… sayılı kararıyla onaylanmıştır.

Enstitü Müdürü

ÖZET
Yüksek Lisans Tezi

DOĞRUSAL REGRESYON MODELİ İÇİN M-TAHMİNCİLERİN İNCELENMESİ

Vural YILDIRIM

Anadolu Üniversitesi
Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı

Danışman: Doç. Dr. Yeliz MERT KANTAR
2012, 87 sayfa

Robust (Sağlam) regresyon tahmincileri, hataların normal dağılıma uymadığı veya veri setinde aykırı değer bulunması durumunda regresyon modelini en güvenilir şekilde tahmin etmek amacı ile geliştirilmiştir.

Bu tez çalışmasının amacı, veri setinde aykırı değer olması durumunda en küçük kareler tahmincisine alternatif olarak geliştirilen robust regresyon tahmincilerinden M-tahmincilerin çeşitli açılardan incelenmesidir.

İlk olarak M-Tahmincilerin hesaplanmasında kullanılan yeniden ağırlıklandırılmış en küçük kareler algoritmasının başlangıç tahminlerinin seçimine olan duyarlılığı ele alınmış ve M-tahmincilerinin kırılma noktaları, grafikler yardımıyla incelenmiştir. Daha sonra hata teriminin dağılımının normal ve normalden farklı olduğu durumlar için M-tahmincilerinin etkinlik açısından performansı değerlendirilmiş ve son olarak başlangıç ölçek tahmincisinin etkinliğe katkısı araştırılmıştır. Ayrıca reel yaşamdan alınan iki örnek üzerinde M-tahminciler uygulanmış ve elde edilen sonuçlar tartışılmıştır.

Anahtar kelimeler: En küçük kareler tahmincisi**,** aykırı değer, robust tahminci, M-tahminci, yeniden ağırlıklandırılmış en küçük kareler algoritması, kırılma noktası, robust ölçek tahmincisi.

ABSTRACT
Master of Science Thesis

EXAMINATION OF M-ESTIMATOR FOR LINEER REGRESSION MODEL

Vural YILDIRIM

Anadolu University
Graduate School of Sciences
Statistics Program

Supervisor: Assoc. Prof. Dr. Yeliz MERT KANTAR
2012, 87 pp

Robust regression estimators have been developed to estimate regression model properly when errors have not normal distribution or there are outliers in data set.

The objective of this thesis is an examination of M-estimators, one of robust regression estimators developed as an alternative to least squares estimator.

Firstly, the sensitivity of iterative reweighted least squares algorithm to the choice of initial estimates is considered and the breakdown points of M-estimators are examined by means of plots. Next, when the distribution of error term is normal and different from normal, the performance of M-estimators is evaluated in terms of efficiency and finally contribution of initial scale estimator to efficiency is assessed. Also, M-estimators are applied on two real life examples and the obtained results are discussed.

Keywords: Least squares estimator, outlier, robust estimator, M-estimator, iterative reweighted least squares algorithm, breakdown point, robust scale estimator.

TEŞEKKÜR

Araştırmaların üzerine azimle gitmeyi öğreten, en sıkıştığım anlarda ufkuyla çalışmalarımı aydınlatan, güler yüzü ve hiç esirgemediği desteğiyle moral ve motivasyonumu her zaman en üst seviyede tutmamı sağlayan, eğitim gördüğüm süre boyunca bilgilerinden faydalandığım süreç içerisinde hoşgörüsünü ve sabrını hiç esirgemeyen, bilgi birikimimin artmasında yaptığı önemli yönlendirmelerle gelişme göstermemi sağlayan, beraber çalışmaktan her zaman zevk aldığım ve onur duyduğum danışmanım Doç. Dr. Yeliz MERT KANTAR'a teşekkür ederim.

Akademik çalışmalar yapma konusunda beni cesaretlendiren ve ilk çalışmalarımın ortaya çıkmasını sağlayan, zamanını ve hoşgörüsünü hiçbir zaman esirgemeyen, bilgi birikimiyle beni kritik alanlarda yönlendiren, çok değerli hocam Prof. Dr. Birdal ŞENOĞLU'na teşekkür ederim.

Yüksek lisans eğitimimin kesintiye uğramaması için gerekli desteği sağlayan, büyük yardımlarını gördüğüm, bilgi ve deneyimlerinden yararlandığım saygıdeğer daire başkanım Harita Mühendisi Hüseyin OKCU'ya teşekkür ederim.

İnsanlığını her zaman hissettiğim, desteğini üzerimden hiçbir zaman eksik etmeyen çok değerli şube müdürüm İstatistikçi Arzu ÇELİK'e teşekkür ederim.

Kendilerini tanımaktan ve aynı ortamda birlikte çalışmaktan mutluluk duyduğum, zor anlarımda desteklerini esirgemeyen, moral ve motivasyonumu kaybetmememi sağlayan çok sevgili iş arkadaşlarıma teşekkür ederim.

Vural YILDIRIM
Temmuz 2012

İÇİNDEKİLER

Sayfa

ŞEKİLLER DİZİNİ

Sayfa

ÇİZELGELER DİZİNİ

Sayfa

SİMGELER VE KISALTMALAR DİZİNİ

OLS : En Küçük Kareler

LTS : En Küçük Budanmış Kareler

LMS : En Küçük Medyan Kareler

LAD : En Küçük Mutlak Sapmalar

ARM : Andrew'in Medyan Regresyon

BLUE : En İyi Lineer Yansız Tahminci

UMVUE : Düzgün En Küçük Varyanslı Yansız Tahmin Edici

IC : Etki eğrisi (Influence Curve)

IF : Etki Fonksiyonu

SC : Duyarlılık eğrisi (Sensitive Curve)

MLE : En Çok Olabilirlik Tahmincisi

IRWLS : Yeniden Ağırlıklandırılmış En Küçük Kareler

QSR : Quadratic Squares Root

Bias : Yanlılık

MSE : Hata Kareler Ortalaması

RMSE : Hata Kareler Ortalaması Karekökü

MAD : Medyan Mutlak Sapma

Min : Minimum

Max : Maximum

cov : Kovaryans

med : Medyan

1. GİRİŞ

Regresyon analizi değişkenler arasındaki ilişkiyi araştıran ve modelleyen istatistiksel bir tekniktir. Bu tekniğin başarısı uygun regresyon modelinin ve uygun tahmincinin seçimine bağlıdır. En yaygın kullanılan regresyon modeli doğrusal regresyon modelidir. Doğrusal regresyon analizi uygulanırken, parametreleri tahmin etmek için en çok kullanılan teknik En Küçük Kareler (OLS) tekniği olup belirli varsayımlar altında en iyi sonucu vermektedir. Fakat gerçek yaşam uygulamalarına bakıldığında, bu varsayımların sağlanması her zaman gerçekleşmemektedir. Örneğin veri setinde aykırı değer olarak adlandırılan ve diğer verilere göre oldukça büyük veya oldukça küçük değerlere sahip olan gözlemler bulunabilir. Bu durumda, hataların normal dağılması varsayımı gerçekleşmemiş olur. Üstelik böyle bir veri setine OLS tekniği uygulanırsa sonuçlar tutarlı olmayacaktır. Bu duruma özel olarak çeşitli teknikler geliştirilmiştir. Bunlardan birisi de aykırı değerlerin veri setinden atılarak OLS tekniğinin uygulanmasıdır. Fakat literatürde kesin kabul edilmiş bir aykırı değer teşhis tekniği mevcut değildir, sadece uygulamalara bakılarak tamamen geleneksel olarak bazı sınırlar önerilmiştir. Bu nedenle de bu yöntemin kullanılması birçok araştırmacı tarafından önerilmemektedir. Veri setinde aykırı değer varlığında önerilen bir diğer teknik de Robust (Sağlam) tekniklerdir. Bu teknikler 1900'lerin ikinci yarısından itibaren çok büyük gelişmeler göstermiş ve giderek popülerlik kazanmıştır. Öyle ki Robust teknikler kendi içinde bile özellikleri bakımından çok farklı gruplara ayrılmıştır. Sıra istatistiklerinin lineer kombinasyonları olarak ifade edilen L-tahminciler, sıra testlerinden türetilen R-tahminciler, en küçük medyan kareler (LMS) tahmincileri, en küçük budanmış kareler (LTS) (Rousseeuw (1984)) tahmincileri, maksimum likelihood tipi olarak bilinen M-tahminciler (Huber, 1964; Andrews, 1974) literatürde bilinen en yaygın robust tahmincilere örnek olarak verilebilir. Bu tezde ise kullanımı ve tanımlanması kolay her tip örneklem büyüklüklerinde kolayca hesaplanabilen M-tahminciler doğrusal regresyon analizi kapsamında incelenecektir. (Wu, 2005)

İkinci bölümde lineer regresyon modeli için OLS ve MLE tahmincileri anlatılacaktır. Ayrıca belli varsayımlar altında OLS'nin üstünlükleri belirtilecek,

OLS'nin çok iyi bir tahminci olmasının sağladığı avantajıyla birlikte çok katı olan ve sağlanması gerekli olan koşulların varlığına da bağlı olmasından dolayı getirdiği dezavantajı belirtilecektir.

Üçüncü bölümde robust (sağlam) regresyon kavramından bahsedilip bazı tanımlamalar yapılacak ve bu tezin ana konusu olan M-tahminciler anlatılacaktır. Literatürde birçok M-tahmincisi yer almaktadır. Bu bölümde geniş bir literatür taramasıyla elde edilen M-tahminciler tanıtılacaktır. Bilinmektedir ki M-tahmincilerin hesaplanabilmeleri için çeşitli nümerik tekniklere ihtiyaç duyulmaktadır. Bu teknikler arasında en çok kabul göreni iteratif olarak yeniden ağırlıklandırılmış en küçük kareler (Iteratively reweighted least squares) yani IRWLS tekniğidir. IRWLS tekniği, M-tahminciler için tanıtılacak, ayrıca algoritmanın bazı avantaj ve dezavantajları tartışılacaktır.

Dördüncü bölümde M-tahmincilerin performanslarına ilişkin yapılan analizlerin sonuçlarına yer verilecektir. İlk olarak M-tahmincilerin IRWLS algoritması kullanılmasından dolayı hesaplanma sürecinde ihtiyaç duyduğu başlangıç tahmincilerinin katsayı sonuçlarına etkileri grafikler ile gösterilecektir. Daha sonra başlangıç tahmincilerinin iterasyon sayılarına etkileri de incelenerek sonuçları karşılaştırılacaktır. Robust regresyon tekniklerinde sık kullanılan bir kavram olan kırılma noktası M-tahminciler için grafikler yardımıyla iki farklı durum için gösterilecektir. Bu çalışma kapsamında M-tahmincilerin farklı hata dağılımları için performansları da araştırılarak sonuçları RMSE kriterine göre karşılaştırılacaktır. Ayrıca M-tahmincilerin aykırı değer teşhis metodu olarak performansı incelenecektir. Son olarak ölçeğin robust tahmincilerinin değerlendirilmesi yapılacaktır.

Beşinci bölümde M-tahmincilerle ilgili iki adet uygulama yapılacak ve OLS'ye göre sonuçlardaki değişimler incelenecektir.

Sonuç bölümünde ise tezde elde edilen sonuçlar özetlenecektir.

Bu tez çalışmasında M- tahminciler ve diğer bahsedilen tahminciler için sonuçlar Matlab R2012a programında yazılan özel kodlar ve LIBRA kütüphanesi yardımıyla elde edilmiştir.

2. DOĞRUSAL REGRESYON MODELİ

İstatistiğin en önemli kullanım alanlarından birisi de "Modelleme" dir. Araştırmacılar aralarında ilişki olan en az iki değişkeni modelleme ihtiyacı hissederler, yani bir değişkeni ikinci bir değişkenle (veya daha fazla değişkenle) açıklamak isterler. Bu ilişki analizi regresyon modeliyle yapılabilir. En yaygın regresyon modeli doğrusal (lineer) regresyon modelidir. İki yada daha fazla bağımsız değişkenle bağımlı bir değişken arasındaki lineer ilişkiyi modellemede kullanılır.

Doğrusal Regresyon modeli;

$$y_{ij} = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_k x_{ki} + \varepsilon_i \;\; i = 1, \dots, n \;, j = 1, \dots, k \qquad (2.1)$$

Matris ve vektörlerle gösterilirse;

$$Y = X\beta + \varepsilon \quad \text{şeklindedir.} \qquad (2.2)$$

Burada, Y: Bağımlı değişken (açıklanan değişken veya yanıt değişkeni), nx1 boyutlu vektör,

X: Bağımsız değişken (açıklayıcı değişken), nxp boyutlu matris,

$(p = k + 1)$

β: Regresyon katsayısı, px1 boyutlu vektör ($p = 2$ olduğunda regresyon modeli basit doğrusal regresyon modeli adını alır),

ε: Rassal hata terimi, nx1 boyutlu vektördür.

2.1. Regresyon Analizinin Amacı

Öne çıkan üç amacı vardır, bunlar;

I. Bağımlı ve bağımsız değişkenler arasındaki ilişkiyi kurmak
II. Bağımsız değişkenin alacağı değeri bağımlı değişkenler yardımıyla tahmin etmek
III. Bağımsız değişkenlerden hangileri bağımlı değişkeni açıklamada daha önemli (daha çok açıklıyor) olduğunu ortaya koymak

2.2. Regresyon Analizinin Adımları

I. Problem belirleme
II. İlgili potansiyel değişkenleri seçme
III. Veri toplama
IV. Model belirleme
V. Doğru tahmini için uygun teknik belirleme
VI. Modeli bu teknik ile uygulama
VII. Modelin uygunluğunu onaylama
VIII. Problemin çözümü için modeli/modelleri uygulama

2.3. Hata Teriminin Önemi

Hata teriminin modelde yer almasının sebepleri (Gujarati, 2001):

I. *Kuramın Belirsizliği*: Bağımlı değişkeni açıklayan bağımsız değişkenlerin tam olarak bilinememesi.
II. *Veri Bulunamaması*: Bağımsız değişkenin verilerinin erişilememesi veya elde edilememesi durumunda o bağımsız değişken modele alınmayacaktır.
III. *Öze – Çevreye İlişkin Değişkenler*: modele katkısı küçük olan bağımsız değişkenler çeşitli sebeplerden ötürü (maliyet, hesaplama zorluğu, zaman vb.) modele alınmaz.
IV. *İnsan Davranışlarında İçerilmiş Rassallık*: Bağımsız değişkenlerin tamamı modele girse bile insan davranışlarındaki rassallık hiçbir zaman belirlenemez.
V. *Güçsüz Yaklaşık Değişkenler*: Gerçek değerlerinin elde edilmesi mümkün olmayan bağımsız değişkenler için yaklaşık değerler kullanılmaktadır. Bu durum aynı zamanda ölçme hatası olarak da bilinmektedir.
VI. *Yanlış Fonksiyon Kalıbı*: Modelleme hatalarından kaynaklanan hatalar olabilir (nonlineer modele lineer model uygulanması gibi). Çok değişkenli regresyonda model belirlemenin zorluğundan dolayı bu tür durumlarla karşılaşılmaktadır.

2.4. En Küçük Kareler Tahmincisi

İlk olarak Legendre (1805) ve Gauss (1809) tarafından yayınlanmıştır. Doğrusal regresyon modelinin parametrelerini bulmada kullanılan en yaygın teknik En Küçük Kareler (Ordinary least squares-OLS) tahmincidir. Hata terimi üzerinde aşağıda verilen belli varsayımların sağlanması halinde OLS tahminci olarak en iyi lineer yansız tahminci yani BLUE özelliklere sahiptir.

i. $E(\varepsilon) = 0$ (2.3)

ii. $E(\varepsilon_i \varepsilon_j) = \sigma^2$ (2.4)

iii. $Cov(\varepsilon_i \varepsilon_j) = 0$ (2.5)

Ayrıca hata terimi normal dağılıma sahip $\varepsilon_i \sim N(0, \sigma^2 I)$ ise OLS tahmincisi düzgün en küçük varyanslı yansız tahmin edici UMVUE özelliğine sahiptir.

OLS'nin varsayımları aşağıdaki gibidir,

I. Hata terimleri rassaldır

II. Hata terimlerinin ortalaması sıfırdır

III. Hata terimleri sabit varyansa sahiptir

IV. Hata terimleri sıfır ortalamalı sabit varyanslı *Normal Dağılım*'a sahiptir

V. Hata terimleri arasında bir ilişki (otokorelasyon) yoktur

$$cov(\varepsilon_i, \varepsilon_j) = E\{[\varepsilon_i - E(\varepsilon_i)][\varepsilon_j - E(\varepsilon_j)]\} \quad (2.6)$$

$$cov(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i, \varepsilon_j) = 0\,, \quad i \neq j \quad (2.7)$$

VI. Bağımsız değişken ile hata terimi arasında ilişki yoktur

$$cov(\varepsilon_i, X_i) = E\{[(\varepsilon_i - E(\varepsilon_i)][(X_i - E(X_i)]\} \quad (2.8)$$

$$= E[\varepsilon(X_i - E(X_i)] \quad (2.9)$$

$$= E(\varepsilon X_i) - E(X_i)E(\varepsilon_i) \quad (2.10)$$

$$= E(\varepsilon_i, X_i) = 0 \quad (2.11)$$

VII. Bağımsız değişkenler arasında ilişki (çoklu doğrusal bağlantı) yoktur

OLS tekniğinin temel mantığı: regresyon katsayılarının tahmini ($\hat{\beta}$) öyle bir değer olmalıdır ki artıkların kareler toplamı minimum olsun. Yani öyle bir doğru elde edilmeli ki tüm veri noktalarının bu doğruya olan uzaklıkları kareleri toplamı en küçük olsun.

Matris formu (2.2) için şu şekilde hesaplama yapılır;

$$S(\beta) = \sum_{i=1}^{n} \varepsilon_i^2 = \varepsilon'\varepsilon = (Y - X\beta)'(Y - X\beta) \quad (2.12)$$

$$= Y'Y - \beta'X'Y - Y'X\beta + \beta'X'X\beta \quad (2.13)$$

$$= Y'Y - 2\beta'X'Y + \beta'X'X\beta \quad (2.14)$$

Burada $\beta'X'Y$ ifadesi 1x1 boyutlu olduğundan transpozunun değeri kendi değerine eşittir,

$$(\beta'X'Y)' = Y'X\beta \quad (2.15)$$

Regresyon katsayılarına göre türev alıp sıfıra eşitlersek

$$\left.\frac{\partial S}{\partial \beta}\right|_{\hat{\beta}} = -2X'Y + 2X'X\hat{\beta} = 0 \quad (2.16)$$

$$X'X\hat{\beta} = X'Y \quad \text{olur.} \quad (2.17)$$

$(X'X)^{-1} \neq 0$ olduğunda, bu denklemin çözümü OLS tahminlerini verir;

$$\hat{\beta} = (X'X)^{-1}X'Y \quad (2.18)$$

Tahmin edilen değerler şu şekilde bulunur;

$$\hat{y} = X\hat{\beta} = X(X'X)^{-1}X'Y \quad (2.19)$$

2.19 denklemi Hat matris yardımıyla da açıklanabilir, $p \times p$ boyutlu hat matris aşağıdaki şekildedir,

$$H = X(X'X)^{-1}X' \quad (2.20)$$

2.19 denklemi Hat matris ile ifade edilirse şu şekle dönüşür;

$$\hat{y} = HY \quad (2.21)$$

Artıklar ise şu şekilde hesaplanır;

$$\varepsilon = Y - \hat{Y} = Y - X\hat{\beta} = Y - HY = (I - H)Y \qquad (2.22)$$

2.5. En Çok Olabilirlik Tahmincisi

Hata terimlerinin dağılımı bilinirse, parametreleri tahmin etmek için alternatif bir yöntem olan en çok olabilirlik tekniği de kullanılabilir. Örneğin (Y_i, X_i) $i = 1,2, \dots, n$ şeklinde veri seti olsun, regresyon modelinde hataların da normal dağıldığı $(\varepsilon_i \sim N(0, \sigma^2))$ varsayılsın. Bu durumda bağımlı değişkenin terimleri normal ve bağımsız olarak $Y_i \sim N(\beta_0 + \beta_1 X_i, \sigma^2)$ dağılır. ε_i normal yoğunluk fonksiyonu;

$$f(\varepsilon_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}\varepsilon_i^2} \qquad (2.23)$$

$\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$ olmak üzere $\varepsilon = Y - X\beta$ şeklindedir, burada $\beta_0, \beta_1 \text{ve } \sigma^2$ bilinmeyen sabitlerdir. Bu durumda olabilirlik fonksiyonu

$$L(Y_i, X_i, \beta_0, \beta_1, \sigma^2) = \prod_{i=1}^{n} (2\pi\sigma^2)^{-1/2} \exp\left[-\frac{1}{2\sigma^2}(Y_i - \beta_0 - \beta_1 X_i)^2\right] \qquad (2.24)$$

$$= (2\pi\sigma^2)^{-n/2} \exp\left[-\frac{1}{2\sigma^2}(Y - X\beta)'(Y - X\beta)\right] \qquad (2.25)$$

Yukarıdaki denklemi maksimum yapan parametre değerleri $\tilde{\beta}_0, \tilde{\beta}_1 ve\ \tilde{\sigma}^2$ şeklinde simgelenir.

$$lnL(Y_i, X_i, \beta_0, \beta_1, \sigma^2) = -\left(\frac{n}{2}\right) ln2\pi - \left(\frac{n}{2}\right) ln\sigma^2 - \left(\frac{1}{2\sigma^2}\right)(Y - X\beta)'(Y - X\beta) \qquad (2.26)$$

veya

$$lnL(Y_i, X_i, \beta_0, \beta_1, \sigma^2) = -\left(\frac{n}{2}\right) ln2\pi - \left(\frac{n}{2}\right) ln\sigma^2 - \left(\frac{1}{2\sigma^2}\right)\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i)^2 \qquad (2.27)$$

Bu fonksiyonu maksimize etmek için $(Y - X\beta)'(Y - X\beta)$ ifadesinin minimize edilmesi gerekmektedir. Kısmi türevler alınırsa

$$\frac{\partial lnL}{\partial \beta_0}\Big|_{\tilde{\beta}_0,\tilde{\beta}_1,\tilde{\sigma}^2} = \frac{1}{\tilde{\sigma}^2}\sum_{i=1}^{n}\left(Y_i - \tilde{\beta}_0 - \tilde{\beta}_1 X_i\right) = 0 \qquad (2.28)$$

$$\frac{\partial lnL}{\partial \beta_1}\Big|_{\tilde{\beta}_0,\tilde{\beta}_1,\tilde{\sigma}^2} = \frac{1}{\tilde{\sigma}^2}\sum_{i=1}^{n}\left(Y_i - \tilde{\beta}_0 - \tilde{\beta}_1 X_i\right)X_i = 0 \qquad (2.29)$$

$$\frac{\partial lnL}{\partial\, \tilde{\sigma}^2}\Big|_{\tilde{\beta}_0,\tilde{\beta}_1,\tilde{\sigma}^2} = -\frac{n}{2\tilde{\sigma}^2} + \frac{1}{2\tilde{\sigma}^4}\sum_{i=1}^{n}\left(Y_i - \tilde{\beta}_0 - \tilde{\beta}_1 X_i\right)^2 = 0 \qquad (2.30)$$

Bu denklemlerin çözümü en çok olabilirlik tahmincilerini verir;

$$\tilde{\beta}_0 = \bar{Y} - \tilde{\beta}_1 \bar{X} \qquad (2.31)$$

$$\tilde{\beta}_1 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad (2.32)$$

$$\tilde{\sigma}^2 = \frac{\sum_{i=1}^{n}\left(y_i - \tilde{\beta}_0 - \tilde{\beta}_1 x_i\right)^2}{n} \quad \text{veya} \qquad (2.33)$$

$$\tilde{\sigma}^2 = \frac{(Y - X\beta)'(Y - X\beta)}{n} \qquad (2.34)$$

Burada dikkat edildiğinde en çok olabilirlik tahmincileri OLS tahmincileri ile aynıdır. Tabi bu durumun hataların normal dağılıma sahip olduğunda geçerli olduğu unutulmamalıdır. Ayrıca $\tilde{\sigma}^2$ σ^2 nin yanlı bir tahmincisidir ($\tilde{\sigma}^2 = [(n-1)/n]\hat{\sigma}^2$) n büyüdükçe yanlılık küçülür.

En çok olabilirlik tahmincileri OLS tahmincilerine göre daha iyi istatistiksel özellikler sunmaktadır.

En çok olabilirlik tahmincileri yansızdır (bu durum $\tilde{\sigma}^2$ içinde geçerlidir, çünkü asimptotik olarak yansızdır veya n büyüdükçe yansız olmaktadır) ve diğer yansız tahmin edicilerle karşılaştırıldığında belli koşullar altında minimum varyansa sahiptir.

Tutarlı tahmin edicilerdir (n büyüdükçe gerçek parametre değerinden sapmalar çok küçük miktarlarda olmaktadırlar).

Yeterli istatistiktir (anakütle için yeterli bilgiye sahiptir).

Diğer bir yandan da en çok olabilirlik tahmincileri OLS göre hata dağılımı hakkında daha katı varsayımlar gerektirmektedir.(Gujarati, 2001; Montgomery ve diğerleri, 2001)

3. ROBUST REGRESYON

Robust (Sağlam) regresyon tahmincileri hataların normal dağılıma uymadığı veya veri setinde aykırı değer bulunması durumunda veri setini en iyi şekilde modellemek için kullanılan bir tekniktir. Robust tahminciler, robust regresyon veya robust regresyon teknikleri olarak da ifade edilmektedirler. Robust regresyon tahmincileri aykırı değerlere karşı duyarlı olmayan tahmincilerdir. Robust regresyon tahmincileri aykırı değerlere ya düşük ağırlık vererek ya da tamamen red ederek güvenilir sonuçlar elde edilmesini sağlar (Winskowski ve diğerleri, 2000). Ayrıca, eğer veri setinde aykırı değer yoksa sonuçları en küçük kareler tahmincisi (OLS) ile benzer çıkması beklenmektedir. (Wang ve Wang, 2007).

$y = X\beta + \varepsilon$ lineer modelinde $E(\varepsilon) = 0$ ve $V(\varepsilon) = \sigma^2 \mathrm{I}$ varsayımları altında OLS tahmincisi $\hat{\beta}$, diğer tahminciler arasında en iyi tahminci olacaktır. Çünkü OLS tekniği "en iyi doğrusal yansız tahminci" (Best Linear Unbiased Estimator – BLUE)'dir. Yani tüm doğrusal yansız $\widehat{\beta_j}$ tahmincileri içerisinde OLS en küçük varyansa sahip olandır. Bununla beraber hataların normal dağılıma sahip olduğu varsayımı altında OLS tahmincisi tüm yansız $\hat{\beta}_j$ tahmincileri içerisinde en küçük varyansa sahip olandır. Bu nedenle OLS tahmincileri "düzgün en küçük varyanslı yansız tahmin edici" (Uniformly Minimum Variance Unbiased Estimator – UMVUE)'dir. Bu özellik katı bir şekilde normallik varsayımına dayanmaktadır. Fakat uygulamada birçok durumda hataların dağılımının normal dağılıma uymadığı görülür. Bunun sebebi hata teriminin dağılımının doğası gereği normal dağılım olmaması ya da aykırı değerler olabilir. (Mcdonald ve diğerleri, 2003) Buna örnek olarak pratikte gözlemlerin dağılımının normal dağılıma göre daha uzun kuyruklu bir dağılıma sahip olması verilebilir. Bu tür durumlarda uzun kuyruklu dağılımlar aykırı değer üretirler ve bu aykırı değerler de OLS tahminini etkileyerek regresyon doğrusunu kendi yönlerine çekerek hata değerlerini büyütürler. Bu durum OLS tahmincisinin her değere karşı oldukça duyarlı olmasından kaynaklanmaktadır. Burada da görüldüğü gibi OLS'nin hata dağılımına olan duyarlılığı OLS tahmincisini en iyi tahminci yaptığı gibi en kötü tahminci de yapabilmektedir.

3.1. Robust Regresyon Neden Gereklidir?

Aykırı değerler çeşitli tekniklere göre tespit edilebilirler. Bu aykırı değerler ölçüm hatalarından veya yanlış kayıt tutmadan kaynaklanabilirler. Bu durumda yeniden ölçüm yapılabilir ise düzeltilebilirler. Bu tür aykırı değerlere *"kötü"* aykırı değer denilmektedir ve bu aykırı değerler gerçek değerler olmadıkları için veri setinden atılabilirler. Fakat bazı aykırı değerler ölçümün gerçek değerini gösterdikleri için veri setinden atılamazlar. Ayrıca aykırı değerler belirli durumlarda kolayca tespit edilebilirler ama değişken sayısı çoğaldıkça veya daha karmaşık olduğu durumlarda aykırı değerleri tespit etmek ve bunları veri setinden atmak oldukça güçtür (Montgomery ve diğerleri, 2001).

Veri setinden aykırı değerler atıldığında klasik tekniklerin performansı robust teknikler kadar iyi çıkabilir. Analizlerde aykırı değerler oldukça büyük öneme sahip olmalarına karşın evrensel olarak kabul görmüş ve iyi tanımlanmış bir aykırı değer kavramı bulunmamaktadır. Ayrıca aykırı değerleri teşhis etmek her zaman kolay olamayabilir (Montgomery ve diğerleri, 2001). Tüm bu sebeplerden dolayı bu tür veriler için robust regresyon uygulamak en iyi çözümü almak için gereklidir. Literatürde, pek çok robust tahminci mevcuttur. Maksimum likelihood tipi olarak bilinen M-tahminciler (Huber, 1964; Andrews, 1974), sıra istatistiklerinin lineer kombinasyonları olarak ifade edilen L-tahminciler, sıra testlerinden türetilen R-tahminciler, En küçük medyan kareler (LMS) ve En Küçük Budanmış Kareler (LTS) (Rousseeuw, 1984) tahminciler literatürde bilinen en yaygın robust tahmincilerdir.

Bu tezde ise kullanımı ve tanımlanması kolay her tip örneklem büyüklüklerinde kolayca hesaplanabilen M-tahminciler ele alınacaktır. (Wu, 2005)

3.2. İyi Bir Robust Tahmincisinde Bulunması İstenilen İstatistiksel Özellikler

İyi bir robust tahmincide bulunması istenilen özellikler aşağıdaki gibi sıralanabilir. Ayrıca M-tahminciler sıralanan bu özellikleri sağlamaktadır.

1. Bir robust tahmincisi standart istatistiksel özellikleri sağlamalıdır. Bunlar tutarlılık, simetrik dağılımlar için yansızlık, asimptotik normallik ve dönüşümlerde equivariance. Kısaca bilgi vermek gerekirse, yansızlık ve tutarlılık bir tahminin, gerçek değere yakın çıkması için gereklidir. Asimptotik normallik $n \to \infty$ iken $\sqrt{n}(\hat{\theta} - \theta)$ nın normal dağılması özelliğidir. Equivariance ise dönüşüm ve değişimlerde gözlemlere bir sabit eklediğimizde veya çarptığımızda tahmincinin de aynı yönde değişmesidir. Kısaca konum ve ölçek equivariance özelliği şu şekilde ifade edilebilir: $(x_1,x_2,\ldots x_n)$ gözlemlerinin fonksiyonu olan $T(x_1,x_2,\ldots x_n)$ tahmincisi, $(ax_1+b,ax_2+b,\ldots ax_n+b)$ gözlemlerinin fonksiyonu olan $T(ax_1+b,ax_2+b,\ldots ax_n+b)$ tahmincisi olsun. Bu durumda $T(ax_1+b,ax_2+b,\ldots ax_n+b)=aT(x_1,x_2,\ldots x_n)+b$ ise T tahmincisi konum ve ölçek equivariance özelliğine sahiptir denir. (Wu 1985; Maronna ve diğerleri, 2006)

2. Robust tahminin değeri varsayımsal dağılımdan küçük sapmalar olduğunda çok az değişmelidir. Söz konusu sapmalar veri setinin küçük bir bölümündeki çok büyük sapma veya veri setinin büyük bir bölümünde çok küçük sapma olabilir. Bu tür tahmin edicilere dayanıklı (resistant) tahmin ediciler denilmektedir. (Wu, 2005)

3. Robust tahminin değeri varsayımsal dağılımdan büyük sapmalar olduğunda da çok sert bir şekilde değişmemelidir. Büyük sapmalar aynı zamanda dağılımın şeklini de değiştirmektedir. (Wu 2005)

4. Robust tahminci gözlem verilerinin dağılımı için etkin bir tahmin edici olmalıdır. Genellikle robust tahmin ediciler normal dağılımın merkezinde çok yüksek etkinliğe sahiptir fakat kuyruklarda ise farklıdır. (Wu, 2005)

5. Robust tahminci pratikte de uygulanabilir olmalıdırlar. Yani esnek, kullanımı ve tanımlanması kolay, maliyeti kabul edilebilir, her tür örneklem büyüklüklerine uygun olmalıdır. (Wu, 2005)

3.3. Robust Regresyonda Bazı Tanımlamalar

3.3.1. Etkileyici (influence) nokta: Eğer bir gözlem regresyon denklemini büyük ölçüde etkiliyorsa yani kendi yönüne büyük ölçüde çekebiliyorsa o nokta etkileyici noktadır. Diğer bir ifadeyle veri setinden bir nokta kaldırıldığında regresyon denkleminde büyük değişik oluyorsa o nokta etkileyici noktadır.(Sarkar ve diğerleri, 2011)

3.3.2. Kaldıraç (leverage) nokta: Bağımsız değişkeni diğerlerine göre çok büyük veya çok küçük değer alan noktadır. Bağımlı değişkenin aldığı değerlere göre regresyon denklemini olumlu veya olumsuz etkiler. Bazı durumlarda aynı zamanda etkileyici nokta da olabilmektedirler. (Montgomery ve diğerleri, 2001) İyi ve kötü olmak üzere ikiye ayrılırlar. (Wilcox, 2010)

iyi kaldıraç (high leverage) nokta: Sadece bağımsız değişkeni değil aynı zamanda bağımlı değişkeni de diğerlerine göre çok büyük veya çok küçük değer alan noktadır. Yani regresyon doğrusun geçtiği yerde bulunan noktadır.

kötü kaldıraç (low leverage) nokta: Sadece bağımsız değişkenin diğerlerine göre çok büyük veya çok küçük değer aldığı noktadır. Regresyon doğrusundan oldukça uzakta olan bir noktadır. Regresyon denklemini olumsuz yönde etkileyici bir noktadır. Aynı zamanda regresyon aykırı değerdir ve etkileyici noktadır.

3.3.3. Aykırı değer: Regresyon analizinde aykırı değer, büyük artığa sahip bir nokta olarak tanımlanmaktadır. Diğer bir ifadeyle, bir noktanın bağımlı değişkeninin diğerlerine göre oldukça farklı (daha büyük veya daha küçük) olmasıdır. Aşağıdaki şekilde sınıflandırılırlar.

regresyon aykırı değer: Bir noktanın veri setinde lineer bağlantısı bulunan diğer noktalardan oldukça farklı bir yönde yer almasıdır. Genelde diğer verilere göre büyük artıklara sahiptirler. Regresyon denklemi üzerinde olumsuz etkiye sahiptirler. Bu nedenle de tespit edilmesi en önemli ve en gerekli olan aykırı değerlerdir. Örnek vermek gerekirse kötü kaldıraç noktaları aynı zamanda

regresyon aykırı değerlerdir. Fakat iyi kaldıraç noktaları regresyon aykırı değer değillerdir (Montgomery ve diğerleri, 2001).

artık aykırı değer: Standartlaştırılmış veya student türü standartlaştırılmış artığı büyük olan noktadır. Artık aykırı değerler aynı zamanda regresyon aykırı değer de olmaktadır. Ama bunun tersi bazı özel durumlarda doğru olmayabilir. Örneğin bir nokta çok fazla etkileyici nokta ise regresyon doğrusunu kendisine çekeceğinden artığı küçük çıkabilir böylece de artık aykırı değeri olmaz ama regresyon aykırı değer olur (Montgomery ve diğerleri, 2001).

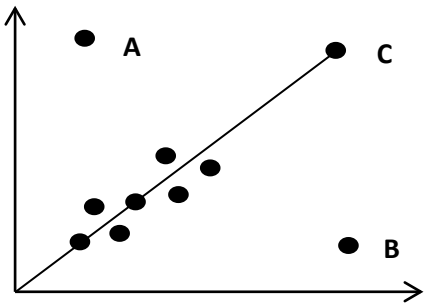
x-yönlü aykırı değer: Bir noktanın x yönünde diğer noktalara göre oldukça uzakta bulunmasıdır (Montgomery ve diğerleri, 2001).

y-yönlü aykırı değer: Bir noktanın y yönünde diğer noktalara göre oldukça uzakta bulunmasıdır (Montgomery ve diğerleri, 2001).

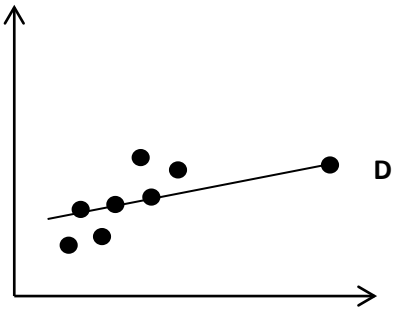
x ve y – yönlü aykırı değer: Bir noktanın hem x yönünde hem de y yönünde diğer noktalara göre oldukça uzakta bulunmasıdır. Bulunduğu konuma göre genellikle regresyon doğrusu üzerinde pek bir etkiye sahip değillerdir (Montgomery ve diğerleri, 2001).

Aykırı değer olan bir nokta yukarıdaki sınıflandırmaların birine girebileceği gibi birden fazlasına da girebilmektedir. Bu durumun oluşmasında temel etken ise aykırı değerin yönü ve etkisidir. Burada aykırı değerin yönü ele alınırsa doğruların genel istikametinde olan bir aykırı değer *regresyon aykırı değer* olmamaktadır. Aykırı değerin etkisi dikkate alındığında ise etkileyici olan bir aykırı değer noktası regresyon doğrusunu kendisine çekerek artık değerini küçültür ve *regresyon aykırı değer* sınıfında olmasına rağmen *artık aykırı değer* sınıfına girmez.

Aykırı değer ve kaldıraç noktaları daha iyi gösterebilmek için rasgele çeşitli noktalar oluşturulmuş ve şekil 3.1. ile şekil 3.2. da grafiklerle gösterilerek ilgili sınıflandırmaları çizelge 3.1.'de verilmiştir.



Şekil 3.1. Çeşitli aykırı değer noktaları



Şekil 3.2. Artık aykırı değer olmayıp regresyon artık değer olan nokta

Şekil 3.1 ve Şekil 3.2 yukarıda ifade edilen aykırı değer türlerini ve kaldıraç noktalarının anlamak bakımından önemlidir. Örneğin A, B ve D noktaları regresyon doğrusunu kendi yönlerine doğru çekeceğinden etkileyici noktalar ve B ve D aynı zamanda kötü kaldıraç nokta iken, C noktası ise doğrunun eğimini etkilemeyeceğinden iyi kaldıraç noktasıdır. Ayrıca A y-yönlü aykırı değer, B ise x-yönlü aykırı değerdir. Aşağıdaki tablo bu noktaları özellikleri bakımından sınıflandırmıştır.

Çizelge 3.1. Aykırı değer noktalarının sınıflandırılmaları

	A	**B**	**C**	**D**
Etkileyici Nokta	+	+	-	+
İyi Kaldıraç (High Leverage) Nokta	-	-	+	-
Kötü Kaldıraç (Low Leverage) Nokta	-	+	-	+
Regresyon Aykırı Değer	+	+	-	+
Artık Aykırı Değer	+	+	-	-
x-yönlü Aykırı Değer	-	+	-	+
y-yönlü Aykırı Değer	+	-	-	-
x ve y – yönlü Aykırı Değer	-	-	+	-

Ayrıca, yukarıda ifade edilen artık türü ne olursa olsun, en küçük kareler yönteminin bu artıklara duyarlı olduğu ve varlığı durumunda tutarsız tahminler ortaya çıkardığı bilinmektedir. (Rousseeuw ve Leroy, 1987).

3.3.4. Student türü standartlaştırılmış artıklar

Student türü artıklar aşağıdaki gibi tanımlanır,

$$r_i = \frac{e_i}{s\sqrt{1-h_{ii}}} \tag{3.1}$$

Burada e_i $i.$ artık, s ölçek tahmini, h_{ii} değeri ise $X(X'X)^{-1}X'$ şapka (hat) matrisinin $i.$ Köşegen elemanıdır ve $i.$ noktanın $x-uzayı$ üzerindeki uzaklığıdır.

Bu ölçeklendirme tekniği ile kaldıraç (leverage) noktalar daha iyi tespit edilirler. Kaldıraç noktaların artıkları genellikle küçüktür. Bu nedenle de $x-uzayı$ merkeze yakın noktalara göre varyansları da küçüktür. Bu durumda da tespit edilmeleri zordur. Standartlaştırılmış artıklar bu noktaları tespit etmekte yetersizdirler. Bu nedenle bu eksikliği ortadan kaldırmak için $x-uzayı$ üzerindeki uzaklığı dikkate alınarak $i.$ gözlemin etkisi standartlaştırmaya dahil edilir ve yeni bir standartlaştırma yapılır. Bu yeni standartlaştırmaya *student* türü standartlaştırma denilir (Montgomery, 2001).

3.4. Tahminciler İçin Kriterler

Genellikle tahminciler varyans veya yanlılık veya ikisinin kombinasyonundan oluşan hata kareler ortalaması (MSE) ölçütlerine göre değerlendirilmektedir. Öte yandan, tahmincilerin aykırı değerlere karşı dayanıklılığını değerlendirmede yaygın olarak kullanılan ölçütler ise etki fonksiyonu (influence function), kırılma noktası (breakdown point), gross hata duyarlılığı (gross error sensitivity) olarak bilinmektedir (Maronna ve diğerleri, 2006). Bu bölümde bu ölçütler tanıtılacaktır.

3.4.1. Etki fonksiyonu (influence curve – IC veya influnce function – IF)

Büyük örneklemlerde çok küçük bir orandaki *kirliliğin (contamination)* etkisinin tahminciyi nasıl etkilediğini gösteren ölçüttür. *T* tahminci, y ise örnek uzayda nokta olsun. Bu durumda *T* için etki fonksiyonu aşağıdaki gibi tanımlanır.

$$IF_{T,F}(y) = \lim_{\epsilon \to 0} \frac{T((1-\epsilon)F+\epsilon\Delta_y)-T(F)}{\epsilon} \qquad (3.2)$$

Burada, F varsayılan dağılım, ϵ küçük bir sayı, Δ_y y noktasında bir olasılık kütlesini ifade eden kümülatif dağılım fonksiyonudur. Etki fonksiyonuna *T* tahmincisinin ϵ' ye göre türevi olarak da bakılabilir. $T((1-\epsilon)F+\epsilon\Delta_y)$ ve $T(F)$ sırasıyla , $(1-\epsilon)F+\epsilon\Delta_y$'nin ve F'nin tahminleri olduğu için bu iki tahmin arasındaki fark etki fonksiyonu ile tahmin edilmektedir. Etki fonksiyonu y deki

kirliliğin (contamination'in) küçük miktarının $T(F)$ deki etkisini ölçmektedir (Rohan, 2011). Wu (1985) etki fonksiyonun varsayılan dağılımda küçük değişim altında tahmincinin asimptotik davranışının niceliksel bir resmini sağlamakta olduğunu ifade etmiştir (Wu, 1985).

Etki fonksiyonun sonlu örnek için ifadesi duyarlılık eğrisi (sensitive curve (SC)) olarak adlandırılmaktadır. $IF_{T,F}(y)$ de F yerine deneysel dağılım fonksiyonu F_n ve ϵ yerine 1/n+1 konulursa duyarlılık eğrisi aşağıdaki gibi ifade edilir (Goodall, 1983):

$$SC_n(y) = (n+1)\left[T\left(\frac{n}{n+1}F_n + \frac{1}{n+1}\Delta_y\right) - T(F_n)\right] \quad (3.3)$$

$$= (n+1)[T_{n+1}(y_1, \dots y_n, y) - T_n(y_1, \dots y_n)] \quad (3.4)$$

3.4.2. Kırılma noktası

Veri setinde aykırı değerler olsa dahi bir tahmincinin istikrarının ne kadar koruduğunun ölçüsüdür. Bir tahmincinin sağlamlığını (robustness) ölçmede kullanılır. Basit bir şekilde ifade etmek gerekirse, β'nın bir tahmincisi olan $\hat{\beta}$'nın kırılma noktası, tahmincinin gerçek parametre değeri hakkında hala doğru bilgi vermeye devam edecek kadar veri setindeki maksimum kirletme miktarıdır. Bir tahminci için en küçük kırılma noktası 1/n dir. (Wu, 1985; Maronna ve diğerleri, 2006, Montgomery, 2006).

n tane veriden oluşan $U=\{x_1,\dots,x_n\}$ örneklemi için $T(x_1,\dots,x_n)$ tahmincisinin değeri bilinmeyen parametre için tahmin değeri olsun. $T(x_1,\dots,x_n)$, $T(U)$ şeklinde gösterilsin.

Orijinal veri noktalarının herhangi m ($m<n/2$) tanesi keyfi değerlerle değiştirilsin ve elde edilen örneklem $\tilde{U}$ şeklinde, buna dayalı tahmin $T(\tilde{U})$ şeklinde gösterilsin. Böylece maksimum yanlılık

$$Bias(m, T, U) = \sup_{\tilde{U}} \left\| T(\tilde{U}) - T(U) \right\| \quad (3.5)$$

olarak tanımlanır. Böylece, sonlu örneklem için T tahmincisinin kırılma noktası

$$\varepsilon = min\{\frac{m}{n}, Bias(m,T,U) = \infty\} \quad (3.6)$$

şeklinde tanımlanır.

3.4.3. Büyük hata duyarlılığı (gross error sensitivity)

T tahmincisinin etki fonksiyonun mutlak değerinin maksimumu Küçük Hata Duyarlılığı (Gross error sensitivity) olarak ifade edilir.

$$\gamma = \sup_y \left| IF_{T,F}(y) \right| \quad (3.7)$$

T tahmincisinin etki fonksiyonu sınırlı ($\gamma < \infty$) ise bu durumda *T*-tahminleri sınırlı sağlam anlamında B-robust olarak adlandırılır (Rohan, 2011).

3.5. Regresyon Analizinde M-Tahminciler

İlk defa Huber tarafından 1964'de ortaya atılan M-tahminciler isimlerini en çok olabilirlik tahmincilerinden (Maximum Likelihood Estimator - MLE) almaktadırlar. Bu tahminciler, OLS tahmincilerinin kullandığı kalıntı kareleri yerine, kalıntıların daha az hızla artan bir fonksiyonunu kullanmayı önerir (montgomery, 2001).

Aşağıdaki iki koşulu sağlayan tahmincilere M-Tahminciler denir;

I. Artıkların bir fonksiyonunun minimize edilmesi

$$\underset{\beta}{\text{minimize}} \sum_{i=1}^{n} \rho(r_i) \quad (3.8)$$

Burada, r_i student türü standartlaştırılmış artıklar ($r_i = \frac{e_i}{s\sqrt{1-h_i}}$), ρ artıkların fonksiyonudur ve karesel fonksiyona göre daha az artacak şekilde seçilir.

II. $$\psi(r) = \frac{\partial}{\partial \beta}\rho(r) \quad (3.9)$$

olmak üzere, $\psi(.)$ etki veya kestirim fonksiyonu olarak isimlendirilir.

$$\sum_{i=1}^{n} \psi(r_i)\, x_{ij} = 0 \qquad j = 0,1,\dots,k \quad (3.10)$$

$\rho(r)$ konveks ve türevlenebilir ise çözüm tektir. (Maronna ve diğerleri, 2006)

Hatalar belirli varsayımları sağladığında ve dağılımı belirli dağılımlara uyduğunda (dağılımın logaritmasının negatif işaretlisi amaç fonksiyonu olarak alındığında) M-tahminciler en çok olabilirlik tahmincisine dönüşmektedirler. Dolayısıyla M-Tahminciler en çok olabilirlik tahmincilerini kapsarlar.

3.5.1. ρ amaç fonksiyonu

İyi tanımlanan bir ρ amaç fonksiyonu aşağıdaki özelliklere sahiptir:

I. $\rho(0) = 0$ (3.11)

II. $\rho(r) \geq 0$ (pozitif tanımlılık) (3.12)

III. $\rho(r) = \rho(-r)$ (simetriktik) (3.13)

IV. $0 < r_1 < r_2$ için $\rho(r_1) < \rho(r_2)$ (3.14)

V. ρ sürekli ve türevlenebilir olmalıdır.

(Maronna ve diğerleri, 2006; Ali ve diğerleri 2006)

3.5.2. ψ fonksiyonu

ψ ile gösterilen fonksiyon, her verinin, parametre tahmin edicisi üzerindeki marjinal etkisini ölçmektedir. Bu fonksiyon o kadar önemlidir ki ψ'nin durumlarına göre M-tahmincisi farklı isimler almaktadır. Regresyon katsayılarının M-tahminleri

$$\sum_{i=1}^{n} \psi(r_i)\, x_{ij} = 0 \qquad j = 0,1,\dots,k \tag{3.15}$$

denkleminin iteratif yöntemlerle çözümü ile bulunur.

ψ'nin bazı özellikleri (Bell, 1980) aşağıdaki gibi sıralanabilir:

I. ψ asimetriktir, $\psi(-r) = -\psi(r)$

II. $\psi \geq 0, r \geq 0$ için

III. $\psi(r)$ sürekli ve parçalı türevlenebilir bir fonksiyondur

IV. $\psi'(0) = 1$

V. $\psi''(0) = 0$

VI. $\psi'''(0) < 0$

3.5.3. w ağırlık fonksiyonu

Ağırlık fonksiyonu da aşağıdaki gibi tanımlanır;

$$w_i = \frac{\psi(r_i)}{r_i} \tag{3.16}$$

Böylelikle (3.10) denklemi aşağıdaki gibi olur;

$$\sum_{i=1}^{n} w_i r_i x_{ij} = 0 \tag{3.17}$$

(3.15) ve (3.17) denklemleri iteratif tekniklerle çözülmektedir. Bu teknikler içinde en uygun kullanılanı Yeniden Ağırlıklandırılmış En Küçük Kareler (Iteratively Reweighted Least Squares-IRWLS) tekniği' dir (O'leary, 1990; Holland and Welsch, 1977).

3.5.4. M-Tahminciler için yeniden ağırlıklandırılmış en küçük kareler (iteratively reweighted least squares-IRWLS) tekniği

IRWLS tekniği adını ağırlıklandırılmış en küçük kareler'in bir döngü içerisinde yenilenmesinden alır. Her döngü içerisinde her bir gözleme karşı gelen ağırlıklar kullanılarak regresyon tahminleri yapılır. Her bir döngüde de ağırlıklar değişmektedir çünkü her yineleme de yeni artıklar hesaplanmaktadır. Burada bir not eklenirse başlangıç tahmini için ağırlıklar olmayacağından (çünkü henüz artıklar elde edilmemiştir) bir başlangıç tahmincisi kullanılması gereklidir. Bu yinelemeler veya döngüler katsayıların bir önceki katsayılardan çok az bir farklılık göstermesine kadar devam eder. IRWLS tekniği genel olarak ağılıkların seçimine dayanmaktadır (Heiberger, 1992; Holland and Welsch, 1977).

IRWLS algoritma adımları:

1. Bir başlangıç tahmincisiyle ilk artıklar hesaplanır (e_i)
2. Bu artıklar yardımıyla ölçek tahmincisi hesaplanır (s)
3. Artıklar ölçek tahmincisi yardımıyla standartlaştırılır (r_i)
4. Standartlaştırılmış artıkların M-tahmincilerinin ağırlık fonksiyonu yardımıyla ağılıkları bulunur (w_i)

5. Ağırlıklar yardımıyla IRWLS tekniği kullanılarak regresyon katsayıları tahmin edilir

   $$\hat{\beta} = (X'WX)^{-1}X'WY \qquad (3.18)$$

   Burada W köşegen elemanları ağırlık olmak üzere köşegen (diagonal) matristir.
6. Bulunan katsayılarla bir önceki katsayılar arasındaki fark çok küçük ise işlem sonlandırılır değilse bulunan katsayılarla yeni artıklar hesaplanarak 3. Adıma dönülerek yakınsama sağlanana kadar işlemlere devam edilir.

3.5.5. M-Tahminci teknikleri ve verdikleri ağırlıklara göre sınıflandırılmaları

M-Tahminciler verdikleri ağırlıklara göre aşağıdaki şekilde sınıflandırırlar;

I. Monoton m-tahminciler

Belirli bir noktaya kadar doğrusal azalan ağırlık, belirli bir noktadan sonra ise ya sabit ağırlık veren ya da çok az artışlarla ağırlık veren tahmincilerdir. Hiçbir zaman sıfır ağırlık vermezler.

II. Yeniden azalan (redescending) m-tahminciler

Yeniden Azalan (Redescending) M-tahminciler en popüler olan M-tahmincilerdir. ψ fonksiyonu orijinin çevresinde azalan değildir fakat orijinden uzaklaştıkça değeri sıfıra gider. ψ fonksiyonları sıfıra düzgün bir biçimde azalacak şekilde seçilir.

Regresyon analizinde bazı yeniden azalan (redescending) tahminciler maksimum kırılma noktasına ulaşmakla beraber bazıları da etkinliği maksimuma çıkarmaktadırlar.

Genellikle yeniden azalan (redescending)M-tahminciler aykırı değerlere karşı monoton tahmincilerden daha dayanıklıdır ancak tahminler elde edilirken kullanılan başlangıç değerlere de daha fazla duyarlıdır. (Rohan, 2011).

İki sınıfa ayrılırlar;

a. Yavaş yeniden azalan (soft redescending) m-tahminciler

Sonsuza giderken büyük artıklara sıfıra yakın ağırlıklar veren tahmincilerdir $\lim_{n\to\infty} \psi(\mathrm{r}) = 0$.

b. Hızlı yeniden azalan (hard redescending) m-tahminciler

Belirli bir noktadan sonraki artıklara sıfır ağırlık veren tahmincilerdir. $X > |c|$ için $\psi(\mathrm{r}) = 0$ sağlanır, c burada minimum reddetme noktasıdır. Bu M-tahminciler direk olmayan aykırı değer teşhis metodu olarak da kullanılabilir (Billor ve Kıral, 2008).

3.5.6. M-tahmincilerin asimptotik normalliği

Regresyon katsayılarının M- tahminlerinin oluşturduğu vektör $\hat{\beta}$ asimptotik olarak normal dağılıma sahiptir.

$$\hat{\beta} \sim N(\beta, V) \tag{3.19}$$

$\hat{\beta}$ için asimptotik varyans matrisi aşağıdaki şekildedir:

$$V = \frac{E(\psi^2)}{[E(\psi')]^2}(X'X)^{-1} \tag{3.20}$$

(Dasiou ve diğerleri, 1999; Maronna ve diğerleri, 2006)

Böylece, $\hat{\beta}$' nın tahmin edilmiş varyans matrisi aşağıdaki şekildedir.

$$\widehat{V(\hat{\beta})} = \frac{n^{-1}\sum[\psi(r_i)]^2}{[n^{-1}\sum\psi'(r_i)]^2}(X'X)^{-1} \tag{3.21}$$

(Rohan, 2011).

3.5.7. Uygulamada en çok kullanılan m-tahminciler

İlk olarak Huber tarafından 1964'de M-tahmincileri tanıtılmış ve Huber OLS ve LAD' yi birlikte ele alan bir $\rho(.)$ fonksiyonunu önermiştir. Bu fonksiyon, belirlenmiş aralıklar arasında gözlemlere eşit ağırlık, bu aralık dışında ise azalarak ağırlık vermektedir. Böylece tahmincinin etkinliği ve robustlığı oldukça yüksek olmaktadır. Huber'in M-tahmincisi etki fonksiyonuna göre monoton özelliğe sahiptir. Benzer şekilde Fair ve Lojistik M-tahmincilerde hemen hemen aynı özelliklere sahiptir ancak etki fonksiyonları belli bir yerden sonra

sabit gitmemektedir. Öte yandan, Andrew (1974)'in sinüs fonksiyonu normal dağılım için en iyi etkinliğe sahip fonksiyonlardan biridir. Andrew (1974) ve Tukey (1974)'in Bisquare M-tahmincilerinin maç fonksiyonu fonksiyonları benzerlik göstermektedir. Fakat Bell tekniğinde tek bir fonksiyon tanımlıdır yani bir reddetme noktası yoktur ama yine de merkezden uzak verilere oldukça küçük ağırlıklar vererek katsayı tahmini üzerindeki etkisini neredeyse sıfıra indirmiştir.

Robust literatüründe, Hampel üç parçalı hızlı yeniden azalan (hard redescending) bir fonksiyonu M-tahmincisi olarak önermesi ile yeniden azalan (redescending) M-tahmincileri popüler olmaya başlamıştır. yeniden azalan (redescending) M-tahmincileri, aykırı değerleri bir anda değil de bir geçiş bölgesinden sonra tamamen atan bir tahmincidir. Bu sayede de etkinliği oldukça yüksektir. Çünkü merkeze yakın (sıfırın komşuluğundaki) gözlemlere ilk kritik değere kadar maksimum ağırlık vermekte ve merkezden uzaklaştıkça da ikinci kritik değere kadar sabit bir ağırlık, üçüncü kritik değere kadar da azalan bir ağırlık vermekte üçüncü kritik değerden sonra da 0 vererek aykırı değerlerin etkisini tamamen ortadan kaldırmaktadır. Asad'ın M tahmincisi (Ali ve diğerleri, 2006) ise belirli bir yere kadar doğrusala yakın daha sonra ise azalacak şekilde bir etki ψ fonksiyonuna sahiptir. Bir noktadan sonra ise gözlemlere sıfır değerini vermektedir. $\psi 1$ ve $\psi 2$ her iki teknikte Winsor'un prensibine dayanılarak oluşturulmuştur (Ali ve diğerleri, 2006). Yani tüm dağılımların ortada aslında normal dağıldığını göz önüne alarak bir tahminci sağlamaktadır. Belirli bir yerden sonra ise hızlı bir şekilde verilen ağırlıkları azaltarak belirli bir noktadan sonra tamamen sıfır vermektedir. Bu tekniğin diğer tekniklerden farkı tama yakın ağırlık verme aralığını diğer tahmincilere göre daha uzun tutmasıdır böylece aykırı değerler dışındaki artıklar normal dağıldığında etkinliği diğer tahmincilere göre yüksek olmaktadır. Insha'nın M-tahmincisi (Ullah ve diğerleri, 2006) her yerde sürekli türevlenebilir bir fonksiyona sahiptir ve Tukey, Asad, $\psi 1$ ve $\psi 2$ ye göre tama yakın ağırlık verilen aralık daha geniştir ve bir reddetme noktası yoktur. Bu nedenle de hiçbir zaman gözlemlere sıfır değeri vermemektedir fakat merkezden uzak verilere çok düşük ağırlıklar vererek tahminci üzerinde aykırı değerlerin etkilerini yok denecek kadar küçültür.

M tahmincileri esnek, hesaplanması kolay ve asimptotik teorisinin bilinmesi bakımından en çok kullanılan robust tahmincilerdir. Ancak, x-yönündeki aykırı değerlere karşı duyarlı olmaları ve uygulamalarda $\rho(.)$'nun seçimine ilişkin bir sonucunun bulunmamasından dolayı dezavantajlara sahiptir.

Literatürde bilinen M-Tahminciler aşağıdaki tabloda etki fonksiyonlarına göre sınıflandırılarak verilmiştir.

Çizelge 3.2. Uygulamada en çok kullanılan M-tahminciler

	Fonksiyon Adı	Sınıfı	Ayaralama Sabiti	$\psi(r)$
1	**Andrew's Wave (veya sine)**	Hızlı Yeniden Azalan	$a = 1{,}339$	$\psi(r) = \begin{cases} \sin(r/a) & r \le a\pi \\ 0 & r > a\pi \end{cases}$
2	**Bell**	Yavaş Yeniden Azalan	---	$\psi(r) = r\left(1 + \frac{r^2}{5}\right)^{-3} \quad \lvert r\rvert < \infty$
3	**Cauchy**	Yavaş Yeniden Azalan	$c = 2{,}385$	$\psi(r) = \frac{r}{1 + \left(\frac{r}{c}\right)^2} \quad \lvert r\rvert < \infty$
4	**Danish**	Hızlı Yeniden Azalan	$c = 2\sim3$	$\psi(r) = \begin{cases} r & \lvert r\rvert \le c \\ \mathrm{r}\exp(-r^2/c^2) & \lvert r\rvert > c \end{cases}$
5	**Fair**	Monoton	$c = 1{,}4$	$\psi(r) = \frac{r}{1 + \lvert r\rvert/c} \quad \lvert r\rvert < \infty$
6	**Geman and Mcclure**	Yavaş Yeniden Azalan	---	$\psi(r) = \frac{2r}{(1 + r^2)^2} \quad \lvert r\rvert < \infty$
7	**Hampel's 17A**	Hızlı Yeniden Azalan	$a = 1{,}7$ $b = 3{,}4$ $c = 8{,}5$	$\psi(r) = \begin{cases} r & \lvert r\rvert \le a \\ asign(r) & a < \lvert r\rvert \le b \\ \frac{asign(r)(c - \lvert r\rvert)}{c - b} & b < \lvert r\rvert \le c \\ 0 & \lvert r\rvert > c \end{cases}$
8	**Huber**	Monoton	$c = 1{,}345$	$\psi(r) = \begin{cases} r & \lvert r\rvert \le c \\ csign(r) & \lvert r\rvert > c \end{cases}$
9	**Logistic**	Monoton	$c = 1{,}205$	$\psi(r) = ctanh(r/c) \quad \lvert r\rvert < \infty$
10	**Quadratic Squares Root Estimator (QSR)**	Yavaş Yeniden Azalan	$c = 1{,}264$	$\psi(r) = \begin{cases} r & \lvert r\rvert \le c \\ c^{3/2}\frac{sign(r)}{\sqrt{\lvert r\rvert}} & \lvert r\rvert > c \end{cases}$
11	**Ramsay's E_α**	Yavaş Yeniden Azalan	$c = 0{,}3$	$\psi(r) = \mathrm{r}\exp(-c\lvert r\rvert) \quad \lvert r\rvert < \infty$
12	**Welsh**	Yavaş Yeniden Azalan	$c = 2{,}985$	$\psi(r) = \mathrm{r}\exp(-(r/c)^2) \quad \lvert r\rvert < \infty$

Çizelge 3.2. (Devam) Uygulamada en çok kullanılan M-tahminciler

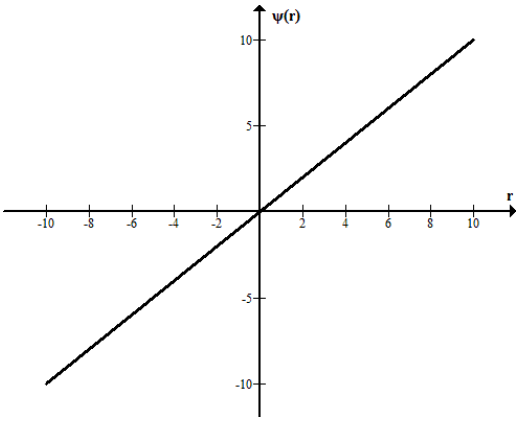
13	**Talwar**	Hızlı Yeniden Azalan	$\boldsymbol{c = 2,795}$	$\boldsymbol{\psi(r) = \begin{cases} r & \lvert r\rvert \le c \\ 0 & \lvert r\rvert > c \end{cases}}$
14	**Tukey'in Biweight (veya Bisquare)**	Hızlı Yeniden Azalan	$c = 4{,}685$	$\psi(r) = \begin{cases} r\left[1 - \left(\frac{r}{c}\right)^2\right]^2 & \lvert r\rvert \le c \\ 0 & \lvert r\rvert > c \end{cases}$
15	**Asad**	Hızlı Yeniden Azalan	$c = 4{,}685$	$\psi(r) = \begin{cases} \frac{2}{3} r\left[1 - \left(\frac{r}{c}\right)^4\right]^2 & \lvert r\rvert \le c \\ 0 & \lvert r\rvert > c \end{cases}$
16	**Insha**	Yavaş Yeniden Azalan	$c = 4{,}685$	$\psi(r) = r\left[1 + \left(\frac{r}{c}\right)^4\right]^{-2} \quad \lvert r\rvert < \infty$
17	**Qadir**	Hızlı Yeniden Azalan	$c = 4{,}685$	$\psi(r) = \begin{cases} \frac{r}{16}\left[1 - \left(\frac{r}{c}\right)^2\right]^2 & \lvert r\rvert \le c \\ 0 & \lvert r\rvert > c \end{cases}$
18	$\boldsymbol{\psi 1}$	Hızlı Yeniden Azalan	$c = 4{,}685$	$\psi(r) = \begin{cases} \frac{r}{2}\left[1 - \left(\frac{r}{c}\right)^6\right]^2 & \lvert r\rvert \le c \\ 0 & \lvert r\rvert > c \end{cases}$
19	$\boldsymbol{\psi 2}$	Hızlı Yeniden Azalan	$c = 4{,}685$	$\psi(r) = \begin{cases} \frac{r}{2}\left[1 - \left(\frac{r}{c}\right)^8\right]^2 & \lvert r\rvert \le c \\ 0 & \lvert r\rvert > c \end{cases}$

Tablonun üçüncü sütunda verilen sabit değerler ayarlama sabiti (tuning constant) olarak da bilinir. Bir tahmin edicinin aykırı değerlere karşı sağlamlığını ve aykırı değerlerin yokluğunda tahmincinin etkinliğini belirlemede yardımcıdır. Burada etkinlikten kasıt hatalar normal dağıldığında tekniğin ne kadar iyi performans sergilediğidir. Buradan da anlaşıldığı üzere tahmin edicinin hata dağılımları üzerindeki performansını etkileyen katsayıdır. Bazı teknikler için ayarlama sabitine küçük değer verildiğinde tekniğin dayanıklılığı artmaktadır fakat aynı zamanda etkinliği azalmaktadır. Bu nedenle ayarlama sabitinin değeri tahmin edicinin hem dayanıklılığını yüksek hem de normal dağılımda etkinliğini yüksek yapacak şekilde optimum olmalıdır (Holland ve Welsch, 1977). Yukarıdaki tablonun üçüncü sütununda verilen ayarlama sabitlerinin çoğu standart normal dağılımda M-tahmincilerinin etkinliği %95 olacak şekilde seçilmiştir, diğerleri için ise araştırmacılar tarafından önerilen sabitler kullanılmıştır.

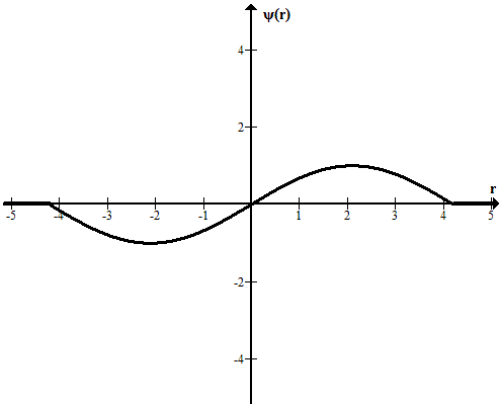
Öte yandan M-tahmincileri red noktası ise belli bir değerden sonra ψ fonksiyonunun sıfır olarak alınmasını belirleyen noktadır. Bu noktanın ilerisinde

kalan noktalar tahmin üzerinde sıfır etkiye sahip olacaklardır. Bu şekilde bir sonlu red noktasına sahip tahmincilere Hızlı Yeniden Azalan Tahminci (Hard Redescending Estimator) denir ve bu tahminciler büyük aykırı değerlere karşı çok iyi korunurlar. Ayrıca, bu tahminciler direk olmayan aykırı değer teşhis metodu olarak da çalışırlar (Billor ve Kiral, 2008)

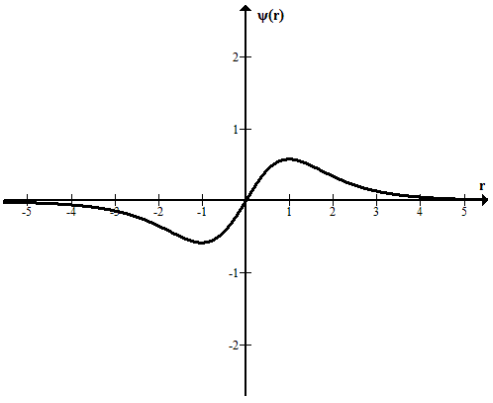
Yukarıda Çizelge 3.2 de verilen M-tahmincilerinin etki fonksiyonları grafikleri aşağıda verilmiştir. Bu grafiklere bakılarak da tahmincilerin aykırı değerlere karşı dayanıklılığı hakkında kabaca yorum yapılabilir. Örneğin EKK tahmincisin etkisi, artık büyüdükçe artmakta, Huber'in M-tahmincisinin etkisi, belli bir noktaya kadar artmakta bu noktadan sonra sabit olarak devam etmektedir. Andrew'in M-tahmincisi ise artık belli bir değeri aştıktan sonra tahminci üzerinde bu değerlerin etkisi olmamaktadır.



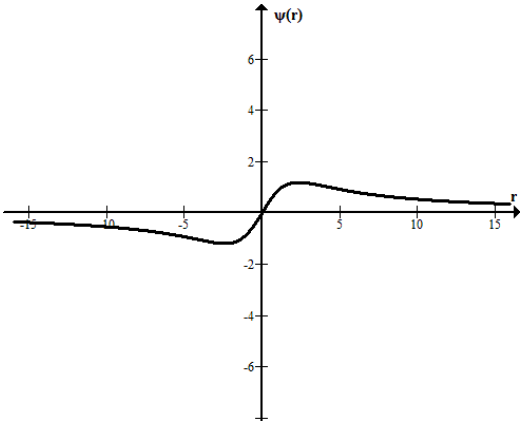
Şekil 3.3. OLS ψ fonksiyonu grafiği



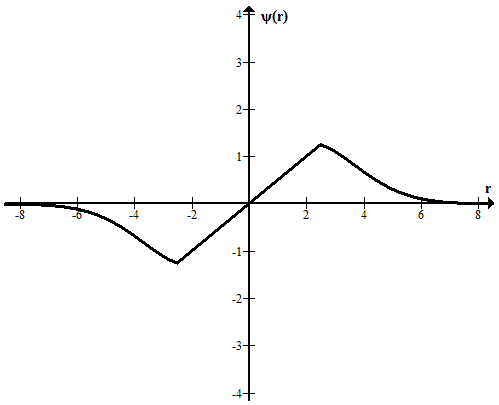
Şekil 3.4. Anrew'in sine ψ fonksiyonu grafiği



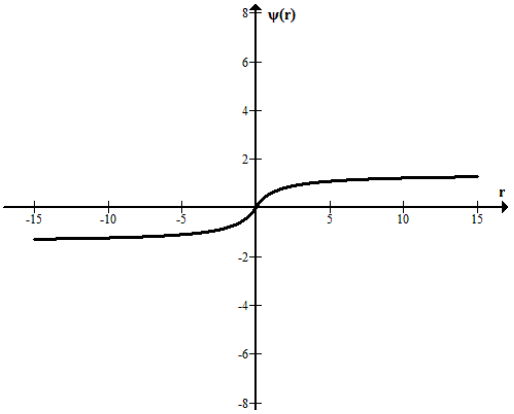
Şekil 3.5. Bell ψ fonksiyonu grafiği



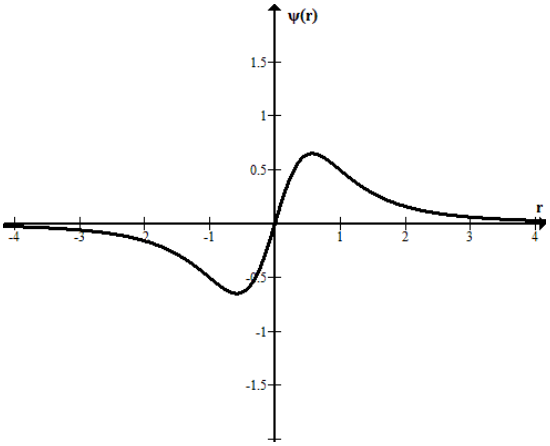
Şekil 3.6. Cauchy ψ fonksiyonu grafiği



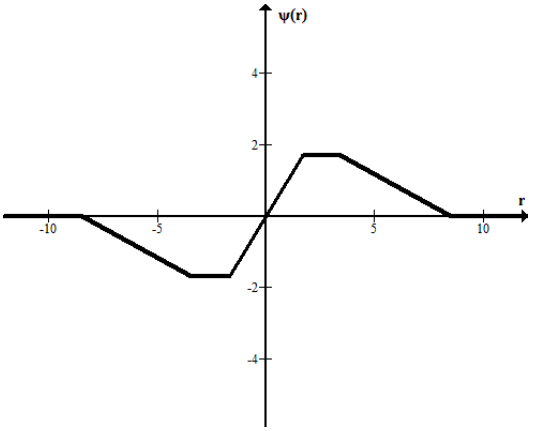
Şekil 3.7. Danish ψ fonksiyonu grafiği



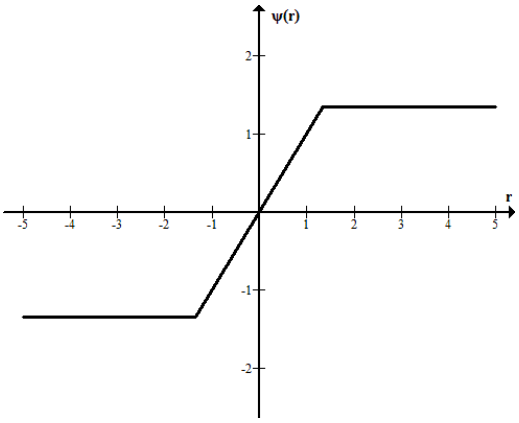
Şekil 3.8. Fair ψ fonksiyonu grafiği



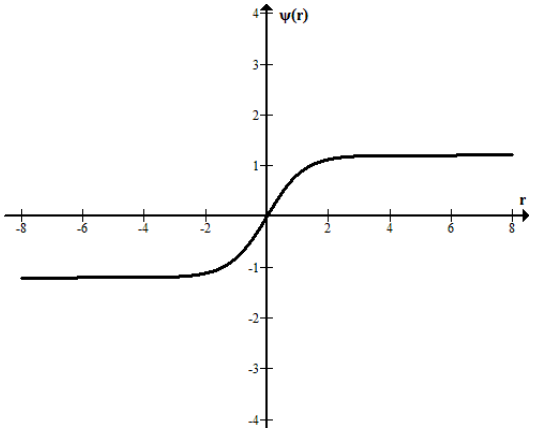
Şekil 3.9. Geman ve Mcclure ψ fonksiyonu grafiği



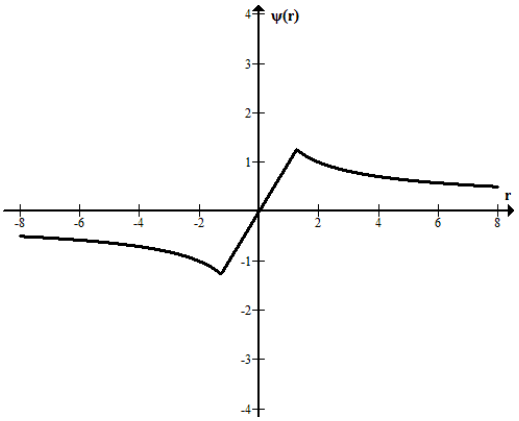
Şekil 3.10. Hampel ψ fonksiyonu grafiği



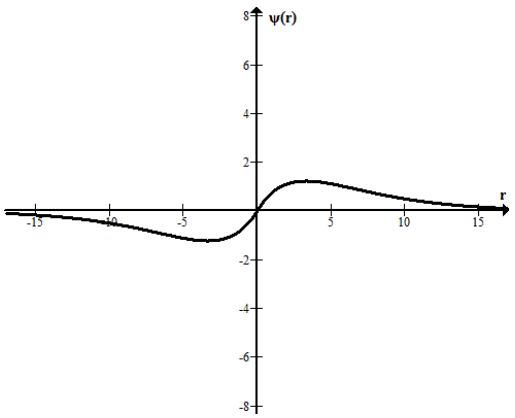
Şekil 3.11. Huber ψ fonksiyonu grafiği



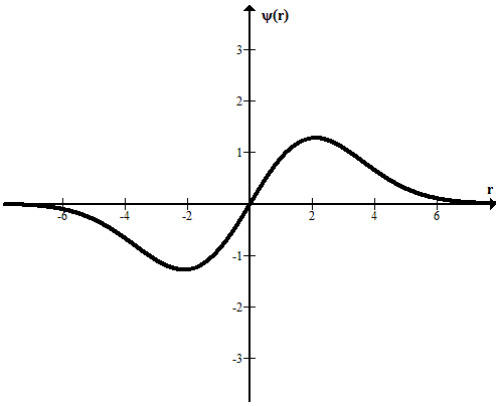
Şekil 3.12. Logistic ψ fonksiyonu grafiği



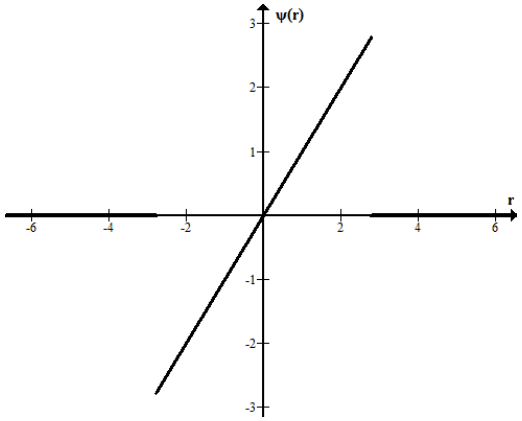
Şekil 3.13. QSR ψ fonksiyonu grafiği



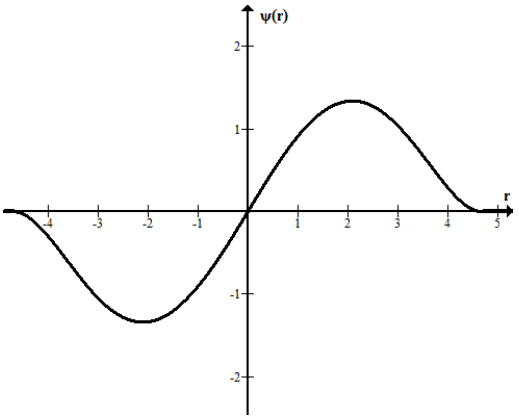
Şekil 3.14. Ramsay ψ fonksiyonu grafiği



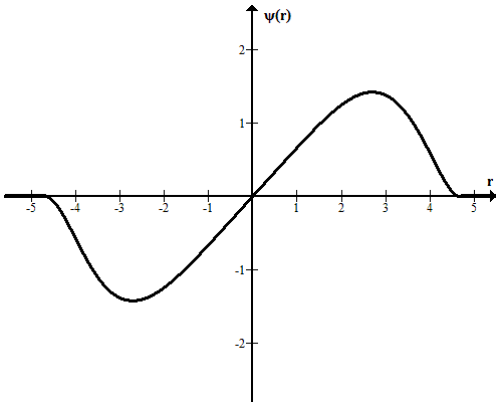
Şekil 3.15. Welsch ψ fonksiyonu grafiği



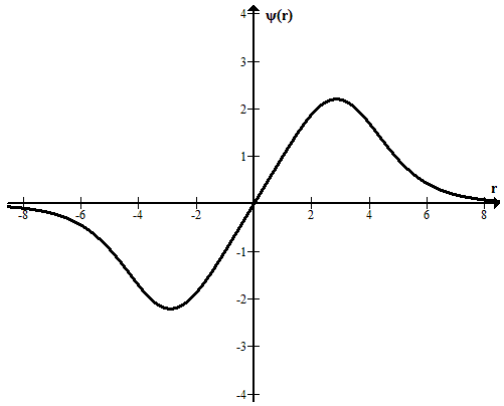
Şekil 3.16. Talwar ψ fonksiyonu grafiği



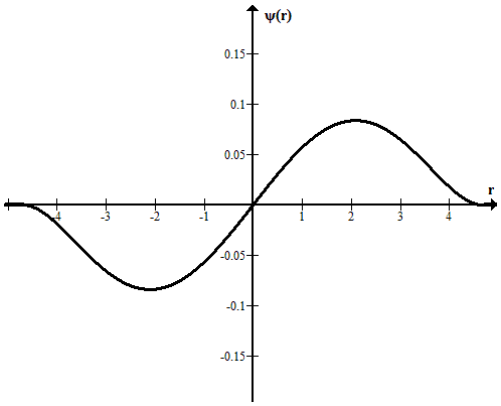
Şekil 3.17. Tukey ψ fonksiyonu grafiği



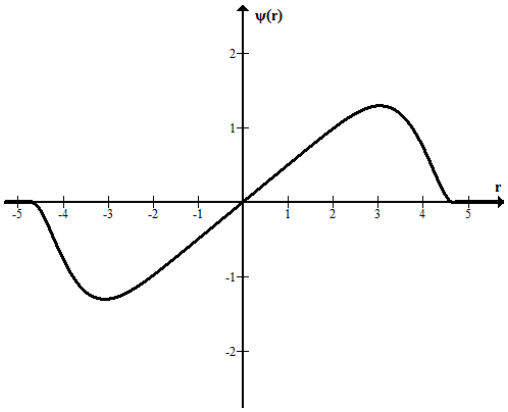
Şekil 3.18. Asad ψ fonksiyonu grafiği



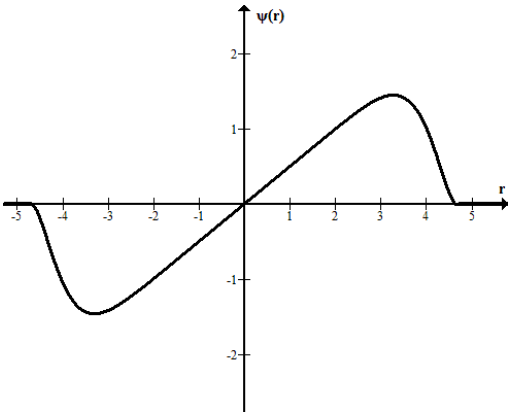
Şekil 3.19. Insha ψ fonksiyonu grafiği



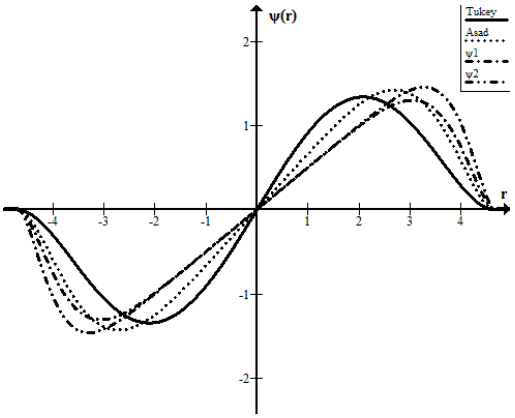
Şekil 3.20. Qadir ψ fonksiyonu grafiği



Şekil 3.21. $\psi 1$ ψ fonksiyonu grafiği



Şekil 3.22. $\psi 2$ ψ fonksiyonu grafiği



Şekil 3.23. Tukey – Asad – $\psi 1$– $\psi 2$ ψ fonksiyonu grafiği

Bir M-Tahmincisinin performansı, seçilen $\rho(r)$ veya $\psi(r)'$ye ve bununla birlikte belirlenen ayarlama sabitine bağlıdır.

3.6. IRWLS Algoritması İçin Önerilen Bazı Başlangıç Tahmincileri

IRWLS Algoritması gereği M-tahminleri başlangıç değerlerine ihtiyaç duymaktadır. Bu başlangıç değerleri için genellikle klasik teknikler kullanılmaktadır. Örneğin artıkların (e) elde edilmesi için OLS ve ölçek Tahmincisi (s) için MAD kullanılmaktadır. Bu başlangıç tahmin değerlerinin dolayısı ile başlangıç tekniklerinin M-tahmincilerin performansını etkilediği bilinmektedir (Holland ve Welsch, 1977). Özellikle hata dağılımı simetrik olmadığında MAD 'ın etkinlik kaybı olduğu görülmüştür. Bu durumu yine Andrews'in (1972) yaptığı similasyon çalışması da göstermektedir. Bu çalışmada ortalamaya veya standart sapmaya dayalı tekniklerin performansları dayanıklı tekniklerin performanslarıyla karşılaştırılmıştır. Öte yandan, Green (1984) de çeşitli örnekler üzerinde iyi başlangıç değerlerin yakınsama hızını etkilediğini ve başlangıç parametrelerinin seçiminin IRWLS algoritmasını duyarlılığını etkilediğini göstermiştir. Genel olarak kabul edilen kanı, yüksek kırılma noktasına sahip tahminlerin regresyon katsayılarının başlangıç tahminleri olarak ele alınmasıdır.

3.6.1. IRWLS tekniğinde başlangıç tahmin değerlerini elde etmek için kullanılan bazı dayanıklı regresyon tahmincileri

Regresyon analizinde yeniden ağırlıklandırılmış en küçük kareler (IRWLS) tekniği ile M-tahminciler hesaplanma süreci artıklara ve ağırlıklara dayanmaktadır. İlk artığın elde edilmesi için M-tahminciler kendi tahmincileri dışında bir tahmincinin kullanılmasını hesaplama tekniğinden dolayı zorunlu kılmaktadırlar. IRWLS tekniğinde ilk artığın elde edilmesi ile M-tahmincilerin ağırlık verme süreci ile başlar ve diğer adımlarla sonuca ulaşılır. Bu süreçte ilk artıklar sınır ve ağırlık değerlerini etkileyeceğinden sonuçlar da etkilenecektir. Genellikle başlangıç için artıkları elde etmede OLS tekniği kullanılmaktadır fakat OLS tekniği aykırı değerlere duyarlı bir teknik olduğundan IRWLS sürecini de etkileyecek ve bu durumda da sonuçlar robust olmayacaktır. Çünkü ilk artıklar elde edilirken aşırı büyük aykırı değerler bazı durumlarda regresyon doğrusunu kendilerine doğru çekeceklerinden kendi hesaplanan artık değerlerini küçültebilmektedirler. Böylece IRWLS sürecinde artıklara ağırlık verilirken aykırı

değer aralığına düşmeyeceklerinden tama yakın ağırlık alacaklar ve tahminci üzerindeki etkilerini devam ettireceklerdir. Bu sebepten dolayı M-Tahmincilerde, özellikle yeniden azalan (redescending) olanlar üzerinde başlangıç tahmincilerinin seçimi kritik bir rol üstlenmektedir. Bu çalışmada başlangıç tahmincisi olarak duyarlı OLS tekniğine alternatif dayanıklı teknikler olan Budanmış En Küçük Kareler (LTS – Least Trimmed Squares) tekniği, En Küçük Mutlak Sapmalar (LAD – Least Absolute Deviations) tekniği, Andrews'in Medyan Regresyon (ARM) Tekniği, Theil tekniği ve Siegel'in Tekrarlı Medyan tekniği kullanılarak sonuçlardaki değişimler incelenecektir.

3.6.1.1. En küçük budanmış kareler (LTS – least trimmed squares) tahmincisi

Rouseeuw tarafından 1984 de önerilmiştir. Amaç fonksiyonu

$$\underset{\hat{\beta}}{\text{minimize}} \sum_{i=1}^{h} e_{(i)}^2 \tag{3.22}$$

şeklindedir, burada artıklar $e_{(1)}^2 < e_{(2)}^2 \ldots < e_{(n)}^2$ şeklinde sıralı artıklar h ise alt grup (subset) tir. Böylelikle LTS tahmini aslında büyük artıklardan arındırılmış h birimlik bir alt grup seçme ve OLS tekniğinin bu alt gruba uygulanmasıdır. Burada h seçimi önemli rol oynamaktadır ve Rouseeuw ve Leroy (1987) tarafından

$$h = [n(1-\alpha)] + 1 \tag{3.23}$$

şeklinde seçilmesi önerilmektedir. α burada atılan veri oranıdır. Eğer $h = n/2$ alınırsa kırılma noktası %50 ye ulaşmaktadır.

3.6.1.2. En küçük mutlak sapmalar (LAD – least absolute deviations) tahmincisi

Tekniğin amacı hata mutlak sapmalar toplamını minimum yapacak katsayı değerlerini tahmin etmektir. En küçük mutlak sapmalar (LAD) tekniği uzun kuyruklu dağılımlar için en uygun tahmincilerden biridir. Bundan dolayı IRWLS'de kullanılması önerilen tahmincilerden de birisidir. Laplace tarafından önerilen tekniğin amaç fonksiyonu,

$$\underset{\beta}{\text{minimize}} \sum_{i=1}^{n} |e_i| \tag{3.24}$$

şeklindedir. Bu tezde LAD tahminleri Barrodale ve Roberts (1973) 'in önerdiği algoritma yardımıyla elde edilmektedir. Bu teknik Huber tarafından tanımlanan M-tahmincide de kullanılmıştır,

$$\rho(r) = \begin{cases} \frac{1}{2}r^2 & |r| < c \\ c|r| - \frac{1}{2}c^2 & |r| \geq c \end{cases} \tag{3.25}$$

3.6.1.3. Andrew'in medyan regresyon (ARM – Andrew's regression by medians) tahmincisi

Andrew tarafından 1974 yılında M-tahmincilerin etkinliğini artırmak için geliştirilmiştir. Andrew normal dağılımayan veri setleri için robust tahmincileri incelemiş ve IRWLS algoritmasında başlangıç tahminlerin önemini fark etmiştir. Bu nedenle kırılma noktası yüksek (bu tezde %25 olacak şekilde p_1 ve p_2 değeri verildi) bir tahminci geliştirerek M-tahmincilerin etkinliğini artırmayı başarmıştır.

ALGORİTMA:

1. Adım: x_i 'ler küçükten büyüğe doğru sıralanır ve p_1 ile p_2 oranları belirlenir

2. Adım: $p_1 n$ kadar en küçük ve en büyük x veri setinden atılır

3. Adım: $p_2 n$ kadar veri $median\{x_i\}$ 'nin sağından ve solundan atılır

4. Adım: Geri kalan ikiye bölünmüş veri setinin küçük olanı L büyük olanı H ile ifade edilir

5. Adım: $\hat{\beta}_1$ ve $\hat{\beta}_0$ hesaplanır

$$\hat{\beta}_1 = \frac{\underset{H}{\text{median}}\{y_{hi}\} - \underset{L}{\text{median}}\{y_{li}\}}{\underset{H}{\text{median}}\{x_{hi}\} - \underset{L}{\text{median}}\{x_{li}\}} \tag{3.26}$$

$$\hat{\beta}_0 = \underset{i}{\text{median}}\left(y_i - \hat{\beta}_1 x_i\right) \tag{3.27}$$

3.6.1.4. Theil tahmincisi

Bu tahminci 1950 yılında Theil tarafından önerilmiştir. Kırılma noktası %29.3'dür (Siegel, 1982). Theil aykırı değerlerden şüphelenildiğinde bu

tahmincinin çok kullanışlı olduğunu kanıtlamıştır. Fakat veri sayısı büyüdükçe ve ikiden fazla değişken olduğunda bu tekniğin uygulanması zordur. Bir doğrunun eğimi tahmininde kullanılan Theil tekniği (x_i, y_i) ve (x_j, y_j) gözlem çiftlerinden hesaplanan $\binom{n}{2}$ kadar eğim değerlerinin medyanı hesabına dayandırılmaktadır. Theil tekniğinde $\hat{\beta}_0$ve $\hat{\beta}_1$ öyle tahmin edilmeli ki e_i artıkların medyanı sıfır olmalıdır.

$$\hat{\beta}_1 = \underset{1\le i<j\le n}{\text{median}} \frac{y_j - y_i}{x_j - x_i} \quad i = 1,2,\dots,n-1 \qquad j = i+1,\dots,n \tag{3.28}$$

$$\hat{\beta}_0 = \underset{i}{\text{median}}(y_i - \hat{\beta}_1 x_i) \tag{3.29}$$

Medyan tahmin edicilerle yapılan kestirimler ortalama ile yapılan tahmin edicilerle karşılaştırıldıklarında aykırı değerlerden daha az etkilenirler.

3.6.1.5. Siegel'in tekrarlı medyanlar tahmincisi

Tekrarlı Medyanlar tahmincisi Siegel tarafından 1982 yılında yüksek kırılma noktasına sahip bir kestirici elde etmek için geliştirilmiştir. Kırılma noktası %50'dir. (x_i, y_i) veri çiftlerinin eğimlerinin ortancası iki şekilde elde edilir;

$$\hat{\beta}_1 = \underset{i}{\text{median}}\ \underset{i\neq j}{\text{median}} \frac{y_j - y_i}{x_j - x_i} \quad i = 1,2,\dots,n \tag{3.30}$$

$$\hat{\beta}_0 = \underset{i}{\text{median}}(y_i - \hat{\beta}_1 x_i) \tag{3.31}$$

3.6.2. Robust ölçek tahmincileri

Regresyon analizinde M-tahminlerin hesaplanma sürecinin en önemli adımlarından biri de bir başlangıç ölçek tahmincisi ile artıkları student türü standartlaştırmaktır. Bu işlem yapılırken basit formül yapısı, daha az hesaplama zamanı, sınırlandırılmış etki fonksiyonu (bounded influence function) içinde çok dayanıklı olması ve %50 kırılma noktasına sahip olmasından dolayı genellikle Medyan Mutlak Sapma (*MAD*- Median Absolute Deviation) kullanılır. Aslında MAD tekniğinin iki dezavantajı vardır, bunlardan birincisi *MAD* simetrik dağılımları amaç edinerek geliştirilmiştir ve ikincisi de Gaussian etkinliği

düşüktür (%37) (Rousseeuw ve Croux, 1993). Bu çalışmada yine aynı kırılma noktasına (%50) sahip fakat Gaussian etkinliği daha yüksek olan S_n (%58) ve Q_n (%82) ölçek tahmincileri kullanılarak M-tahmincilerin performanslarındaki değişim incelenmiştir.

3.6.2.1. En küçük mutlak sapma (MAD – median absolute deviations) tahmincisi

Önerilen robust ölçek tahminlerinden birisi olan MAD (Median Absolute Deviations) aşağıdaki gibi hesaplanır;

$$MAD = \frac{Median\{|e_i - median(e_i)|\}}{0{,}6745} \tag{3.32}$$

0,6745 sabiti n büyük olduğunda ve hatalar normal dağıldığında MAD'ı σ nın yansız bir tahmincisi yaptığı için kullanılmaktadır.

3.6.2.2. S_n tahmincisi

MAD'a göre daha yüksek etkinliğe sahip olan S_n yalnızca simetrik dağılımlar için değil aynı zamanda simetrik olmayan dağılımlar için de yüksek etkinliğe sahiptir. S_n tahmincisi aşağıdaki gibi hesaplanır;

$$S_n = c_0 c_1 \, med_i\{med_j|x_i - x_j|\} \quad i,j = 1,2,\dots,n \tag{3.33}$$

Burada $c_1 = 1{,}1926$ alınırsa ve düzeltme faktörü c_0 kullanıldığında sonlu örneklemde S_n tahmincisi yansız olacaktır (Rousseeuw and Croux 1993). c_0 değerleri çizelge 3.3 ve Çizelge 3.4'de verilmiştir.

Çizelge 3.3. S_n için $n < 9$ için c_0 değerleri

n	2	3	4	5	6	7	8	9
c_0	0,743	1,851	0,954	1,351	0,993	1,198	1,005	1,131

Çizelge 3.4. S_n için $n > 9$ için c_0 değerleri

n tek ise	n çift ise
$c_0 = \frac{n}{n - 0{,}9}$	$c_0 = 1$

MAD gibi S_n de medyan ve mutlak değerlerin kombinasyonu ile hesaplanması kolay bir tahmincidir. Burada mutlak değer yerine karelerini alıp sonunda da bu değerin karekökünü kullanırsa aynı sonucu verecektir (Rousseeuw and Croux 1993).

3.5.8.2.3. Q_n tahmincisi

MAD ve S_n nin etki fonksiyonları sürekli değildir, fakat Q_n fonksiyonu gaussian dağılımı için süreklidir.

$$Q_n = d_0 d_1 \{|x_i - x_j|; i < j\}_{(k)} \tag{3.34}$$

$$k = \binom{h}{2} \approx \binom{h}{2}/4 \tag{3.35}$$

$$h = [n/2] + 1 \tag{3.36}$$

$$d_1 = 2{,}2219$$

d_0 değerleri Çizelge 3.5 ve 3.6'da verilmiştir.

Çizelge 3.5. Q_n için $n < 9$ için d_0 değerleri

n	2	3	4	5	6	7	8	9
d_0	0,339	0,994	0,512	0,844	0,611	0,857	0,669	0,872

Çizelge 3.6. Q_n için $n > 9$ için d_0 değerleri

n tek ise	n çift ise
$d_0 = \frac{n}{n + 1{,}4}$	$d_0 = \frac{n}{n + 3{,}8}$

Q_n tanımından dolayı simetrik olmayan dağılımlar için de uygundur ve yüksek etkinliğe sahiptir. %50 kırılma noktasına sahip olan Q_n gaussian dağılımında %82 etkinliğe ulaşmaktadır. Fakat bu avantajlarına rağmen hesaplanmasında algoritmalar nedeniyle zorluklar da vardır (Rousseeuw and Croux 1993).

4. M-TAHMİNCİLERİN PERFORMANLARININ İNCELENMESİ

Bu bölümde robust regresyon tekniklerinden M-tahminciler için çeşitli analizler yapılarak, farklı açılardan performansları incelenmiştir. İlk olarak M-tahmincilerin başlangıç tahmincilerine olan duyarlılığı uygulamalı olarak gösterilmiştir. Duyarlılığın bu etkisi hem regresyon katsayıları için hem de IRWLS algoritmasının iterasyon sayıları için incelenmiştir. Bununla bağlantılı olarak daha sonra M-tahmincilerin kırılma noktaları (breakdown point) iki farklı durum üzerinde incelenmiştir. Birinci durumda aralarındaki korelasyonu sıfır olan xy çiftleri kullanılmış, ikinci durumda ise yalnızca y aykırı değer kullanılarak kırılma noktaları grafiklerle gösterilmiştir. Bir sonraki aşamada da hata dağılımları standart normal dağılımdan farklı olduğunda M-tahmincilerin performansları hata kareler ortalamasının karekökü (RMSE) açısından incelenmiştir. Son olarak başlangıç ölçek tahminlerinin normal dağılım altında etkinliği değerlendirilmiştir.

4.1. M-Tahmincilerin Başlangıç Tahmincilerine Olan Duyarlılığı

M-tahmincilerin hesaplanmasında IRLWS algoritması kullanmaktadır. Bilinmektedir ki bu algoritmanın adımları gereği bir başlangıç tahmincisine ihtiyaç duyulmaktadır, ve yine bilinmektedir ki monoton M-tahminciler için başlangıç tahmincileri yalnızca iterasyon sayısını etkilemekte ve final çözümü etkilememektedir. Ancak yeniden azalan (redescending) M-tahminciler için bu başlangıç tahminlerinin seçimi büyük önem taşımaktadır, çünkü amaç fonksiyonunu minimum değer veren çözümler local minimumlarda olabilmektedir. Bu bakımından robust başlangıç değerleri kullanmak önem arz etmektedir. Örneğin OLS tahminleri gibi robust olmayan başlangıç değerler kullanıldığında, çözüm kötü local minimum değerlerde olabilmektedir. Özet olarak, IRLWS algoritması başlangıç tahminlerine karşı robust değildir. (Serneels ve diğerler, Coakley and Thomas P. Hettamansperger, 1993; Filzmoser ve diğerleri 2009). Bu bölümde, bu durum çeşitli analizlerle açıklanmaya çalışılacaktır.

4.1.1. Başlangıç tahmincilerinin m-tahmincilerin katsayı sonuçlarına etkisi

Başlangıç tahmincileri ilk artığın elde edilmesi sürecinde gereklidir. Bu nedenle de çok önemli bir aşamadır. Bu durumu uygulamada karşılaşılan iki aykırı değer tipi üzerinden açıklamak başlangıç katsayılarına etkisinin daha iyi anlaşılmasını sağlayacaktır. İkinci bölümde artık aykırı değer ve regresyon aykırı değer kavramlarından söz edildi. Bu iki aykırı değer kısaca tekrar tanımlanırsa, regresyon aykırı değer bir noktanın diğer noktaların genel istikametinden farklı bir yönde yer alması, artık aykırı değer ise bir noktanın artığının diğerlerine göre oldukça büyük çıkmasıdır. Fakat bazı özel durumlarda regresyon artık değer olan nokta o kadar etkileyicidir ki regresyon doğrusunu kendi yönüne doğru çeker ve artık değerini küçültür böylece de artık aykırı değer olmaz. Bu duruma Şekil 3.2 örnek olarak verilmiştir. M-tahminciler de artıkların büyüklüklerine göre ters oranda bir ağırlık verdiklerinden artığı küçük çıkan aykırı değer işlem sürecinde aykırı değer olarak işlem görmeyeceğinden regresyon doğrusu üzerindeki etkinliğini sürdürecektir.

Başlangıç tahmincilerinin M-tahminciler üzerindeki etkisini araştırmak için literatürde kabul görmüş iki farklı deney düzeni kurulmuş ve bu etki incelenmiştir. İlk deney düzeninde, veri hem x hem y aykırı değerler, ikinci deney düzeninde ise yalnızca y aykırı değerler içermektedir.

4.1.1.1. $x-y$ yönlü aykırı değer bulunması durumunda başlangıç tahmincilerinin m-tahmincilerin katsayı sonuçlarına etkisi

Bu deney düzeni, Wu (1985) tarafından Bell tekniği için araştırılmıştır. Bu bölümde de benzer şekilde bir deney düzeni kurularak başlangıç tahmincilerinin etkileri araştırılacaktır. Burada tasarlanan deneyde aykırı değerler o kadar etkilidir ki OLS regresyon doğrusunu kendi yönlerine çekmeyi başarmış bu sayede de artık değerlerini küçültmüşlerdir.

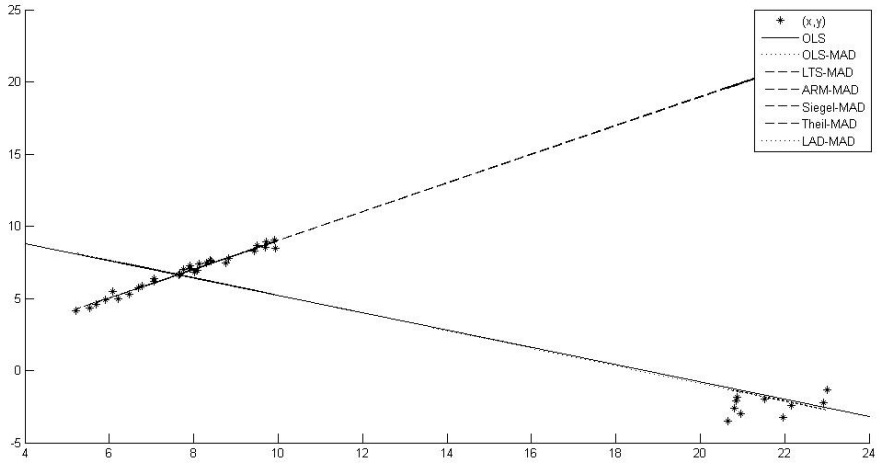
Bu duruma özel olarak, $x \sim U(5,10)$ ve $\varepsilon \sim N(0,0.2)$ dağılımına sahip 30 adet veri türetilmiş $y = -1 + x + \varepsilon$ hesaplanmıştır. Daha sonra 10 adet $(x,y) \sim N\left(\begin{bmatrix}22\\-2\end{bmatrix}, \begin{bmatrix}0.6 & 0\\0 & 0.6\end{bmatrix}\right)$ dağılımına sahip veri türetilerek toplam 40 veri için

M-tahminciler hesaplanmıştır. Böylelikle veri %25 oranında x-y yönünde kirletilmiştir.

Başlangıç değerlere olan duyarlılığı göstermek için bu tezde ele alınan tüm M-tahminciler için yukarıda açıklanan deney düzeni kullanılarak analizler yapılmış ancak elde edilen sonuçlar benzer çıktığı için sadece hızlı yeniden azalan (hard redescending) M-tahmincilerden Andrew ve Tukey, yavaş yeniden azalan (soft redescending) M-tahmincilerden Bell ve Welsch ile monoton M-tahmincilerden Huber ve Fair'in sonuçlarına bu bölümde yer verilmiştir.

Şekil 4.1 ve Şekil 4.2 de Hızlı yeniden azalan (hard redescending) M-tahminciler olan Andrew ve Tukey tahmincileri için başlangıç tahmincisi olarak en küçük kareler (OLS), en küçük budanmış kareler (LTS), Andrew'in medyan regresyon tekniği (Andrew's Regression by Medians – ARM), Siegel, Theil ve en küçük mutlak sapmalar (LAD) alındığında M-tahminciler için elde edilen regresyon doğrusunun çizimleri gösterilmektedir. Ayrıca bu bölümde yapılan tüm hesaplamalarda ARM için $p_1 = 0{,}15$ ve $p_2 = 0{,}10$ alınmış, LTS tahminlerini elde etmek için ise ve h=0,75 alınarak LIBRA kütüphanesi kullanılmıştır.

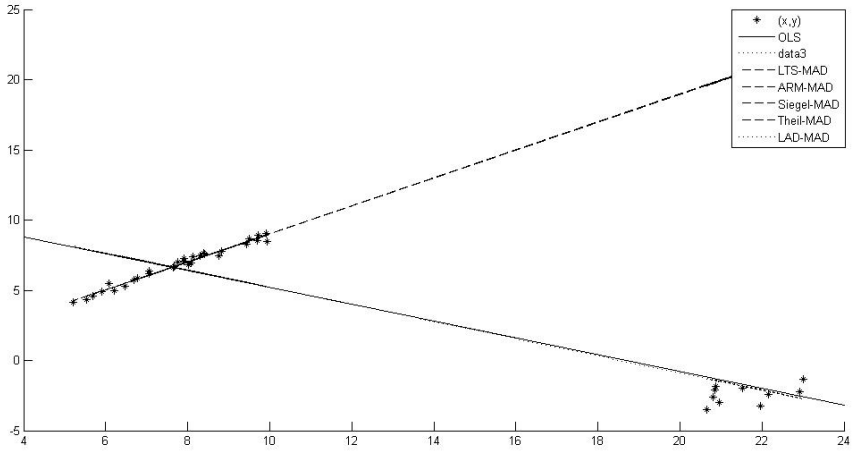
Grafiklerde aynı doğruyu veren başlangıç tahminleri aynı çizgi türü ile gösterilmiştir (…… veya ------ şeklinde).



Şekil 4.1. Andrew'in sine M-tahmincisi için başlangıç tahmincilerinin, veride x-y aykırı değer olması durumunda regresyon katsayılarına etkisi grafiği

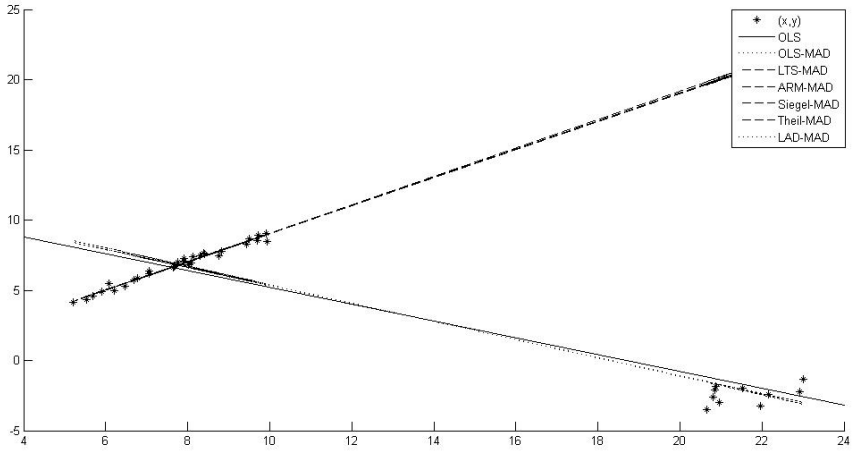
Şekil 4.1'e bakıldığında başlangıç tahmincileri olarak OLS ve LAD alındığında hızlı yeniden azalan (hard redescending) M-tahmincisi olan Andrew

tahmincisinin x aykırı değere doğru eğildiği görülmektedir. Ancak başlangıç tahmincileri olarak LTS, ARM, Siegel ve Theil alındığıda x aykırı değerden etkilenmeden bir tahminleme yapıldığını görülmektedir. Bu sonuç, literatürde sürekli vurgulanan başlangıç tahminlerin yüksek kırılma noktalı tahminler olması gerekliliğini göstermektedir. (Coakley and Hettamansperger, 1993)

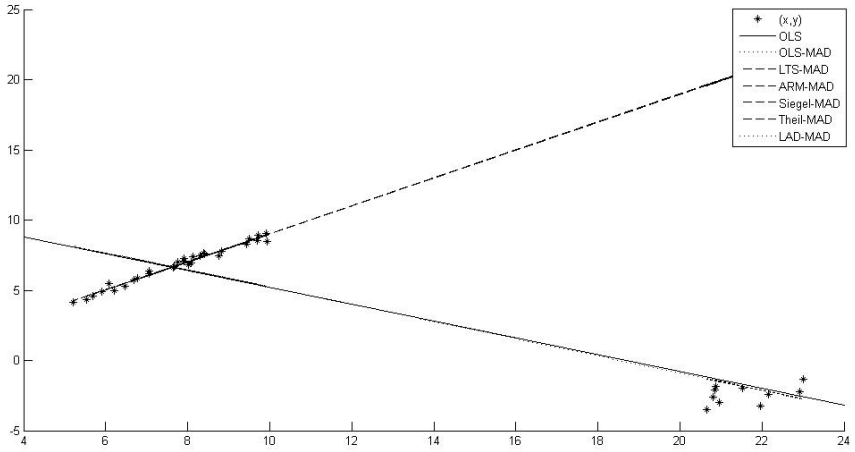


Şekil 4.2. Tukey'in Bi-square M-tahmincisi için başlangıç tahmincilerinin, veride x-y aykırı değer olması durumunda regresyon katsayılarına etkisi grafiği

Bir önceki paragraf da yapılan yorumlar Tukey'in M-tahmincisi içinde geçerlidir. Yüksek kırılma noktasına sahip olan tahmincilerden elde edilen tahminler, başlangıç tahminleri olarak verildiğinde elde edilen regresyon doğrularının x-yönlü aykırı değerlere duyarlı olmadığını göstermektedir. Bu sonuçlar ise IRLWS algoritmasının başlangıç değerin seçimine robust olmadığını göstermektedir.



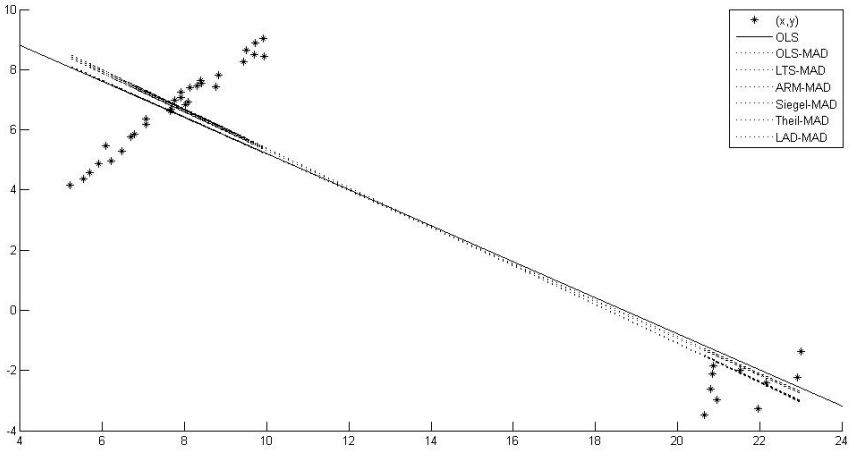
Şekil 4.3. Bell M-tahmincisi için başlangıç tahmincilerinin, veride x-y aykırı değer olması durumunda regresyon katsayılarına etkisi grafiği



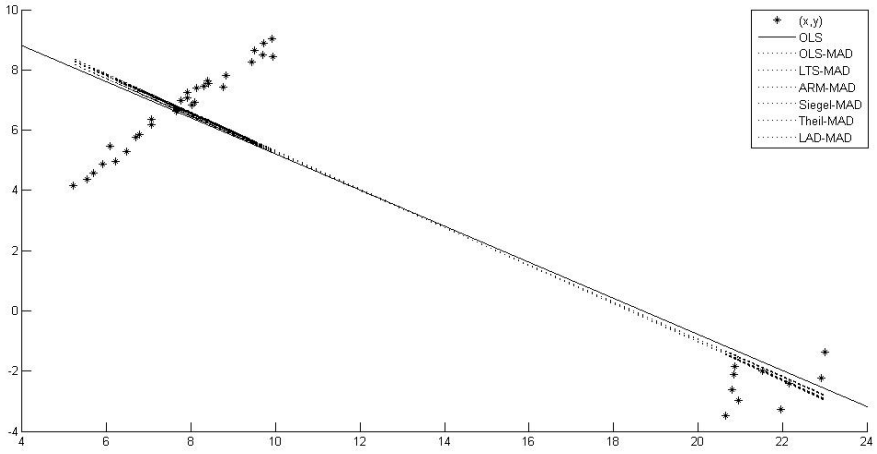
Şekil 4.4. Welsch M-tahmincisi için başlangıç tahmincilerinin, veride x-y aykırı değer olması durumunda regresyon katsayılarına etkisi grafiği

Şekil 4.3 ve Şekil 4.4 incelendiğinde yavaş yeniden azalan (soft redescending) Bell ve Welsch M-tahmincilerinin sonuçlarının hızlı yeniden azalan (hard redescending) M-tahminciler ile aynı çıktığı görülmektedir.

Şekil 4.5 ve Şekil 4.6 ise monoton M-tahmincilerden olan Huber ve Fair'in tahmincileri için başlangıç tahmincileri değiştikçe elde edilen doğruları göstermektedir.



Şekil 4.5. Huber M-tahmincisi için başlangıç tahmincilerinin, veride x-y aykırı değer olması durumunda regresyon katsayılarına etkisi grafiği



Şekil 4.6. Fair tekniği için başlangıç tahmincilerinin xy aykırı değer olması durumunda regresyon katsayılarına etkisi grafiği

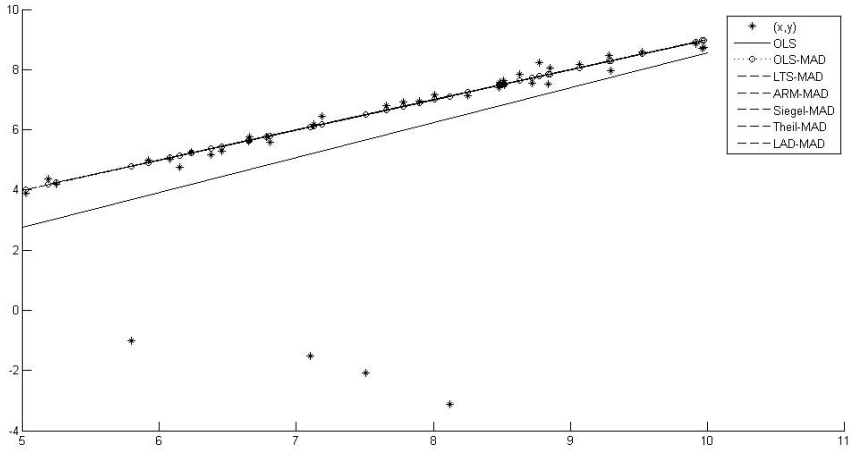
Şekil 4.5 ve Şekil 4.6'dan görüldüğü üzere, monoton ve kesin konveks amaç fonksiyonuna sahip olan bu tahminciler üzerinde başlangıç tahminlerin etkisi olmamakta ve ayrıca x-yönlü aykırı değerlere de oldukça duyarlı oldukları görülmektedir. Hatta bu duyarlılık OLS ile hemen hemen aynıdır.

Yukarıdaki grafikler topluca incelendiğinde başlangıç tahmincilerinin ve aykırı değer türünün etkisi açık bir şekilde görülmektedir. Yeniden azalan olanlar için bu etkiler hem eğim katsayılarının değerini değiştirdiği gibi hem de regresyon katsayılarının yönünü de değiştirmektedir. Öte yandan bu grafiklerde dikkat çekici

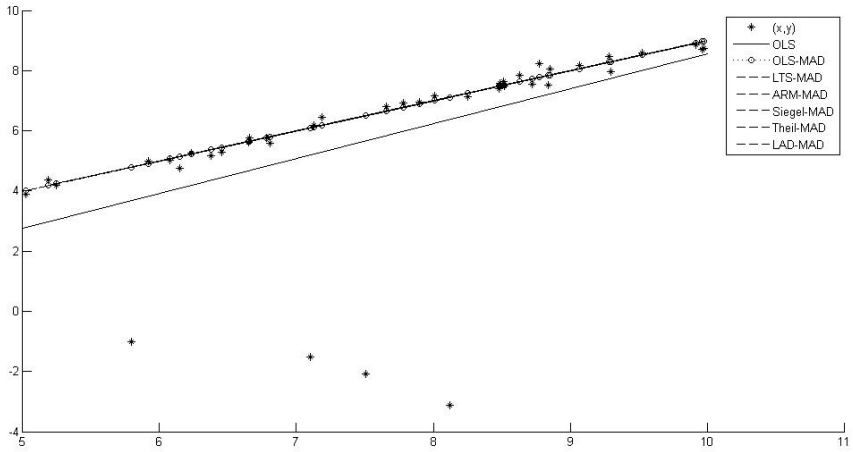
diğer bir önemli nokta LAD tekniğinin robust bir teknik olmasına rağmen, başlangıç değer tahmincisi olarak ele alındığında, Yeniden azalan (redescending) M-tahminciler üzerinde OLS ile genelde aynı etkiye sahip olmasıdır. Yani sonuçlarının OLS ile benzer çıkmasıdır. Bu durumun nedeni LAD tekniğinin x aykırı değerlere karşı olan duyarlılığından kaynaklanmaktadır. Yani LAD'nin x'e karşı kırılma noktası OLS ile aynı olup 0'dır.

4.1.1.2. y yönlü aykırı değer bulunması durumunda başlangıç tahmincilerinin m-tahmincilerin katsayı sonuçlarına etkisi

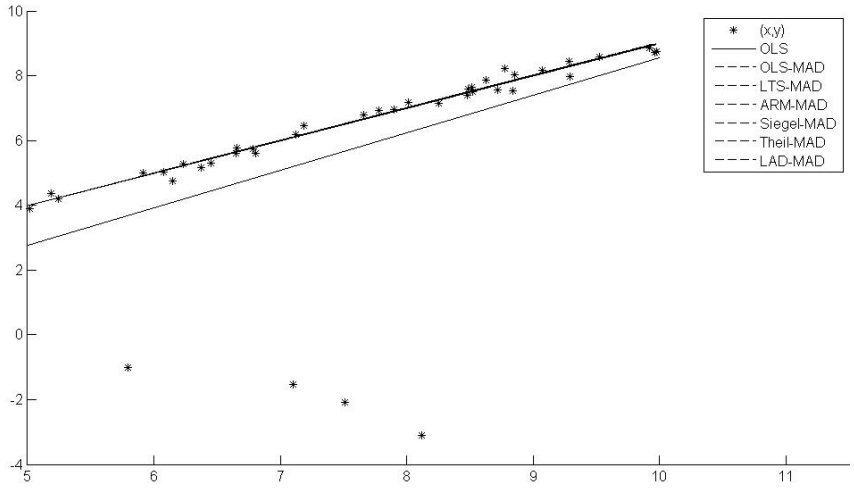
M-tahmincilerin yalnızca y aykırı değer bulunması durumunda sonuçlarını araştırmak için bölüm 4.1.1.1 deki deney düzeninde biraz değişiklik yaparak x aykırı değer kısmı atılmıştır. Böylelikle yeni deney düzenine özel olarak, $x \sim U(5,10)$ ve $\varepsilon \sim N(0,0.2)$ dağılımına sahip 36 adet veri türetilmiş $y = -1 + x + \varepsilon$ hesaplanmıştır. Daha sonra $(x, y) \sim N\left(\begin{bmatrix} 7.5 \\ -2 \end{bmatrix}, \begin{bmatrix} 0.6 & 0 \\ 0 & 0.6 \end{bmatrix}\right)$ dağılımına sahip 4 adet veri türetilerek M-tahminciler hesaplanmıştır. Böylelikle veri %10 y-yönünde kirletilmiştir. Bu veri için bu tezde ele alınan tüm M-tahminciler uygulanmıştır. Sonuçların benzer çıkmasından dolayı bir önceki bölümde olduğu gibi, hızlı yeniden azalan (hard redescending) M-tahmincilerden Andrew ve Tukey'in (Şekil 4.7 - Şekil 4.8), yavaş yeniden azalan (soft redescending) M-tahmincilerden Bell ve Welsch (Şekil 4.9 – Şekil 4.10) monoton M-tahmincilerden Huber ve Fair'in (Şekil 4.11 - Şekil 4.12) sonuçları grafiksel olarak gösterilmiştir. Sonuçlar tüm tahminciler için benzerdir. Katsayılar arasında küçük farklılıklar söz konusudur.



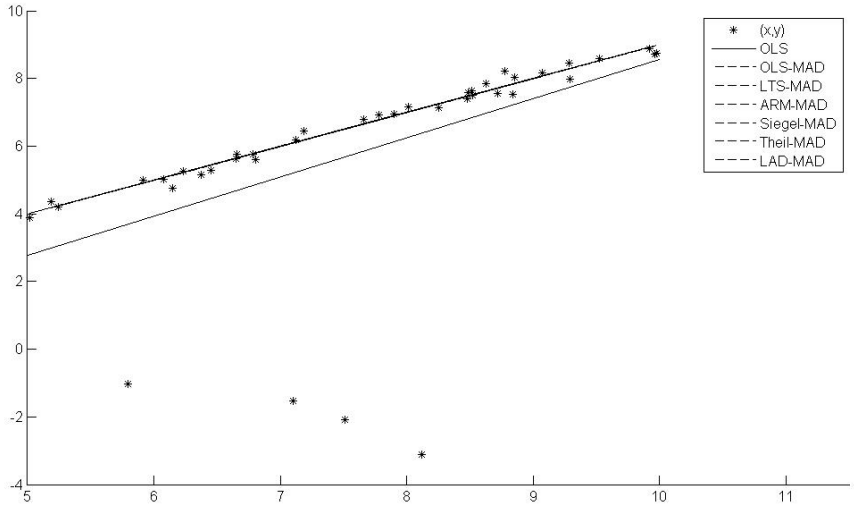
Şekil 4.7. Andrew'in sine M-tahmincisi için başlangıç tahmincilerinin, veride y aykırı değer olması durumunda regresyon katsayılarına etkisi grafiği



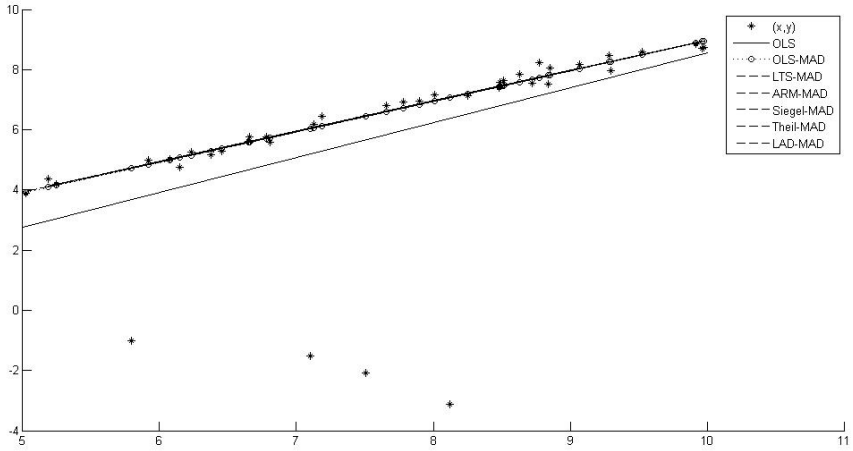
Şekil 4.8. Tukey M-tahmincisi için başlangıç tahmincilerinin, veride y aykırı değer olması durumunda regresyon katsayılarına etkisi grafiği



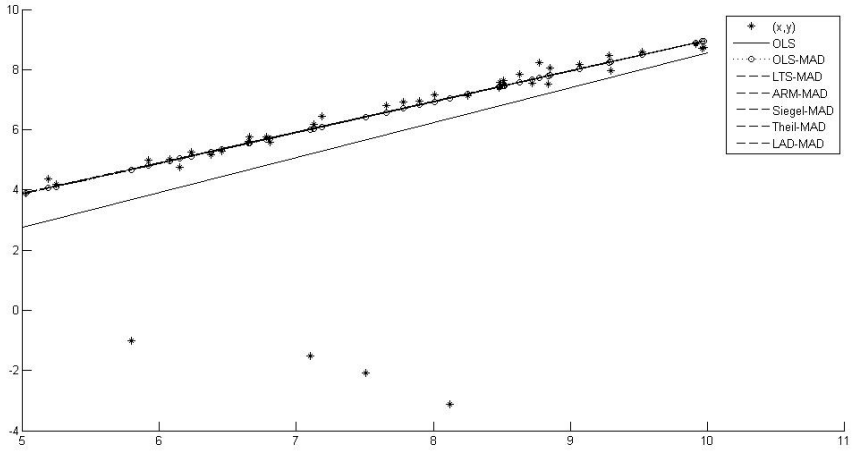
Şekil 4.9. Bell M-tahmincisi için başlangıç tahmincilerinin, veride y aykırı değer olması durumunda regresyon katsayılarına etkisi grafiği



Şekil 4.10. Welsch M-tahmincisi için başlangıç tahmincilerinin, veride y aykırı değer olması durumunda regresyon katsayılarına etkisi grafiği



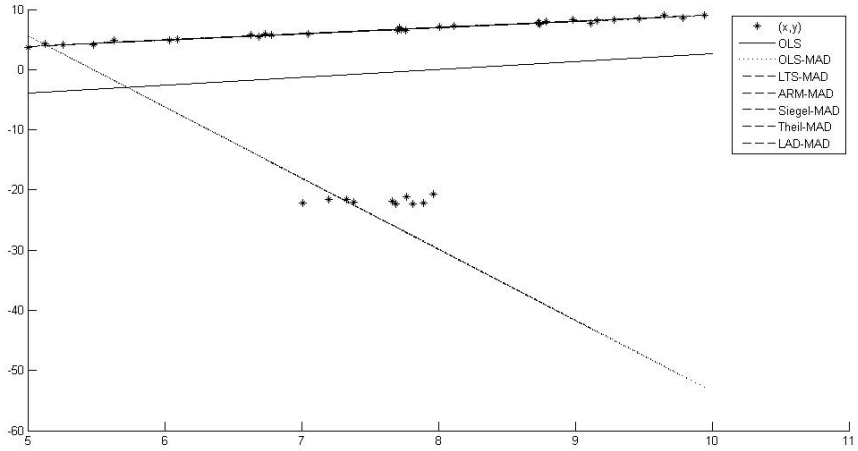
Şekil 4.11. Huber M-tahmincisi için başlangıç tahmincilerinin, veride y aykırı değer olması durumunda regresyon katsayılarına etkisi grafiği



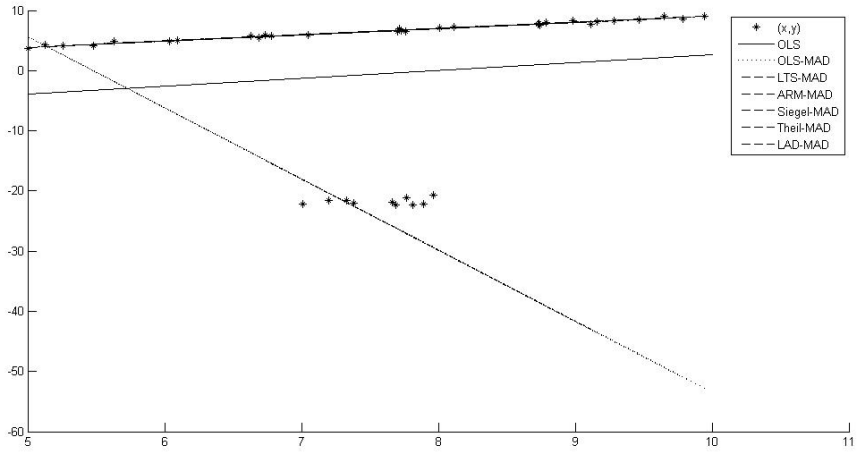
Şekil 4.12. Fair M-tahmincisi için başlangıç tahmincilerinin, veride y aykırı değer olması durumunda regresyon katsayılarına etkisi grafiği

Şekil 4.7'den şekil 4.12'e kadar grafikler topluca incelendiğinde %10 y-yönlü aykırı değer olması durumunda tüm m-tahmincilerinin başlangıç değere bağlı olmaksızın aynı sonucu verdiği görülmektedir.

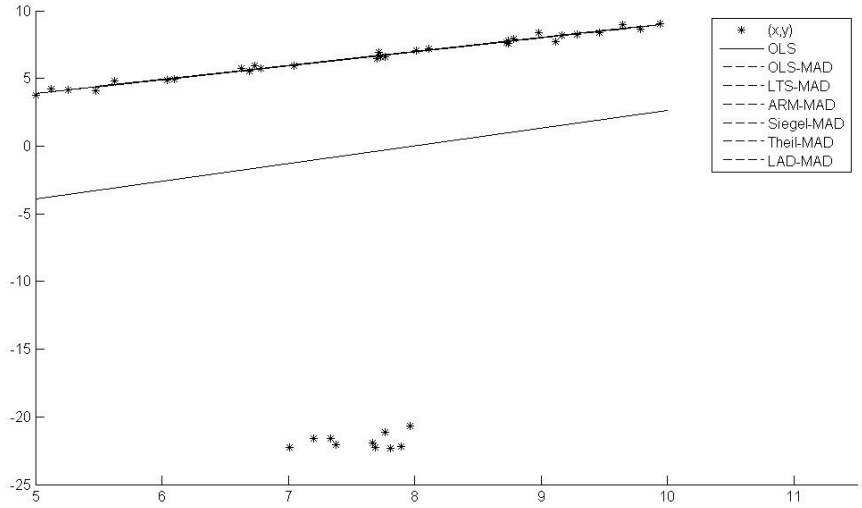
Aynı deney düzeninde temiz veri %25 $(x,y) \sim N\left(\begin{bmatrix} 7.5 \\ -22 \end{bmatrix}, \begin{bmatrix} 0.2 & 0 \\ 0 & 0.2 \end{bmatrix}\right)$ şeklinde y- yönünde kirletildiğinde ise elde edilen grafikler aşağıda verilmiştir. Burada dikkat edilmesi gereken y-yönündeki aykırı değerlerin verinin genelinden daha uzak konumda olmasıdır.



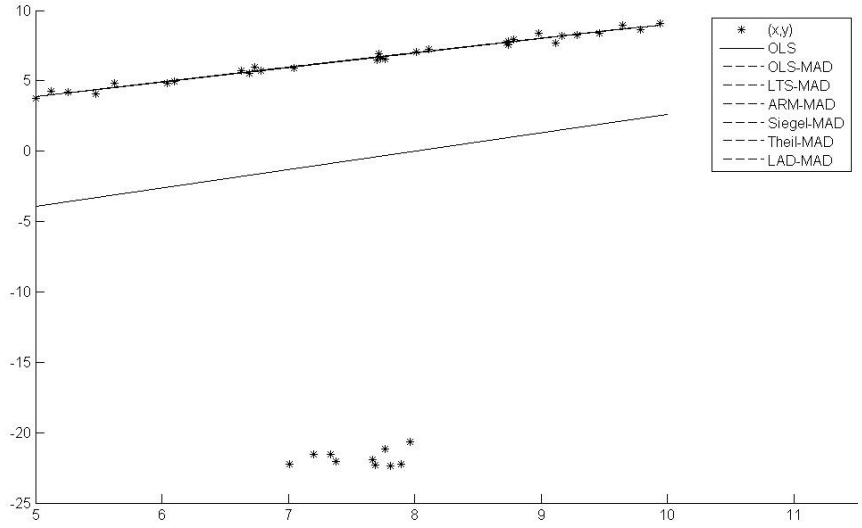
Şekil 4.13. Andrew'in sine M-tahmincisi için başlangıç tahmincilerinin, veride uzak y aykırı değer olması durumunda regresyon katsayılarına etkisi grafiği



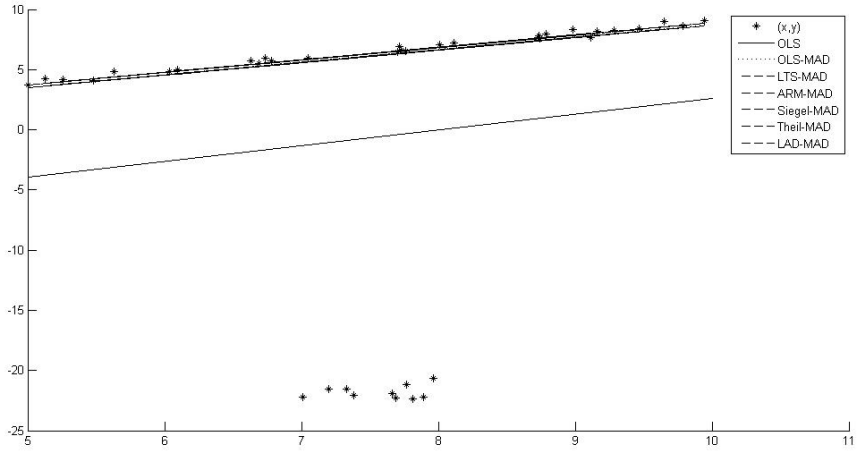
Şekil 4.14. Tukey M-tahmincisi için başlangıç tahmincilerinin, veride uzak y aykırı değer olması durumunda regresyon katsayılarına etkisi grafiği



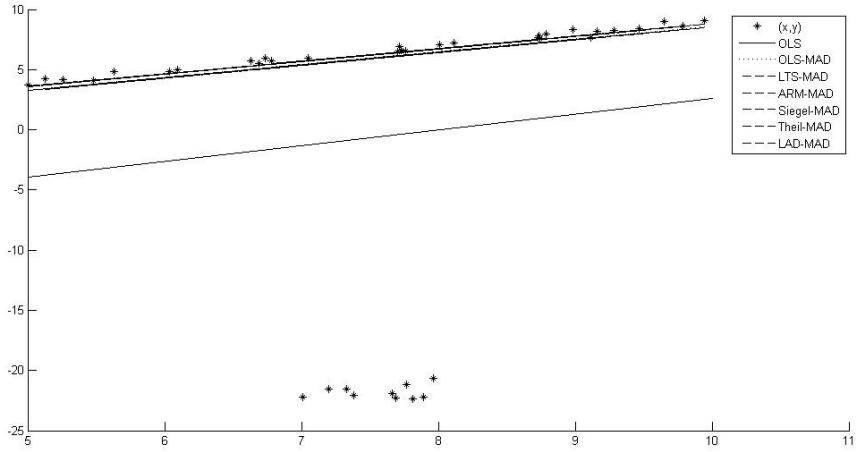
Şekil 4.15. Bell M-tahmincisi için başlangıç tahmincilerinin, veride uzak y aykırı değer olması durumunda regresyon katsayılarına etkisi grafiği



Şekil 4.16. Welsch tekniği için başlangıç tahmincilerinin uzak y aykırı değer olması durumunda regresyon katsayılarına etkisi grafiği



Şekil 4.17. Huber M-tahmincisi için başlangıç tahmincilerinin, veride uzak y aykırı değer olması durumunda regresyon katsayılarına etkisi grafiği



Şekil 4.18. Fair M-tahmincisi için başlangıç tahmincilerinin, veride uzak y aykırı değer olması durumunda regresyon katsayılarına etkisi grafiği

Şekil 4.13-Şekil 4.14 hızlı yeniden azalan M-tahmincilerin sonuçlarıdır. Bu grafiklerden görülmektedir ki başlangıç tahminleri olarak OLS alındığında hızlı yeniden azalan M-tahmincilerin sonuçları robust olmamaktadır. Bu ise IRWLS algoritmasının başlangıç tahmincilerine karşı robust olmadığının bir göstergesidir. Öte yandan yavaş yeniden azalan (soft redescending) ve monoton tahmincilerin çözümleri başlangıç değerden bağımsız olarak tektir.

Tüm grafikler incelendiğinde, doğal olarak OLS tahminlerinin x-y ve y – yönlü aykırı değerlere doğru eğildiği yani büyük ölçüde etkilendiği görülmektedir.

y- aykırı değer durumunda monoton M-tahmincilerin sonuçları robust ve başlangıç değerin seçiminden bağımsız olduğu için daha güvenilirdir yorumu yapılabilir. y- aykırı değer durumunda yeniden azalan M-tahmincilerin sonuçları da monoton M-tahmincilerle pareler çıkmıştır.

x-y aykırı değer olduğunda başlangıç tahminine bağlı olarak yeniden azalan M-tahminciler monoton tahmincilere göre aykırı değerlere karşı daha dirençlidir yorumu yapılabilir.

Öte yandan bu sonuçların aykırı değer türüne ve oranına bağlı olarak değişebileceği vurgulanmalıdır. Çünkü M-tahmincilerin kırılma noktası y- aykırı değere karşı %30'u geçmemektedir.

4.1.2. Başlangıç tahmincilerinin IRWLS algoritmasında iterasyon sayılarına etkisi

Bir tahmincide bulunması istenilen özelliklerden biriside kolay hesaplanması ve kısa işlem sürecine sahip olmasıdır. M-tahminciler hesaplama olarak kolay hesaplanabilir tahmincilerdir fakat hesaplamada kullanılan IRLWS tekniğinin iterasyon sayıları köke yakınsamak için bazı durumlarda artabilmektedir. Bu ise uygulamada istenilmeyen bir durumdur. M-tahmincilerin diğer robust tahmincilere göre en önemli avantajı olan kolay hesaplanabilirliği, iterasyon sayısının artması durumunda işlem sürecinin uzamasının gölgesinde kalmakta ve yerini diğer robust tahmincilere bırakmaktadır. Bu süreci kısaltabilmek için başlangıç tahminlerini uygun seçmek gerekmektedir. Bu bölümde IRWLS algoritmasının başlangıç tahminlerinin iterasyon sayısına etkisi araştırılmaktadır. Bu amaçla, daha önceki bölümde olduğu gibi normal dağılmış veri belli oranda kirletilmiştir ve bir simülasyon modeli üzerinde iterasyon sayıları incelenmiştir. Simülasyon modelinde $n = 50$ olacak şekilde $x \sim N(0,1)$ ve $\varepsilon \sim N(0,1)$ veri türetilmiş, türetilen bu veri %10 oranında $\varepsilon \sim N(0,10)$ veri ile

kirletilmiştir. Daha sonra $y = -1 + x + \varepsilon$ ile y değerleri elde edilmiştir. 1000 yineleme yapılmış ve ortalama iterasyon sayısı elde edilmiştir. Sonuçlar Çizelge 4.1 de sunulmaktadır.

Çizelge 4.1. Başlangıç tahmincilerinin iterasyon sayılarına etkisi

	Başlangıç Tahmincileri					
M-tahminciler	**OLS**	**LTS**	**ARM**	**Siegel**	**Theil**	**LAD**
Andrew	5	3	4	4	4	4
Bell	8	6	7	6	6	6
Cauchy	5	4	5	4	4	4
Danish	3	2	3	3	3	3
Fair	5	4	5	5	4	5
Geman and Mcclure	13	10	12	10	10	9
Hampel	4	3	3	3	3	3
Huber	4	3	4	3	3	4
Logistic	5	4	5	4	4	4
QSR	5	4	5	4	4	4
Ramsay	5	4	5	4	4	4
Talwar	3	2	2	2	2	2
Welsh	5	3	4	4	4	4
Tukey	5	3	4	4	4	4
Asad	4	3	3	3	3	3
insha	4	3	3	3	3	3
Qadir	5	3	4	4	4	4
$\psi 1$	3	3	3	3	3	3
$\psi 2$	3	3	3	3	3	3

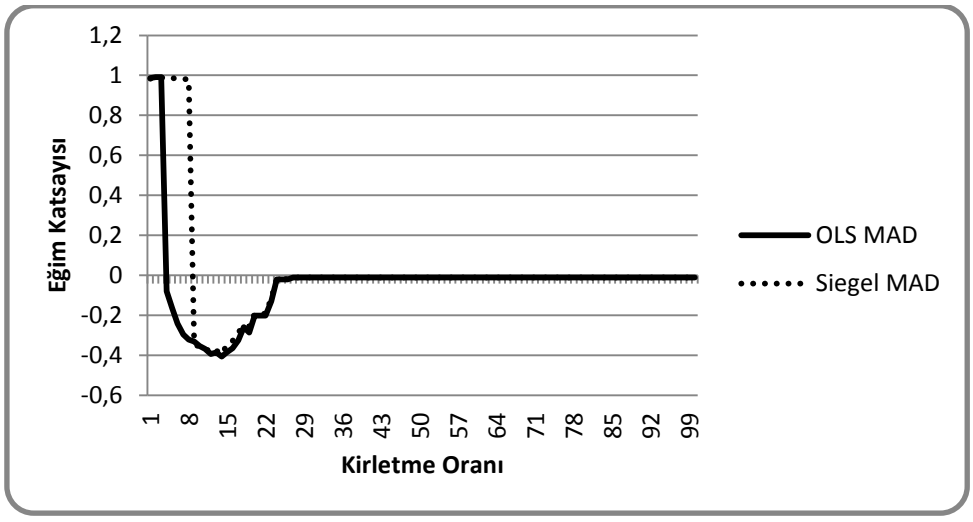
Çizelge 4.1'den görülmektedir ki başlangıç tahminleri iterasyon sayılarını genelde etkilemektedir. İlk göze çarpan sonuç, OLS başlangıç tahmincisi olarak alındığında, M-tahmincilerin tahminlerine ulaşırken algoritmanın iterasyon sayısında artma olduğudur. Başlangıç değerin iterasyon sayısına etkisi en fazla Geman and Mcclure tahmincisinde olmuştur. Daha önceki bölüm ile birlikte ele alındığında başlangıç tahmincilerinin iki tür etkisi söz konusudur. Bunlardan ilki yeniden azalanlarda (redesceding) local minimuma yakınsama, ikincisi de tüm tahmincilerde iterasyon sayısında artma veya azalma şeklindedir.

4.2. M-Tahmincilerin Kırılma Noktaları

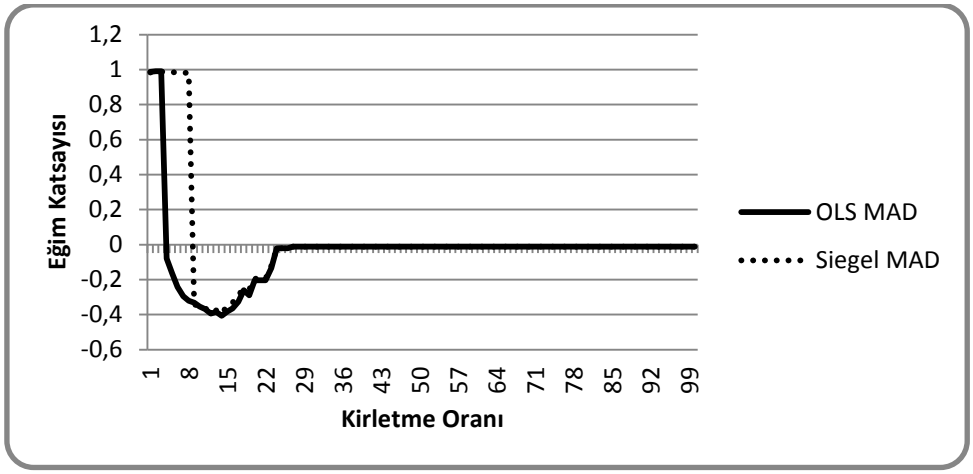
M-tahmincilerin kırılma noktaları aykırı değerlerin türlerine göre değişiklik göstermektedirler. Bu durumu göstermek üzere iki farklı deney düzeni dizayn edilmiş ve sonuçlar grafikler halinde gösterilmiştir.

4.2.1. $x - y$ yönlü aykırı değerler için m-tahmincilerin kırılma noktaları

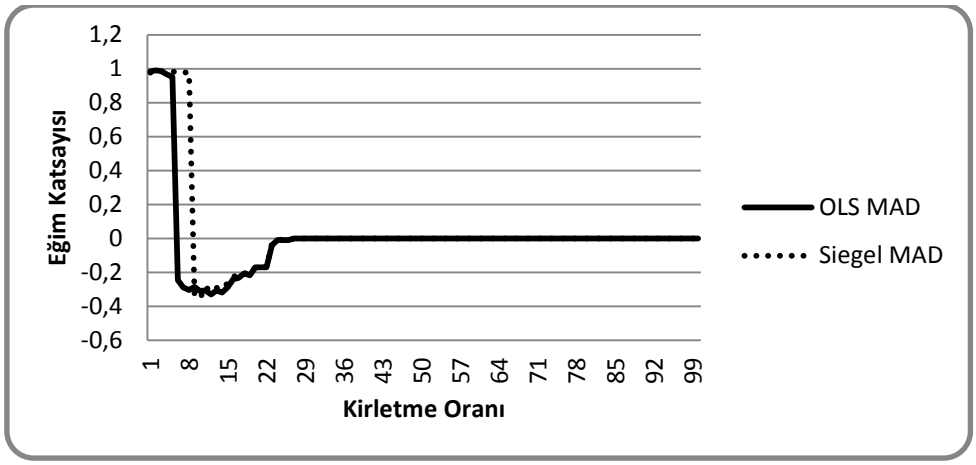
Bu grafikler için Boyer (2003)'ün çalışmasına benzer yönde $n = 100$ olacak şekilde $x \sim U(1,4)$ ve $\varepsilon \sim N(0,0.2)$ dağılımına sahip veri türetilmiş ve bağımsız değişken değerleri $y = 2 + x + \varepsilon$ ile elde edilmiştir. Daha sonra $(x, y) \sim N\left(\begin{pmatrix}7\\2\end{pmatrix} \quad \begin{pmatrix}0.5 & 0\\0 & 0.5\end{pmatrix}\right)$ dağılımlarına sahip 100 veri türetilerek %1 den başlayarak %100'e kadar (x, y) çiftleri yerine (x, y) çiftleri kullanılarak veri seti kirletilmiştir.



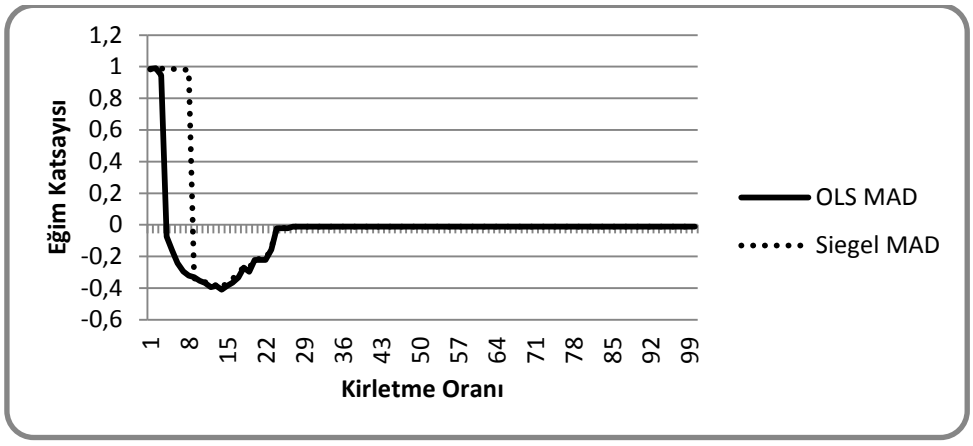
Şekil 4.19. Andrew'in sine M-tahmincisinin, veride x-y aykırı değer olması durumunda kırılma noktası grafiği



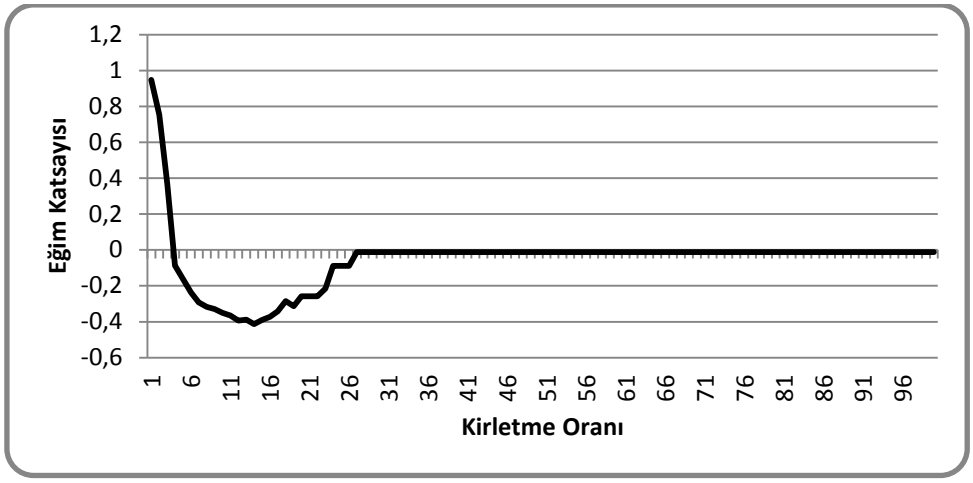
Şekil 4.20. Tukey M-tahmincisinin, veride x-y aykırı değer olması durumunda kırılma noktası grafiği



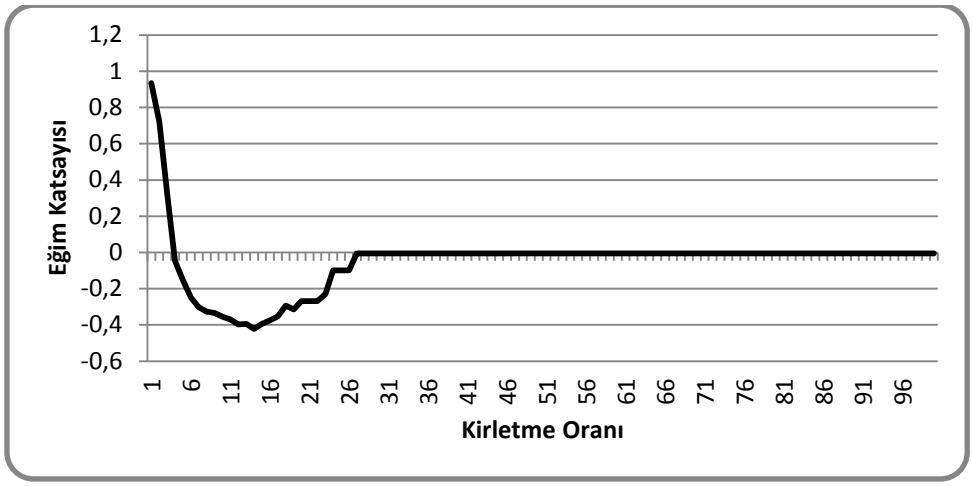
Şekil 4.21. Bell M-tahmincisinin, veride x-y aykırı değer olması durumunda kırılma noktası grafiği



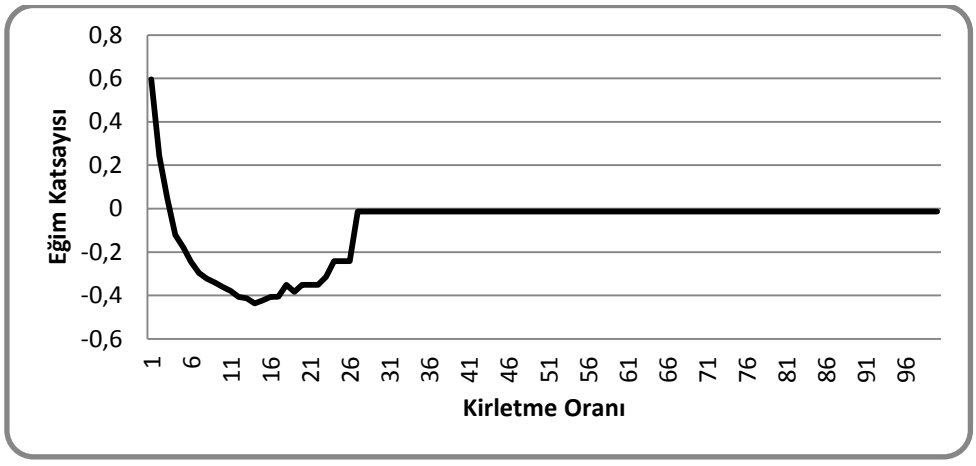
Şekil 4.22. Welsch M-tahmincisinin, veride x-y aykırı değer olması durumunda kırılma noktası grafiği



Şekil 4.23. Huber M-tahmincisinin, veride x-y aykırı değer olması durumunda kırılma noktası grafiği



Şekil 4.24. Fair M-tahmincisinin, veride x-y aykırı değer olması durumunda kırılma noktası grafiği



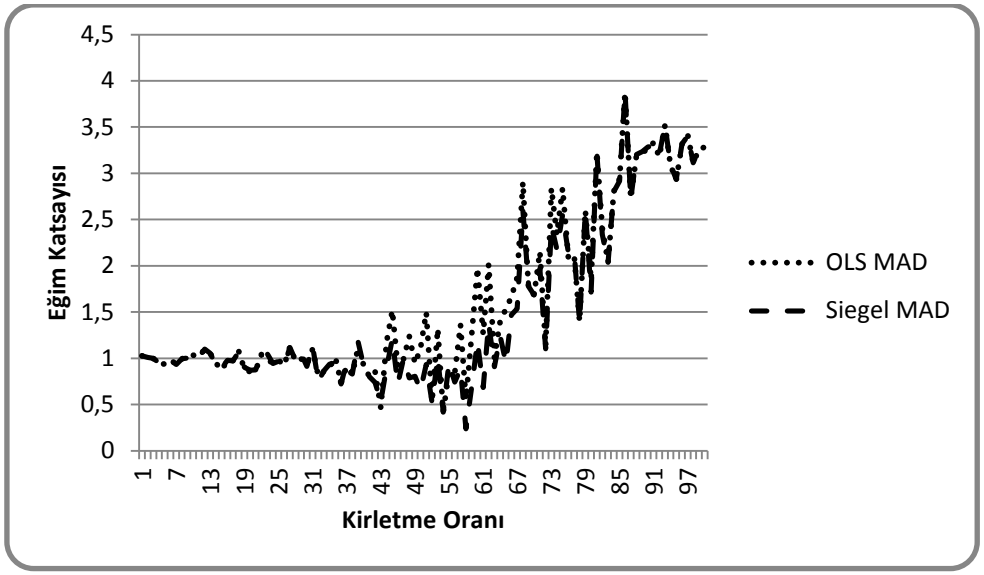
Şekil 4.25. OLS tahmincisinin, veride x-y aykırı değer olması durumunda kırılma noktası grafiği

Şekil 4.19 - Şekil 4.20'de hızlı yeniden azalan (hard redescending) M-tahmincilerden olan Tukey ve Andrew ve Şekil 4.21 - Şekil 4.22'de yavaş yeniden azalan (soft redescending) M-tahmincilerden olan Bell ve Welsch için kırılma noktası grafikleri verilmiştir. Grafiklerde görülen düz çizgi OLS ve MAD sırsıyla regrsyon ve ölçek başlangıç tahmini olarak alınırken kesikli çizgi Siegel ve MAD yine sırasıyla regrsyon ve ölçek başlangıç tahmini olarak alınmıştır. Bu grafiklerden hızlı yeniden azalan M-tahminciler yaklaşık %7 oranında x- aykırı değere karşı dirençli oldukları görülmektedir. Öte yandan Şekil 4.23 ve Şekil 4.24 den görülmektedir ki monoton M-tahmincilerin x-yönlü aykırı değerlere %1'lik bir oranında dahi duyarlılığı söz konusudur. Bu sonuçlar, tezde sunulan tüm monoton M-tahminciler için geçerlidir. Ayrıca Şekil 4.19, Şekil 4.20, Şekil 4.21 ve Şekil 4.22 den bir önceki bölüme paralel olarak, yeniden azalan M-tahminciler için OLS başlangıç tahmincisi olarak alındığında M-tahmincilerin kırılma noktası grafikleri OLS ile benzer sonuçlar vermektedir. Böylelikle yeniden azalan M-tahmincilerin başlangıç tahmincilerine duyarlılığı bir kez daha ortaya çıkmaktadır. Fakat robust başlangıç tahmincileri kullanıldığında M-tahmincilerin kırılma noktaları %10 a kadar çıkmaktadır. Bu sonuçlar tezde sunulan Danish, Geman and Mcclure, Ramsay, Talwar, Smith, Asad, insha, Qadir, $\psi 1$ ve $\psi 2$ için de geçerlidir.

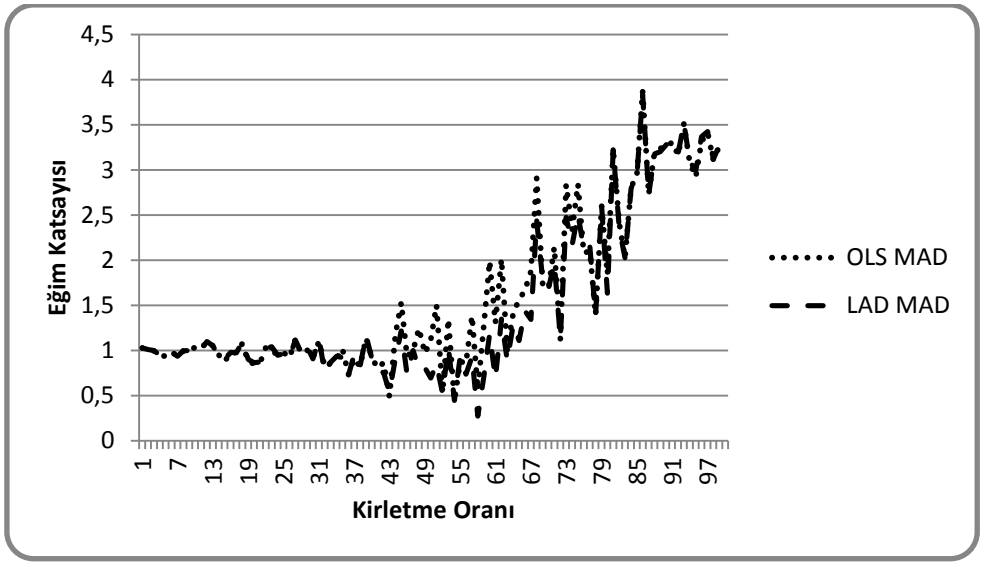
4.2.2. y-yönlü aykırı değerler için m-tahmincilerin kırılma noktaları

M-tahmincilerin x aykırı değerlere karşı duyarlı y aykırı değerlere karşı robust olduğu bilindiğinden bu durum için bir önceki bölümden farklı olarak yeni bir deney dizayn edilmiş ve kırılma noktası grafikleri incelenmiştir.

Bu grafikler için $n = 100$ olacak şekilde $x \sim U(1,4)$, $\varepsilon \sim N(0,1)$ ve $\varepsilon \sim N(0,10)$ dağılımlarına sahip veri türetilmiş daha sonra %1 den başlanarak %100'e kadar veri seti kirletilmiştir. Bu deney modelinde bağımlı değişken değerleri $y = -1 + x + \varepsilon$ ile elde edilmiştir. Aşağıda sunulan grafikler tüm M-tahminciler için incelenmiş sonuçların benzerliği dikkate alınarak daha önceki analizlerde olduğu gibi Andrew, Tukey, Bell, Welsch, Huber ve Fair'e ilişkin sonuçlar verilmiştir.

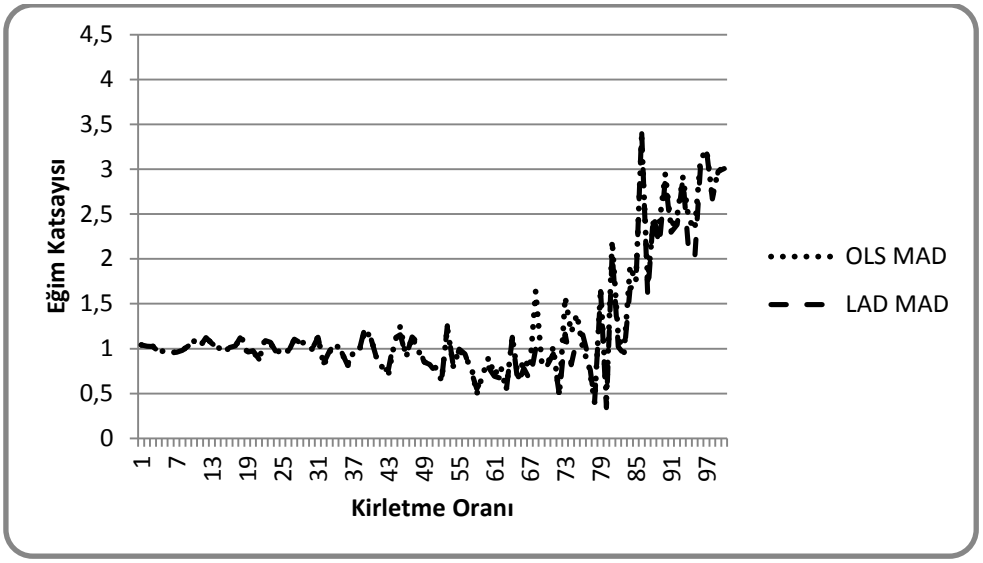


Şekil 4.26. Andrew'in M-tahmincisinin, veride y aykırı değer olması durumunda kırılma noktası grafiği

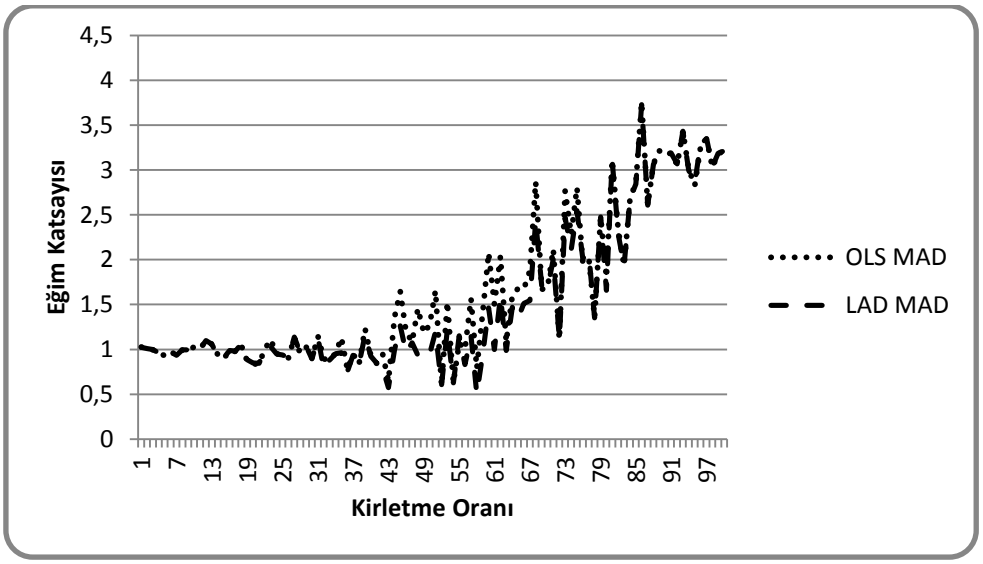


Şekil 4.27. Tukey M-tahmincisinin, veride y aykırı değer olması durumunda kırılma noktası grafiği

Şekil 4.26 – Şekil 4.27'den görüldüğü üzere y-aykırı değer durumunda Andrew ve Tukey'in kırılma noktası %30 a kadar çıkmaktadır. %30-%40 arasında ise değişkenlik bir miktar artmakta fakat %40 dan sonra ise değişkenlik oldukça fazla olmaktadır.

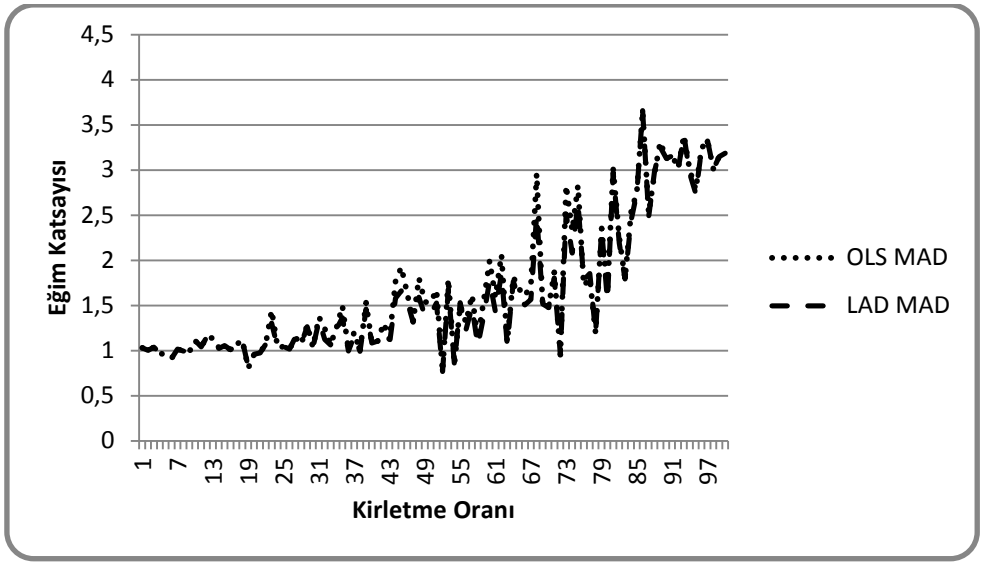


Şekil 4.28. Bell M-tahmincisinin, veride y aykırı değer olması durumunda kırılma noktası grafiği

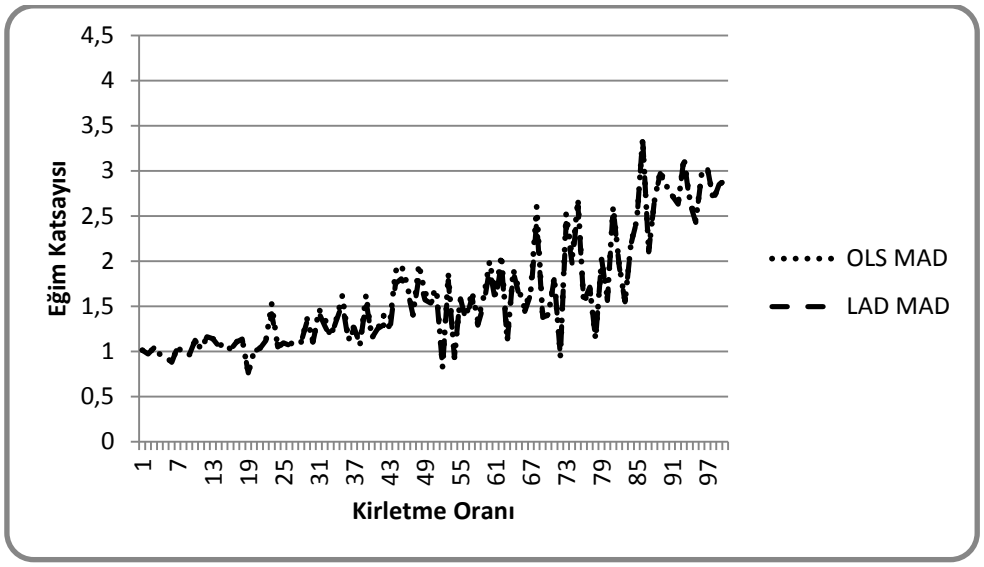


Şekil 4.29. Welsch M-tahmincisinin, veride y aykırı değer olması durumunda kırılma noktası grafiği

Şekil 4.28 – Şekil 4.29'dan görüldüğü üzere y-aykırı durumunda Bell ve Welsch'in tahmincilerinin kırılma noktası grafikleri Andrew ve Tukey'e benzemektedir.

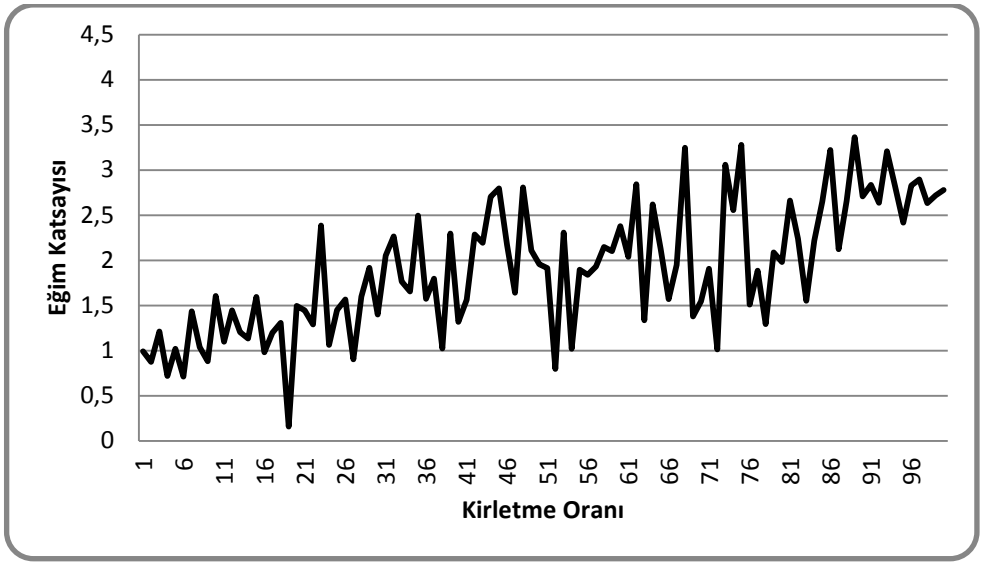


Şekil 4.30. Huber M-tahmincisinin, veride y aykırı değer olması durumunda kırılma noktası grafiği



Şekil 4.31. Fair M-tahmincisinin, veride y aykırı değer olması durumunda kırılma noktası grafiği

Şekil 4.30 – Şekil 4.31'den görüldüğü üzere y-aykırı değer durumunda monoton olan Huber ve Fair'in kırılma noktası yaklaşık %18'e kadar çıkmaktadır. %18'den sonra ise giderek artan bir değişkenlik gözlemlenmektedir.



Şekil 4.32. OLS tahmincisinin, veride y aykırı değer olması durumunda kırılma noktası grafiği

Son olarak OLS tahmincisinin kırılma noktası grafiğine bakıldığında, OLS'nin en küçük oranda bir aykırı değerden bile oldukça etkilendiği görülmektedir.

4.3. Hata Dağılımlarına Göre M-Tahmincilerin Performansı

Genellikle tahminciler, varyans, yanlılık (bias) veya hata kareler ortalamasının karekökü (RMSE) ölçütlerine göre değerlendirilmektedir. Bilinmektedir ki OLS tahmincileri modelin hata dağılımı normal olduğunda düzgün en küçük varyanslı yansız tahmin edici yani UMVUE'dir. Bu özellik katı bir şekilde normallik varsayımına dayanmaktadır. Fakat uygulamada birçok durumda hataların dağılımının normal dağılıma uymadığı görülür. Bu bölümde hata teriminin dağılımı normal dağılım olduğu durum başta olmak üzere, hata teriminin dağılımı normalden farklı olduğu durumlarda, ele alınan M-tahmincilerin performansı hata kareler ortalamasının karekökü (RMSE) ölçütüne göre değerlendirilecektir.

Hata dağılımı olarak simetrik dağılımlardan normal dağılım, Laplace dağılımı, kirletme oranları farklı iki scale-contaminated normal dağılım, 3 serbestlik dereceli Student- t dağılımı ele alınmıştır. Yukarıda ifade edilen

dağılımlardan Laplace dağılımı, scale-contaminated normal dağılımlar, 3 serbestli dereceli Student- t dağılımı kalın kuyruklu ve aykırı değerleri üretmede kullanılan dağılımlardır. Asimetrik dağılım olarak da lognormal dağılımı kullanılmıştır. Bilinmektedir ki asimetrik dağılımlar için M-tahminciler yanlı sonuçlar üretmektedir. Ancak yine de ortaya çıkardıkları RMSE değerlendirmesi OLS'ye göre yapılacaktır. Bunlara ilave olarak bu tezde başlangıç tahmincileri olarak kullanılan LTS, LAD, ARM, Theil ve Siegel teknikleri için de RMSE analizi yapılacaktır.

Similasyon dizaynı olarak Hsieh & Manski (1987)'nin çalışmasındaki dizayn kullanılmıştır. (Kantar ve diğerleri 2011; Usta ve Kantar, 2011; McDonald ve White 1993; Ramirez ve diğerleri 2003).

$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad i = 1,...,n$ doğrusal regresyon modelinde $\beta_0 = -1$ ve $\beta_1 = 1$ alınmış elde edilen $y_i = -1 + x_i + \varepsilon_i$ for $i = 1,...,n$ modelinde, $x_i \sim N(0,1)$ dağılımına sahiptir ve yukarıda bahsedilen dağılımlar hata teriminin dağılımı olarak düşünülerek veri üretilmiş ve elde eldilen veriler için bahsedilen tahminciler yardımıyla tahmincilerin RMSE değerleri 1000 tekrar yapılarak hesaplanmıştır. Regresyon eğim parametresi için RMSE formülü aşağıdaki eşitlikte verilmektedir.

$$RMSE\left(\tilde{\beta}_1\right) = \sqrt{\frac{1}{1000}\sum_{i=1}^{1000}\left(\tilde{\beta}_{1i} - 1\right)^2} \qquad (4.1)$$

Normal dağılım için elde edilen sonuçlar Çizelge 4.2 de verilmektedir. Tabi ki OLS normal dağılımda en iyi sonucu vermelidir. Bununla birlikte normal dağılım durumunda M-tahmincilerinin kendi içinde performans analizi yapılmıştır. Ayrıca M-tahmincileri hesaplanırken, ölçeğin robust tahmini olarak MAD, regresyon katsayılarının başlangıç tahminleri olarak LTS'den elde edilen tahminler kullanılmıştır.

Çizelge 4.2. Hata dağılımı normal olduğu durum için tahmincilerin RMSE değerleri

	n=10	n=20	n=30	n=50	n=100	n=250
OLS	0,3687	0,2423	0,1917	0,1411	0,0989	0,0651
Andrew	0,4731	0,2619	0,2076	0,1505	0,1035	0,0665
Bell	0,5321	0,3127	0,2447	0,1799	0,1242	0,0748
Cauchy	0,4131	0,2511	0,1987	0,1478	0,1033	0,0661
Danish	0,4692	0,2669	0,2130	0,1525	0,1038	0,0676
Fair	0,3789	0,2483	0,1965	0,1466	0,1034	0,0661
Geman and Mcclure	0,5926	0,3784	0,2960	0,2243	0,1567	0,0926
Hampel	0,3844	0,2459	0,1958	0,1446	0,1009	0,0657
Huber	0,3914	0,2506	0,1982	0,1477	0,1028	0,0661
Logistic	0,3863	0,2496	0,1976	0,1473	0,1032	0,0661
QSR	0,4416	0,2635	0,2070	0,1552	0,1071	0,0679
Ramsay	0,4042	0,2475	0,1962	0,1459	0,1023	0,0658
Talwar	0,4771	0,2802	0,2184	0,1550	0,1043	0,0682
Welsch	0,4555	0,2564	0,2017	0,1492	0,1034	0,0663
Tukey	0,4708	0,2604	0,2072	0,1504	0,1035	0,0665
Asad	0,4526	0,2495	0,1978	0,1450	0,1001	0,0656
İnsha	0,4214	0,2445	0,1949	0,1432	0,0997	0,0654
Qadir	0,4708	0,2604	0,2072	0,1504	0,1035	0,0665
$\psi 1$	0,4456	0,2468	0,1957	0,1435	0,0993	0,0654
$\psi 2$	0,4420	0,2455	0,1949	0,1428	0,0991	0,0653
Theil	0,4199	0,2614	0,2067	0,1538	0,1059	0,0677
Siegel	0,5079	0,3134	0,2428	0,1869	0,1275	0,0805
ARM	1,3806	0,7633	0,5618	0,4228	0,2909	0,1771
LAD	0,4690	0,3065	0,2429	0,1788	0,1313	0,0795
LTS	0,4770	0,2927	0,2332	0,1701	0,1167	0,0716

Çizelge 4.2 dikkate alındığında, OLS tahmincisinin doğal olarak en düşük RMSE değeri verdiği görülmektedir. Bununla birlikte n=10 için Fair, Hampel, Huber ve Logistic M-tahmincileride OLS'ye yakın RMSE değerleri ortaya çıkartmıştır. Bu tahminciler, Hampel hariç monoton M-tahmincilerdir. Bu tahmincilerin etki fonksiyonlarını incelediğimiz zaman görülmektedir ki etkisi belli bir aralıkta OLS ye benzemektedir. Örnek birim sayısı arttıkça M-tahmincilerin birçoğu OLS ye yakın RMSE değeri üretmektedir. Bunlar arasından en kötü performansı sergileyen Geman&Mcclure ve Bell M-tahmincileridir. Öte yandan diğer kırılma noktası yüksek olan tahminciler ve LAD pek çok örnek boyutunda OLS ve M-tahmincilerinden daha yüksek RMSE değerleri ortaya çıkarmışlardır. Bunların kırılma noktaları tekrar hatırlanırsa, Theil, Siegel, ARM ve LTS sırası ile kırılma noktaları %29, %50, %25, %25 şeklindedir. Kırılma

noktası arttıkça normal dağılımda etkinliğin azaldığı görülmektedir. Bu tahminciler arasında büyük örnek boyutlarında bile etkinliği en kötü olanın ARM olduğu görülmektedir.

Hata dağılımı Laplace olduğunda M-tahmincilerin performanslarına ilişkin RMSE değerleri Çizelge 4.3'de verilmiştir.

Çizelge 4.3. Hata dağılımı Laplace olduğu durum için tahmincilerin RMSE değerleri

	n=10	**n=20**	**n=30**	**n=50**	**n=100**	**n=250**
OLS	0,3744	0,2454	0,1910	0,1475	0,1018	0,0637
Andrew	0,4055	0,2290	0,1728	0,1295	0,0906	0,0545
Bell	0,4453	0,2395	0,1727	0,1263	0,0872	0,0513
Cauchy	0,3507	0,2184	0,1652	0,1247	0,0876	0,0532
Danish	0,4058	0,2325	0,1797	0,1362	0,0950	0,0578
Fair	0,3449	0,2179	0,1650	0,1244	0,0871	0,0528
Geman and Mcclure	0,4835	0,2589	0,1911	0,1322	0,0908	0,0515
Hampel	0,3450	0,2218	0,1713	0,1300	0,0909	0,0561
Huber	0,3482	0,2194	0,1673	0,1267	0,0888	0,0545
Logistic	0,3444	0,2175	0,1648	0,1246	0,0874	0,0532
QSR	0,3746	0,2223	0,1657	0,1257	0,0881	0,0535
Ramsay	0,3451	0,2186	0,1665	0,1255	0,0881	0,0536
Talwar	0,4197	0,2407	0,1838	0,1394	0,0989	0,0602
Welsh	0,3898	0,2237	0,1690	0,1269	0,0891	0,0537
Tukey	0,4041	0,2285	0,1723	0,1292	0,0905	0,0544
Asad	0,3934	0,2304	0,1794	0,1338	0,0940	0,0572
insha	0,3688	0,2260	0,1758	0,1325	0,0930	0,0571
Qadir	0,4041	0,2285	0,1723	0,1292	0,0905	0,0544
$\psi 1$	0,3862	0,2332	0,1823	0,1364	0,0956	0,0584
$\psi 2$	0,3871	0,2344	0,1834	0,1377	0,0964	0,0591
Theil	0,3808	0,2249	0,1732	0,1277	0,0888	0,0539
Siegel	0,4492	0,2419	0,1824	0,1307	0,0919	0,0536
ARM	1,4682	0,5724	0,4448	0,2968	0,2247	0,1328
LAD	0,3779	0,2316	0,1702	0,1234	0,0851	0,0478
LTS	0,4187	0,2406	0,1809	0,1386	0,0947	0,0577

Çizelge 4.3 dikkate alındığında, n>29 olduğu durumda, LAD tahmincisinin en iyi sonucu verdiği görülmektedir. Bu durum OLS de olduğu gibi beklenen bir durumdur. Bununla birlikte n=10 için bazı M-tahmincilerin LAD den bile daha küçük RMSE değerleri üretmeleri göze çarpmaktadır. Normal dağılıma sahip hatalar için sonuçlara benzer olarak Geman&Mcclure ve Bell M-tahminciler arasında en kötü ancak OLS den yine de daha küçük RMSE değerleri

sağlamışlardır. Yine bir önceki duruma benzer olarak ARM'nin performansı en kötü çıkmıştır. Theil ve Siegel ise M-tahmincilere yakın sonuçlar sergilemektedir.

Hata dağılımı %10 Scale Contaminated (0.90N(0,1) – 0.10N(0,10)) olduğunda M-tahmincilerin performanslarına ilişkin RMSE değerleri Çizelge 4.4'de verilmiştir.

Çizelge 4.4. Hata dağılımı Scale-Contaminated (%10) olduğu durum için tahmincilerin RMSE değerleri

	n=10	**n=20**	**n=30**	**n=50**	**n=100**	**n=250**
OLS	1,1547	0,7688	0,6080	0,4689	0,3277	0,2109
Andrew	0,5110	0,2978	0,2224	0,1624	0,1111	0,0719
Bell	0,5530	0,3352	0,2517	0,1866	0,1258	0,0798
Cauchy	0,4858	0,2977	0,2256	0,1646	0,1154	0,0749
Danish	0,5060	0,3003	0,2247	0,1635	0,1106	0,0719
Fair	0,5496	0,3431	0,2621	0,1884	0,1340	0,0863
Geman and Mcclure	0,6250	0,3903	0,3005	0,2307	0,1544	0,0971
Hampel	0,5046	0,3181	0,2423	0,1753	0,1243	0,0808
Huber	0,4956	0,3122	0,2387	0,1719	0,1216	0,0786
Logistic	0,5041	0,3172	0,2428	0,1750	0,1239	0,0801
QSR	0,4856	0,3012	0,2283	0,1668	0,1156	0,0744
Ramsay	0,5011	0,3004	0,2281	0,1666	0,1176	0,0765
Talwar	0,5124	0,3053	0,2302	0,1654	0,1117	0,0723
Welsch	0,5011	0,2949	0,2199	0,1623	0,1116	0,0725
Tukey	0,5105	0,2973	0,2220	0,1624	0,1111	0,0719
Asad	0,5110	0,2958	0,2183	0,1610	0,1109	0,0726
insha	0,5086	0,2955	0,2194	0,1621	0,1130	0,0743
Qadir	0,5105	0,2973	0,2220	0,1624	0,1111	0,0719
$\boldsymbol{\psi 1}$	0,5120	0,2978	0,2203	0,1624	0,1123	0,0740
$\boldsymbol{\psi 2}$	0,5175	0,3002	0,2222	0,1637	0,1136	0,0751
Theil	0,5544	0,3409	0,2519	0,1802	0,1271	0,0822
Siegel	0,6271	0,3742	0,2777	0,2007	0,1419	0,0929
ARM	3,9630	0,8715	0,6292	0,4611	0,3283	0,2026
LAD	0,5679	0,3582	0,2699	0,2043	0,1401	0,0882
LTS	0,5112	0,3138	0,2330	0,1725	0,1163	0,0735

Yukarıda bahsedilen diğer tabloların sonuçlarında olduğu gibi n>10 olduğunda monoton M-tahmincilerin sonuçları Çizelge 4.4'de de en iyi sonuçları üretmemektedir. Burada özellikle yeniden azalanlar daha iyi performans sergilemektedir. Theil ve LTS diğer ele alınan başlangıç tahmincileri içinde en küçük RMSE'ye sahiptirler.

Hata dağılımı %20 Scale Contaminated (0.80N(0,1) – 0.20N(0,10)) olduğunda M-tahmincilerin performanslarına ilişkin RMSE değerleri Çizelge 4.5'de verilmiştir.

Çizelge 4.5. Hata dağılımı Scale-Contaminated (%20) olduğu durum için tahmincilerin RMSE değerleri

	n=10	**n=20**	**n=30**	**n=50**	**n=100**	**n=250**
OLS	1,7533	1,1235	0,9010	0,6774	0,4721	0,2862
Andrew	0,6336	0,3063	0,2477	0,1740	0,1293	0,0803
Bell	0,6621	0,3299	0,2567	0,1795	0,1322	0,0830
Cauchy	0,6482	0,3210	0,2655	0,1961	0,1435	0,0864
Danish	0,6314	0,3099	0,2498	0,1748	0,1306	0,0804
Fair	0,8965	0,4296	0,3468	0,2661	0,1821	0,1076
Geman and Mcclure	0,7133	0,3813	0,3149	0,2150	0,1581	0,0971
Hampel	0,7760	0,3792	0,3116	0,2392	0,1683	0,0999
Huber	0,7369	0,3574	0,2926	0,2215	0,1571	0,0937
Logistic	0,7661	0,3691	0,3028	0,2302	0,1620	0,0962
QSR	0,6339	0,3084	0,2512	0,1824	0,1338	0,0820
Ramsay	0,6707	0,3377	0,2802	0,2089	0,1524	0,0911
Talwar	0,6323	0,3177	0,2511	0,1756	0,1320	0,0819
Welsh	0,6459	0,3053	0,2501	0,1770	0,1319	0,0814
Tukey	0,6337	0,3047	0,2478	0,1741	0,1294	0,0803
Asad	0,6425	0,3209	0,2619	0,1869	0,1383	0,0859
insha	0,6613	0,3322	0,2738	0,1983	0,1467	0,0896
Qadir	0,6337	0,3047	0,2478	0,1741	0,1294	0,0803
$\boldsymbol{\psi 1}$	0,6510	0,3322	0,2732	0,1980	0,1455	0,0906
$\boldsymbol{\psi 2}$	0,6532	0,3432	0,2813	0,2071	0,1505	0,0939
Theil	0,7811	0,3876	0,3137	0,2355	0,1639	0,0984
Siegel	0,8159	0,4058	0,3278	0,2430	0,1697	0,1036
ARM	5,4165	1,0967	0,8294	0,5366	0,3649	0,2223
LAD	0,8051	0,3751	0,3181	0,2277	0,1639	0,0985
LTS	0,6332	0,3152	0,2497	0,1723	0,1289	0,0785

Yukarıdaki tablodan Normal dağılım ve Laplace dağılımında görülen monoton M-tahmincilerin sonuçlarının burada iyi olmadığı yeniden azalan M-tahmincilerin performansının daha iyi olduğu görülmektedir. Ayrıca M-tahmincilerin yanı sıra, LTS de en iyi sonucu vermektedir.

Hata dağılımı Student-t(3) olduğunda M-tahmincilerin performanslarına ilişkin RMSE değerleri Çizelge 4.6'da verilmiştir.

Çizelge 4.6. Hata dağılımı Student-t(3) olduğu durum için tahmincilerin RMSE değerleri

	n=10	n=20	n=30	n=50	n=100	n=250
OLS	0,3981	0,2142	0,1969	0,1384	0,1039	0,0656
Andrew	0,3553	0,1893	0,1450	0,1116	0,0765	0,0486
Bell	0,3809	0,2072	0,1542	0,1184	0,0794	0,0485
Cauchy	0,3257	0,1751	0,1421	0,1099	0,0761	0,0482
Danish	0,3521	0,1935	0,1474	0,1150	0,0787	0,0500
Fair	0,3230	0,1766	0,1461	0,1120	0,0781	0,0491
Geman and Mcclure	0,4058	0,2384	0,1815	0,1356	0,0906	0,0540
Hampel	0,3208	0,1760	0,1458	0,1127	0,0786	0,0499
Huber	0,3201	0,1758	0,1430	0,1110	0,0768	0,0488
Logistic	0,3197	0,1751	0,1435	0,1107	0,0769	0,0485
QSR	0,3337	0,1800	0,1422	0,1104	0,0755	0,0478
Ramsay	0,3265	0,1756	0,1437	0,1109	0,0770	0,0488
Talwar	0,3607	0,1998	0,1494	0,1174	0,0796	0,0516
Welsh	0,3428	0,1803	0,1431	0,1103	0,0761	0,0482
Tukey	0,3533	0,1882	0,1446	0,1114	0,0764	0,0485
Asad	0,3533	0,1850	0,1484	0,1139	0,0795	0,0506
insha	0,3348	0,1805	0,1478	0,1137	0,0796	0,0507
Qadir	0,3533	0,1882	0,1446	0,1114	0,0764	0,0485
$\psi 1$	0,3521	0,1885	0,1509	0,1158	0,0815	0,0518
$\psi 2$	0,3515	0,1902	0,1523	0,1170	0,0828	0,0524
Theil	0,3415	0,1943	0,1485	0,1143	0,0789	0,0497
Siegel	0,3880	0,2245	0,1612	0,1253	0,0856	0,0540
ARM	1,2134	0,5527	0,4085	0,2793	0,1970	0,1226
LAD	0,3602	0,2014	0,1597	0,1232	0,0818	0,0513
LTS	0,3584	0,2061	0,1516	0,1179	0,0805	0,0488

Yukarıdaki tabloya dikkat edilirse, çoğu M-tahminciler OLS, Theil, Siegel, ARM ve LAD den küçük örneklemlerde iyi sonuç sergilemektedir. Örnek birim sayısı n=250 olduğunda dahi M-tahmincilerinin performansı diğer tahmincilere göre tatminkârdır.

Daha önceki analizlerden farklı olarak Hata dağılımı lognormal olduğunda M-tahmincilerinin ortaya çıkardıkları RMSE değerleri Çizelge 4.7'de verilmiştir.

Çizelge 4.7. Hata dağılımı lognormal olduğu durum için tahmincilerin RMSE değerleri

	n=10	n=20	n=30	n=50	n=100	n=250
OLS	0,7947	0,5090	0,4105	0,3140	0,2218	0,1344
Andrew	0,4984	0,2447	0,1968	0,1382	0,0983	0,0581
Bell	0,5046	0,2433	0,1851	0,1279	0,0905	0,0526
Cauchy	0,5220	0,2583	0,2109	0,1526	0,1061	0,0623
Danish	0,5068	0,2425	0,1968	0,1386	0,0998	0,0594
Fair	0,5584	0,3020	0,2471	0,1826	0,1269	0,0745
Geman and Mcclure	0,5553	0,2782	0,1987	0,1396	0,0933	0,0533
Hampel	0,5263	0,2716	0,2216	0,1610	0,1122	0,0653
Huber	0,5222	0,2624	0,2140	0,1543	0,1055	0,0610
Logistic	0,5343	0,2778	0,2268	0,1659	0,1148	0,0672
QSR	0,4846	0,2358	0,1905	0,1343	0,0924	0,0535
Ramsay	0,5331	0,2705	0,2212	0,1623	0,1131	0,0662
Talwar	0,5100	0,2471	0,1981	0,1427	0,1039	0,0628
Welsh	0,4994	0,2452	0,2008	0,1412	0,0993	0,0587
Tukey	0,4981	0,2444	0,1970	0,1382	0,0983	0,0581
Asad	0,5018	0,2572	0,2139	0,1510	0,1067	0,0631
insha	0,5251	0,2666	0,2193	0,1572	0,1104	0,0649
Qadir	0,4981	0,2444	0,1970	0,1382	0,0983	0,0581
$\psi 1$	0,5078	0,2671	0,2233	0,1593	0,1127	0,0666
$\psi 2$	0,5131	0,2727	0,2293	0,1650	0,1164	0,0688
Theil	0,4977	0,2389	0,1845	0,1370	0,0905	0,0515
Siegel	0,5971	0,2936	0,2278	0,1742	0,1196	0,0720
ARM	3,3418	0,9717	0,6481	0,4526	0,2942	0,1697
LAD	0,5913	0,3161	0,2533	0,1849	0,1294	0,0782
LTS	0,5086	0,2371	0,1870	0,1278	0,0887	0,0527

Yukarıdaki tablodan görüldüğü üzere, hata dağılımı asimetrik olan lognormal dağılım olduğunda dahi iyi tanınan M-tahmincileri OLS den daha küçük RMSE değeri ortaya çıkarmaktadır. Ayrıca monoton M-tahminciler yeniden azalan (redescending) olanlara göre daha büyük RMSE değerleri ortaya çıkarmışlardır, yani performansları daha kötüdür. Yüksek kırılma noktasına sahip olanlar arasından, RMSE kriterine göre en iyi tahminci Theil olarak görülmektedir.

4.4. Aykırı Değer Teşhis Metodu Olarak M-Tahminciler

M-tahmincilerin red noktası (cut point) belli bir değerden sonra ψ fonksiyonunun sıfır olarak alınmasını belirleyen noktadır. Bu noktanın ilerisinde kalan noktalar tahmin üzerinde sıfır etkiye sahip olacaklardır. Bu şekilde bir sonlu red noktasına sahip tahmincilere Hızlı Yeniden Azalan Tahminci (Hard Redescending Estimator) denir ve bu tahminciler büyük aykırı değerlere karşı çok iyi korunurlar. Ayrıca, bu tahminciler direk olmayan aykırı değer teşhis metodu olarak da çalışırlar (Billor ve Kiral, 2008).

Bu bölümde aykırı değer teşhis metodu olarak M-tahmincilerin performansı değerlendirilecektir. Daha önceki bölümde olduğu gibi veri %10 oranında kirletilecek yani belli sayıda aykırı değer veri setine atılacak ve bu sayının en iyi şekilde tahminini veren M-tahmincisi belirlenmeye çalışılacaktır. Yani aykırı değer teşhis (outliers detection) metodu olarak M-tahminciler incelenecektir.

Çizelge 4.8. Aykırı Değer Teşhis Eden M-Tahminciler

n	Andrew	Danish	Talwar	Tukey	Asad	Qadir	T1	T2
10	1	1	1	1	1	1	1	1
30	2	3	3	2	2	2	2	2
50	3	5	4	3	3	3	3	3
75	4	6	6	4	4	4	4	4
100	7	9	8	6	6	6	6	6
250	16	21	19	15	15	15	15	15

Çizelge 4.8'e bakıldığında n=10 için tüm M-tahminciler %10 oranında aykırı değeri doğru bir şekilde tespit etmektedir. n=30 için ise Danish ve Talwar'ın M-tahmincisi verideki 3 aykırı değeri tam olarak tespit etmektedir. n sayısı arttıkça tahmincilerin aykırı değeri yakalama oranlarında biraz azalma söz konusudur. Çizelgeye bakıldığında, ele alınan tahminciler arasında aykırı değerleri doğru bir şekilde tespit etmeleri bakımından en iyisi Danish M-tahmincisi olduğu görülmektedir.

4.5. Ölçek Tahmincilerinin Etkinliği

Regresyon analizinde M-tahmincileri hesaplanırken başlangıç ölçek tahmini olarak genellikle, %50 kırılma noktasına sahip olmasından dolayı Medyan Mutlak Sapma (MAD- Median Absolute Deviation) kullanılmaktadır. Ancak MAD tahmincisinin etkinliği düşüktür (Rousseeuw and Croux, 1993). Bu bölümde MAD'a alternatif iki ölçek tahmincisinin etkinliği araştırılmıştır. Etkinlikten kasıt hata dağılımı normal olduğunda bu tahmincilerin MSE açısından değerlendirilmesidir. Aşağıda Çizelge 4.9- Çizelge 4.11 hata dağılımı normal olduğunda regresyon eğim parametresinin MSE değerleri, örnek birim sayısı n=10, 20 ve 30 olduğu durumlar için verilmiştir.

Çizelge 4.9. Hata dağılımı normal iken başlangıç ölçek tahmincilerinin etkinliği (n=10)

	n=10		
Başlangıç Ölçek Tahmincisi	MAD	Sn	Qn
Andrew	0,1748	0,1701	0,1590
Bell	0,2438	0,2501	0,2428
Cauchy	0,1492	0,1472	0,1448
Danish	0,1607	0,1587	0,1460
Fair	0,1418	0,1414	0,1406
Geman and Mcclure	0,3036	0,3053	0,3037
Hampel	0,1431	0,1416	0,1400
Huber	0,1465	0,1456	0,1434
Logistic	0,1440	0,1431	0,1415
QSR	0,1672	0,1656	0,1604
Ramsay	0,1439	0,1424	0,1403
Talwar	0,1537	0,1496	0,1423
Welsh	0,1662	0,1615	0,1545
Tukey	0,1746	0,1696	0,1589
Asad	0,1607	0,1510	0,1482
insha	0,1507	0,1450	0,1384
Qadir	0,1746	0,1696	0,1589
$\boldsymbol{\psi 1}$	0,1502	0,1423	0,1392
$\boldsymbol{\psi 2}$	0,1463	0,1412	0,1374

Çizelge 4.10. Hata dağılımı normal iken başlangıç ölçek tahmincilerinin etkinliği (*n*=20)

Başlangıç Ölçek Tahmincisi	n=20		
	MAD	Sn	Qn
Andrew	0,0635	0,0635	0,0621
Bell	0,0876	0,0888	0,0873
Cauchy	0,0615	0,0616	0,0612
Danish	0,0626	0,0615	0,0607
Fair	0,0612	0,0613	0,0611
Geman and Mcclure	0,1306	0,1283	0,1287
Hampel	0,0596	0,0598	0,0595
Huber	0,0615	0,0615	0,0611
Logistic	0,0614	0,0615	0,0612
QSR	0,0663	0,0665	0,0654
Ramsay	0,0605	0,0606	0,0603
Talwar	0,0626	0,0617	0,0602
Welsh	0,0626	0,0625	0,0616
Tukey	0,0634	0,0634	0,0620
Asad	0,0601	0,0598	0,0592
insha	0,0592	0,0591	0,0588
Qadir	0,0634	0,0634	0,0620
$\boldsymbol{\psi 1}$	0,0594	0,0590	0,0590
$\boldsymbol{\psi 2}$	0,0593	0,0588	0,0589

Çizelge 4.11. Hata dağılımı normal iken başlangıç ölçek tahmincilerinin etkinliği (*n*=30)

Başlangıç Ölçek Tahmincisi	*n*=30		
	MAD	Sn	Qn
Andrew	0,0406	0,0398	0,0392
Bell	0,0549	0,0551	0,0552
Cauchy	0,0388	0,0387	0,0385
Danish	0,0411	0,0403	0,0391
Fair	0,0384	0,0383	0,0383
Geman and Mcclure	0,0843	0,0846	0,0844
Hampel	0,0379	0,0377	0,0375
Huber	0,0386	0,0386	0,0385
Logistic	0,0386	0,0385	0,0384
QSR	0,0412	0,0413	0,0412
Ramsay	0,0381	0,0380	0,0379
Talwar	0,0409	0,0400	0,0397
Welsh	0,0396	0,0392	0,0388
Tukey	0,0404	0,0398	0,0391
Asad	0,0387	0,0379	0,0374
insha	0,0376	0,0373	0,0371
Qadir	0,0404	0,0398	0,0391
$\boldsymbol{\psi 1}$	0,0380	0,0374	0,0370
$\boldsymbol{\psi 2}$	0,0376	0,0372	0,0369

Yukarıdaki tablolardan görülmektedir ki MAD'a alternatif Sn ve Qn başlangıç ölçek tahmini olarak alındığında eğim parametresinin MSE değerleri düşmüş yani normal dağılımda etkinliği artmıştır. Ayrıca alternatif ölçek tahmincileri arasından eğim parametresi tahminine en düşük MSE değeri sağlayan ölçek tahmincisi Qn olduğu tablolardan görülmektedir.

5. UYGULAMA

Bu bölümde iki farklı uygulama üzerinde M-Tahmincilerin sonuçları incelenecektir. İki uygulamanın birbirinden farklılığı ilkinde aykırı değer barındıran gerçek gözlemler kullanılırken, ikincisinde bir gözlemin değerinin değiştirilmesiyle yapay bir aykırı değer elde edilmiştir. İkinci uygulamanın kullanılmasının bir avantajı da gözlemlerin yanlış kayıt edilmesinden kaynaklanan *"kötü"* aykırı değerlerin M-Tahminleme süreciyle etkisinin giderilerek veri setinin yeniden kazanılmasıdır.

İlk uygulama olarak Rousseeuw ve Leroy (1987)'dan alınan *"Çin'de fiyatların büyümesinin geleneksel oranları (Annual Rates of Growth of Prices in China)"* veri seti üzerinde M-tahmincilerin önemi gösterilmiştir.

Çizelge 5.1. Çin'de fiyatların büyümesinin geleneksel oranları veri seti

Gözlem No	**Yıllar** (x_i)	**Fiyatların Büyümesi (%)** (y_i)
1	40	1,62
2	41	1,63
3	42	1,9
4	43	2,64
5	44	2,05
6	45	2,13
7	46	1,94
8	47	15,5
9	48	364

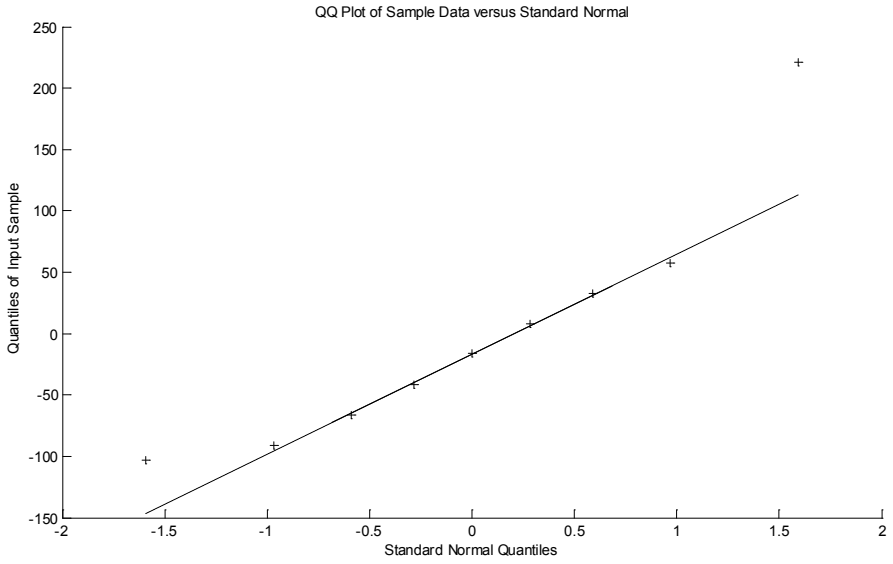
Çizelge 5.1. Çin'in ana şehirlerinde 1940 ile 1948 yılları arasındaki ortalama fiyatların büyümesinin oranlarının bilgilerini içermektedir. Örneğin 1940 yılında fiyatlar bir önceki yıla göre %1.62 oranında bir artış göstermiştir. 1948 yılında ise hükümetin çok büyük harcamaları, bütçe açıkları ve savaş durumu çok büyük oranda fiyatların artmasına sebep olmuştur. Bu duruma *"Hiper Enflasyon"* denilmektedir.

Bu örnek için OLS artıklarına ilişkin betimleyici istatistikler Çizelge 5.2.'de verilmiştir.

Çizelge 5.2. Çin'de fiyatların büyümesinin geleneksel oranları modeli OLS artıkları için betimleyici istatistikler ve Jarque-Bera normallik testi

Betimsel İstatistik						Normalik Testi
Min	Max	ort	Varyans	Çarpıklık	Basıklık	JB
-102,75	220,91	-2,6874e-012	9816,42	1,20	3,79	2,3848

Çizelge 5.2.'de OLS artık değerlerinin betimsel istatistik değeri ve yapılan Jarque-Bera (JB) normallik testi sonuçları verilmiştir. JB testinin kritik değeri 2,3352 olarak çıkmış ve dolayısıyla OLS artık değerlerinin normal dağıldığı hipotez testi α=0,05 anlamlılık düzeyinde reddedilmiştir.



Şekil 5.1. Çin'de fiyatların büyümesinin geleneksel oranları modeli OLS artıkları için Q-Q grafiği

Artıklar için çizilen Q-Q grafiğine bakıldığında y-yönlü olası iki artık değerin olduğu görülmektedir.

Araştırılan regresyon modeli için katsayı tahminleri ve teşhis edilen aykırı değer sayıları Çizelge 5.3.'de verilmiştir.

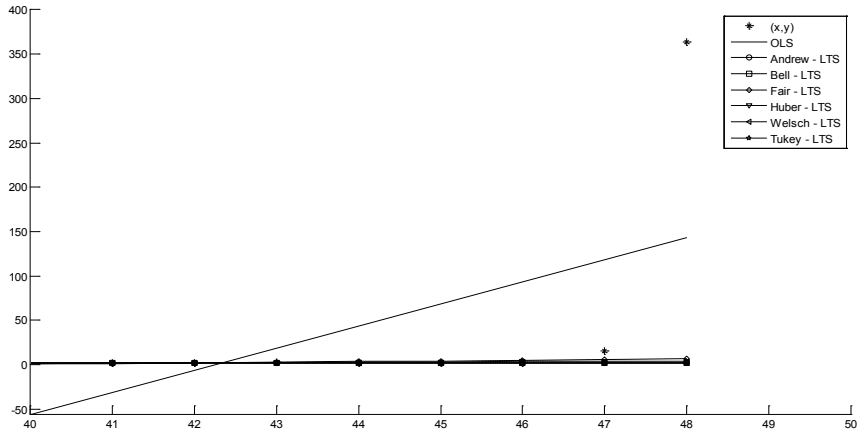
Çizelge 5.3. Çin'de fiyatların büyümesinin geleneksel oranları veri seti uygulama sonuçları

Tahminciler	Regresyon Parametreleri		Aykırı değer Sayısı
	$\hat{\beta}_0$	$\hat{\beta}_1$	
OLS	-1049,47	24,85	-
Andrew	-1,35	0,08	2
Bell	-1,55	0,08	-
Cauchy	-1,59	0,08	-
Danish	-1,25	0,08	2
Fair	-28,43	0,72	-
Geman and Mcclure	-2,15	0,09	-
Hampel	-12,61	0,35	-
Huber	-12,46	0,34	-
Logistic	-13,76	0,38	-
QSR	-1,67	0,08	-
Ramsay	-1,43	0,08	-
Talwar	-1,25	0,08	2
Welsch	-1,36	0,08	-
Tukey	-1,35	0,08	2
Asad	-1,27	0,08	2
insha	-1,27	0,08	-
Qadir	-1,35	0,08	2
$\psi 1$	-1,26	0,08	2
$\psi 2$	-1,25	0,08	2
Theil	-10,88	0,31	-
Siegel	-6,71	0,21	-
ARM	1,61	0,01	-
LAD	-19,07	0,50	-
LTS	-1,25	0,08	-

Çizelge 5.3. incelendiğinde görüldüğü gibi OLS'nin katsayı tahminleri aykırı değerlerden etkilenmektedir. Ayrıca Q-Q grafiğinde görülen aykırı değerlerden oldukça etkilendiği görülmektedir. Tablo dikkatle incelendiğinde OLS'nin eğim katsayısı çok yüksek çıkmış fakat birer M-tahminci olan Fair,

Hampel, Huber ve Logistic ile LAD, Theil ve Siegel'ın da değerleri diğer M-tahmincilere göre oldukça büyük çıkmıştır. Ayrıca tablo incelenildiğinde iki adet gözlem değeri, hızlı yeniden azalan M-tahminciler aykırı değer olarak teşhis etmişler ve etkilerini sıfır vermek suretiyle yok etmişlerdir.

Regresyon katsayılarına ilişkin grafik şekil 5.2.'de verilmiştir.



Şekil 5.2. Çin'de fiyatların büyümesinin geleneksel oranları veri seti eğim katsayıları grafiği

Şekil 5.2.'de ise tahmin edilmiş regresyon doğruları görülmektedir. OLS doğrusun aykırı değere doğru çekildiği yine görülmektedir. Fakat M-tahminciler aykırı değerlerden etkilenmemişlerdir. OLS doğrusu ve M-tahmincilerin doğrusunun aralarındaki eğim katsayısı farkı oldukça yüksektir. OLS doğrusu incelendiğinde tek bir verinin etkisinden dolayı o veri istikametinde gitmekte fakat diğer noktaları kapsamamaktadır. Böylece aslında aykırı değer olmayan veriler büyük artıklara sahip olacaklar ve artık aykırı değer olacaklardır. Bu ise tamamıyla yanlış bir bilgiye götürecek ve yanlış yorumlar yapılmasına sebebiyet verecektir. Bu durum çizelge 5.4.'de daha iyi bir şekilde görülmektedir, tahmin edilen bağımlı değişken verileri ile gerçek değerleri kıyaslandığında bu fark çok net ortaya çıkmaktadır.

Çizelge 5.4. Çin'de fiyatların büyümesinin geleneksel oranları veri seti tahmin edilen bağımlı değişken değerleri

Gözlem No	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Gerçek y	**1,62**	**1,63**	**1,90**	**2,64**	**2,05**	**2,13**	**1,94**	**15,50**	**364,00**
OLS	-55,67	-30,82	-5,98	18,87	43,71	68,56	93,40	118,25	143,09
Andrew	1,71	1,79	1,86	1,94	2,02	2,09	2,17	2,25	2,32
Bell	1,65	1,73	1,81	1,89	1,97	2,05	2,13	2,21	2,29
Cauchy	1,70	1,78	1,86	1,94	2,03	2,11	2,19	2,27	2,35
Danish	1,76	1,84	1,91	1,99	2,06	2,14	2,21	2,29	2,36
Fair	0,54	1,27	1,99	2,71	3,44	4,16	4,89	5,61	6,34
Geman and Mcclure	1,63	1,72	1,82	1,91	2,01	2,10	2,20	2,29	2,38
Hampel	1,29	1,63	1,98	2,33	2,68	3,02	3,37	3,72	4,07
Huber	1,32	1,67	2,01	2,35	2,70	3,04	3,39	3,73	4,08
Logistic	1,24	1,62	1,99	2,37	2,74	3,12	3,49	3,87	4,24
QSR	1,68	1,76	1,84	1,93	2,01	2,09	2,18	2,26	2,35
Ramsay	1,71	1,79	1,87	1,95	2,03	2,10	2,18	2,26	2,34
Talwar	1,76	1,84	1,91	1,99	2,06	2,14	2,21	2,29	2,36
Welsch	1,71	1,78	1,86	1,94	2,02	2,09	2,17	2,25	2,32
Tukey	1,71	1,79	1,86	1,94	2,02	2,09	2,17	2,25	2,32
Asad	1,75	1,82	1,90	1,97	2,05	2,12	2,20	2,27	2,35
insha	1,75	1,82	1,90	1,97	2,05	2,13	2,20	2,28	2,35
Qadir	1,71	1,79	1,86	1,94	2,02	2,09	2,17	2,25	2,32
$\psi 1$	1,76	1,83	1,91	1,98	2,06	2,13	2,21	2,28	2,36
$\psi 2$	1,76	1,84	1,91	1,99	2,06	2,14	2,21	2,29	2,36
Theil	1,33	1,63	1,94	2,24	2,55	2,85	3,16	3,46	3,77
Siegel	1,49	1,70	1,90	2,11	2,31	2,52	2,72	2,93	3,13
ARM	2,01	2,02	2,03	2,04	2,05	2,06	2,07	2,08	2,09
LAD	1,13	1,63	2,13	2,64	3,14	3,65	4,15	4,66	5,16
LTS	1,76	1,84	1,91	1,99	2,06	2,14	2,21	2,29	2,36

Yukarıdaki tablo incelendiğinde OLS tahmin değerlerinin ne kadar yanlış sonuçlar çıkardığı görülmektedir. OLS 364 değerinin çok büyük etkisinde kalmasına rağmen o değeri doğru tahmin edemediği gibi diğer iyi noktalardaki tahminlerini de bozmuştur. Bu durumda OLS tamamen kullanışsız bir tekniğe dönüşmüştür. Eğer robust teknikler olmasaydı bu veri seti doğru bir şekilde

modellenemeyecekti, fakat robust tekniklerin başarısı bu veri seti için modelleme yapılmasına imkan vermiştir.

İkinci uygulama olarak Hampel (1986, syf 309-310)'dan alınan *"Su Akıntısı Veri Seti (Water Flow Data)"* üzerinde m-tahmincilerin uygulaması yapılmıştır. Bu veri setinin özelliği bir gözlemin gerçek değerinin yapay olarak başka bir değerle değiştirilmesi ve bu durumun da kayıt tutma hatalarından kaynaklanan aykırı değerler olması durumuna örnek teşkil etmesidir. Bu örnekte değiştirilmiş olan veri seti için M-tahminciler uygulanacaktır. Bu veri seti Hampel (1986, syf 309-310)'dan alınmıştır; Hampel da Ezekiel and Fox (1959, pp. 57-58)'dan almıştır.

Kootenay nehrinde su akıntıları iki farklı noktada (Libby, Mont. And Newgate, B.C.) ölçülmüştür. Bu iki nokta arasında Ezekiel ve Fox (1959) tarafından bir regresyon denklemi kurulmuş ve Newsgate'deki su akıntısı Libby'deki su akıntısı ile modellenmiştir. Fakat Hampel ve diğerleri (1986) aykırı gözlemlerin OLS üzerindeki etkisini göstermek için 1934 yılına ait olan (77,6, 44,9) gözlemi (20,0, 44,9) olarak değiştirmişlerdir.

Çizelge 5.5. Su akıntısı veri seti

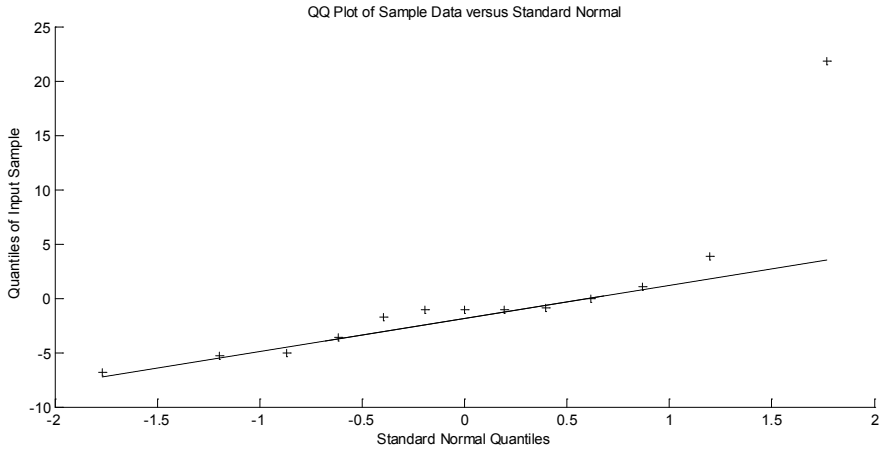
Yıl	Gerçek Gözlemler		Değiştirilmiş Gözlemler	
	x_i	y_i	x_i	y_i
1931	27,1	19,7	27,1	19,7
1932	20,9	18	20,9	18
1933	33,4	26,1	33,4	26,1
1934	77,6	44,9	20	44,9
1935	37	26,1	37	26,1
1936	21,6	19,9	21,6	19,9
1937	17,6	15,7	17,6	15,7
1938	35,1	27,6	35,1	27,6
1939	32,6	24,9	32,6	24,9
1940	26	23,4	26	23,4
1941	27,6	23,1	27,6	23,1
1942	38,7	31,3	38,7	31,3
1943	27,8	23,8	27,8	23,8

Çizelge 5.6. Su akıntısı modeli OLS artıkları için betimleyici istatistikler ve Jarque-Bera normallik testi

Betimsel İstatistik						Normalik Testi
Min	Max	ort	Varyans	Çarpıklık	Basıklık	JB
-6,84	21,81	4,4546e-14	51,006	233	7,99	25,2195

Bir önceki örnekle pareler olarak JB testinin kritik değeri 3.0310 olarak çıkmış ve dolayısıyla OLS artık değerlerinin normal dağıldığı hipotez testi α=0,05 anlamlılık düzeyinde reddedilmiştir.

Artıklara ilişkin Q-Q grafiği şekil 5.3.'de verilmiştir.



Şekil 5.3. Su akıntısı modeli OLS artıkları için Q-Q grafiği

Yukarıdaki grafik incelendiğinde artıkların normal dağılıma sahip olmadığı görülmektedir. Ayrıca bir değer de olası bir aykırı değer olarak görülmektedir.

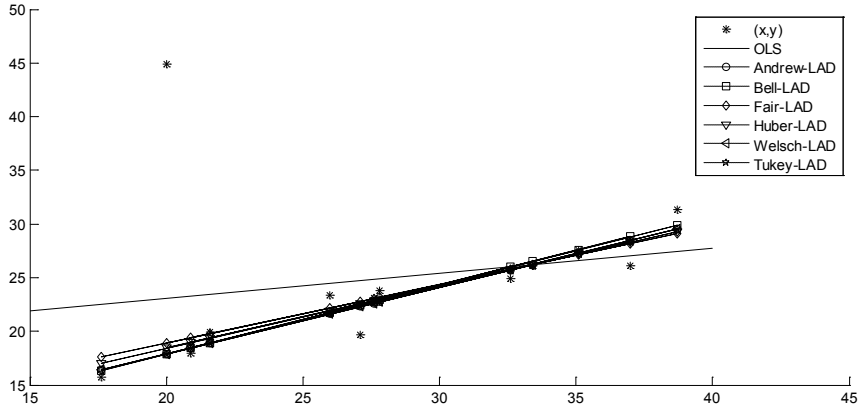
Araştırılan regresyon modeli için katsayı tahminleri ve teşhis edilen aykırı değer sayıları Çizelge 5.7.'de verilmiştir.

Çizelge 5.7. Su akıntısı veri seti uygulama sonuçları

Tahminciler	Regresyon Parametreleri		Aykırı değer Sayısı
	$\hat{\beta}_0$	$\hat{\beta}_1$	
OLS	18,48	0,23	-
Andrew	5,46	0,62	1
Bell	5,05	0,64	-
Cauchy	5,83	0,61	-
Danish	5,49	0,62	1
Fair	8,00	0,55	-
Geman and Mcclure	4,56	0,66	-
Hampel	7,11	0,57	-
Huber	6,84	0,58	-
Logistic	7,02	0,57	-
QSR	5,72	0,62	-
Ramsay	5,56	0,62	-
Talwar	5,49	0,62	1
Welsch	5,45	0,62	-
Tukey	5,46	0,62	1
Asad	5,49	0,62	1
insha	5,49	0,62	-
Qadir	5,46	0,62	1
$\psi 1$	5,49	0,62	1
$\psi 2$	5,49	0,62	1
Theil	6,68	0,59	-
Siegel	7,24	0,57	-
ARM	4,57	0,66	-
LAD	6,54	0,60	-
LTS	5,49	0,62	-

Çizelge 5.7.'de yine OLS tahminlerinin aykırı değerden etkilendiği görülmektedir. Eğim katsayıları incelendiğinde en çok 0.62 değerinin tahmin edildiği görülmektedir. Burada yine monoton tahminciler 0.62'den farklı sonuçlar elde etmişler fakat yeniden azalan (redescending) tahmincilerin büyük bir kısmı hep aynı değeri 0.62'yi tahmin etmişlerdir. Fakat OLS'nin eğim katsayı tahmini değeri yaklaşık olarak robust tekniklerin tahmininin yarısı kadardır.

Regresyon katsayılarına ilişkin grafik şekil 5.4.'de verilmiştir.



Şekil 5.4. Su akıntısı veri seti eğim katsayıları grafiği

Şekil 5.4.'den yine görülmektedir ki OLS doğrusu aykırı değere doğru çekilmektedir ve veri noktalarının büyük bir kısmının üzerinden geçmemektedir. Fakat M-tahminciler bu veri seti için aykırı değerlere karşı daha sağlam tahminleme yapmışlar ve veri noktalarının büyük bir kısmının üzerinden geçmişlerdir.

Çizelge 5.8.'de tahmin edilen bağımlı değişken değerleri verilmiştir. Tablo incelendiğinde robust tekniklerin daha iyi sonuçlar verdiği gözlemlenmektedir, yani artık değerlerini küçülttüğü görülmektedir. Bu durum küçük y değerlerinde daha açık bir şekilde kendini göstermektedir.

Çizelge 5.8. Su akıntısı veri seti tahmin edilen bağımlı değişken değerleri

Gözlem No	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
Gerçek y	**19,7**	**18**	**26,1**	**44,9**	**26,1**	**19,9**	**15,7**	**27,6**	**24,9**	**23,4**	**23,1**	**31,3**	**23,8**
OLS	24,7	23,3	26,2	23,1	27,0	23,5	22,5	26,6	26,0	24,5	24,8	27,4	24,9
Andrew	22,3	18,4	26,2	17,9	28,5	18,9	16,4	27,3	25,7	21,6	22,6	29,5	22,7
Bell	22,5	18,5	26,5	17,9	28,8	18,9	16,4	27,6	26,0	21,7	22,8	29,9	22,9
Cauchy	22,4	18,6	26,2	18,0	28,4	19,0	16,6	27,3	25,7	21,7	22,7	29,5	22,8
Danish	22,3	18,4	26,2	17,9	28,4	18,9	16,4	27,2	25,7	21,6	22,6	29,4	22,7
Fair	22,8	19,4	26,2	18,9	28,2	19,8	17,6	27,1	25,8	22,2	23,0	29,1	23,2
Geman and Mcclure	22,5	18,4	26,7	17,8	29,0	18,8	16,2	27,8	26,1	21,8	22,8	30,2	22,9
Hampel	22,6	19,0	26,2	18,5	28,2	19,4	17,1	27,1	25,7	21,9	22,9	29,2	23,0
Huber	22,6	19,0	26,2	18,4	28,3	19,4	17,0	27,2	25,7	21,9	22,8	29,3	23,0
Logistic	22,6	19,0	26,2	18,5	28,3	19,4	17,1	27,2	25,8	22,0	22,9	29,3	23,0
QSR	22,4	18,6	26,3	18,1	28,5	19,0	16,6	27,4	25,8	21,8	22,7	29,6	22,9
Ramsay	22,3	18,5	26,2	17,9	28,5	18,9	16,5	27,3	25,7	21,7	22,6	29,5	22,8
Talwar	22,3	18,4	26,2	17,9	28,4	18,9	16,4	27,2	25,7	21,6	22,6	29,4	22,7
Welsch	22,3	18,4	26,2	17,9	28,5	18,9	16,4	27,3	25,7	21,6	22,6	29,5	22,7
Tukey	22,3	18,4	26,2	17,9	28,4	18,9	16,4	27,3	25,7	21,6	22,6	29,5	22,7
Asad	22,3	18,4	26,2	17,9	28,4	18,9	16,4	27,2	25,7	21,6	22,6	29,4	22,7
insha	22,3	18,4	26,2	17,9	28,4	18,9	16,4	27,2	25,7	21,6	22,6	29,4	22,7
Qadir	22,3	18,4	26,2	17,9	28,4	18,9	16,4	27,3	25,7	21,6	22,6	29,5	22,7
ψ**1**	22,3	18,4	26,2	17,9	28,4	18,9	16,4	27,2	25,7	21,6	22,6	29,4	22,7
ψ**2**	22,3	18,4	26,2	17,9	28,4	18,9	16,4	27,2	25,7	21,6	22,6	29,4	22,7
Theil	22,8	19,1	26,6	18,6	28,7	19,5	17,2	27,6	26,1	22,1	23,1	29,7	23,2
Siegel	22,8	19,2	26,4	18,7	28,5	19,7	17,4	27,4	26,0	22,2	23,1	29,5	23,2
ARM	22,4	18,3	26,5	17,7	28,8	18,7	16,1	27,6	26,0	21,6	22,7	30,0	22,8
LAD	22,8	19,1	26,6	18,5	28,7	19,5	17,1	27,6	26,1	22,1	23,1	29,8	23,2
LTS	22,3	18,4	26,2	17,9	28,4	18,9	16,4	27,2	25,7	21,6	22,6	29,4	22,7

6. SONUÇLAR

Robust regresyon tahmincileri, hataların normal dağılıma uymadığı veya veri setinde aykırı değer bulunması durumunda regresyon modelini en güvenilir şekilde tahmin etmek amacı ile geliştirilmiştir.

Bu tezde, lineer regresyon modeli için geliştirilen robust tahmincilerden M-tahmincileri çeşitli açılardan incelenmiştir. Elde edilen sonuçlar aşağıdaki gibi sıralanabilir:

i. M-Tahmincilerin hesaplanmasında kullanılan yeniden ağırlıklandırılmış en küçük kareler algoritmasının başlangıç tahminlerinin seçimine olan duyarlılığı çeşitli analizler yardımıyla gösterilmiştir. Veride x-y-yönlü aykırı değer olması durumunda, yeniden azalan M-tahminciler için IRWLS algoritmasında başlangıç değerlerinin seçiminin önemli olduğu görülmüştür. Yeniden azalan M-tahmincileri için IRWLS algoritmasının başlangıç tahminlerinin seçimine robust yani sağlam olmadığı analizler yardımıyla gösterilmiştir. Veride yalnızca y-yönlü aykırı değer olması durumunda, yeniden azalan M-tahminciler için IRWLS algoritmasında başlangıç değerlerinin seçiminden daha az etkilendiği görülmüştür. Özellikle veride aykırı değer bulunduğunda yüksek kırılma noktasına sahip tahmincileri başlangıç tahmincisi olarak ele almanın gerekliliği vurgulanmıştır.

ii. Yapılan çeşitli analizler yardımıyla, M-tahmincilerinin kırılma noktası grafikleri incelenmiş ve yeniden azalan M-tahmincilerin kırılma noktalarının monotonlara göre yüksek olduğu sonucu görülmüştür.

iii. Hata teriminin dağılımının normal ve normalden farklı olduğu durumlar için M-tahmincilerinin etkinlik açısından performansı incelenmiş ve hata terimi normalden farklı olduğunda bu tezde ele alınan tüm M-tahmincilerin OLS tahmincisine göre daha düşük RMSE değerleri ortaya çıkardığı sonucuna ulaşılmıştır.

iv. Başlangıç ölçek tahmincisinin, regresyon eğiminin tahminin etkinliğine olan katkısı araştırılmıştır. Başlangıç ölçek tahmini Qn

olarak alındığında regresyon eğiminin tahminin MSE değerlerini küçülttüğü görülmüştür.

v. Ayrıca reel yaşamdan alınan iki örnek üzerinde M-tahminciler uygulanmış ve sonuçları tartışılmıştır.

KAYNAKLAR

Ali, A., Qadir, M.F. (2005), "A modified m-estimator for the detection of outliers", *Pakistan Journal of Statistics and Operation Research*, **1**, 49-64.

Ali, A., Qadir, M.F., Salahuddin, (2006), "Regression outliers: new m-class ψ-functions based on winsor's principle with improved asymptotic efficiency", *Journal of Statistics*, **13**, 67-83.

Allende, H., Frery, A.C., Galbiati, J. ve Pizarro L., (2006), "M-estimators with asymmetric influence functions: the G0A distribution case", *Journal of Statistical Computation and Simulation*, **76**, 941–956.

Alma, Ö.G., (2011), "Comparison of robust regression methods in linear regression", *International Journal of Contemporary Mathematical Sciences*, **6**, 409 – 421.

Andrews, D.F. (1974), "A robust method for multiple linear regression", *Technometrics*, **16**, 523-531

Barrodale, I. ve Roberts, F.D.K., (1973), "An improved algorithm for discrete L1 approximation", *SIAM J. Numerical Analysis*, **10**, 839-848.

Baryamureeba, V. ve Steihaug, T., (2007), "On the properties of preconditioners for. robust linear regression", *International Journal of Computing and ICT Research*, **1**, 50-66.

Bassett, G. Jr. ve Koenker, R., (1978), "Asymptotic theory of least absolute error regression", *Journal of the American Statistical Association*, **73**, 618-622.

Beaton, A.E. ve Tukey, J.W., (1974), "The firing of power series, meaning polynomials, illustrated on band spectroscopic data", *Technometrics*, **16**, 147-185.

Belsley, D.A., Kuh, E., Welsch, R.E., (1980), *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, Wiley Series, A.B.D..

Birkes, D., Dodge, Y., (1990), *Alternative Methods of Regression*, Wiley Series, New York, A.B.D..

Black, M.J., Rangaranjan, A., (1996), "On the unification of line processes, outlier rejection, and robust statistics with applications in early vision", *International Journal of Computer Vision*, **19**, 57-91.

Boyer, B.H., McDonald, J.B., Newey, W.K., (2003), "A comparison of partially adaptive and reweighted least squares estimation", *Econometric Reviews*, **22**, pp. 115-134.

Can Mutan, O., (2004), *Comparison of regression techniques via monte carlo simulation*, Yüksek Lisans Tezi, Orta Doğu Teknik Üniversitesi, Fen Blimleri Enstitüsü, Ankara.

Chen, J.H., Chen, C.S. ve Chen, Y.S., (2003), "Fast algorithm for robust template matching with m-estimators", *Ieee Transactions On Sıgnal Processing*, **51**, 230–243.

Chen, C. ve Yin, G., (2002), "Computing the efficiency and tuning constants for m-estimation", *Proceedings of the 2002 Joint Statistical Meetings*, 478–482.

Chen, K., Ying, Z., Zhang, H., Zhao, L., (2008), "Analysis of least absolute deviation", *Biometrika*, **95**, 107-122.

Clint W. Coakley ve Thomas P. Hettamansperger, (1993), "A bounded influence, high breakdown, Efficient Regression Estimator", *Journal of the American Statistical Association*, **88**, 872-880.

Dasiou, D., Angelis L. ve Moyssiadis, C., (1999), "Alternative m-estimators of location and their linear convex combinations", *Communications in Statistics – Theory and Methods*, **28**, 19-33

Dennis, J. E., Welsch, R. E., (1976), "Techniques for nonlinear least squares and robust regression", *Proceedings of the Statistical Computing Section American Statistical Association*, Washington, D. C, A.B.D., 83-87.

Dufrenois, F., Colliez, J., Hamad, D., (2007), "Crisp weighted support vector regression for robust single model estimation: application to object tracking in image sequences", *International Joint Conference on Neural Networks*, 586 – 591.

Fair, R.C., (1974), "On the robust estimation of econometric models", *Annals of Economic and Social Measurement*, **3,** 667-678.

Filzmoser, P., Serneels, S., Maronna, R. ve Van Espen, P.J., (2009), "Robust multivariate methods in chemometrics", *Comprehensive Chemometrics*, 681-722.

Geman S. ve McClure, D.E., (1987), "Statistical methods for tomographic image reconstruction", *Bulletin of the ISI*, **52**, 5-21.

Gökalp, E., Boz, Y., (2005), "Robust m-kestirimlerin gps ağlarındaki uyuşumsuz baz vektörlerini belirlemede karşılaştırılması", *Harita Dergisi*, **134**, 1-17.

Gökalp, E., Güngör, O. ve Boz,Y., (2008) "Evaluation of different outlier detection methods for GPS networks.", *Sensors*, **8**, 7344-7358.

Green, P. G., (1984), "Iteratively rewighted least squares for maximum likelihood estimation and some robust and resistant alternatives (with discussion)", *Journal of the Royal Statistical Society*, **46**, 149-192.

Gujarati, D., (2001), *Temel Ekonometri*, Çeviren: Ümit Şenesen, Literatür Yayıncılık, İstanbul

Hampel, F.R., (1974), "The influence curve and its role in robust estimation", *Journal of the American Statistical Association*, **69**, 383-393.

Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J. ve Stahel, W.A., (1986), *Robust Statistics: The Approach Based on Influence Functions*, Wiley Series, New York, A.B.D..

Hampel, F.R., Rousseeuw, P.J. ve Ronchetti, E., (1981), "The change of variance curve and optimal redescending m-estimators", *Journal of the American Statistical Association*, **76**, 643-648.

Heiberger, R.M. ve Becker R.A., (1992), "Design of an S function for robust regression using iteratively reweighted least squares", *Journal of Computational and Graphical Statistics*, **1**, 181–196.

Hekimoglu, S., Berber, M., (2003), "Effectiveness of robust methods in heterogeneous linear models", *Journal of Geodesy*, **76**, 706-713.

Heritier, S., Cantoni, E., Copt, S., Victoria-Feser, M.P., (2009), *Robust Methods in Biostatistics*, John Wiley & Sons, Ltd, Chichester, İngiltere.

Hinich, M.J. ve Talwar, P.P., (1975), "A simple method for robust regression", *Journal of the American Statistical Association*, **70**, 113-119.

Hsieh, D.A. ve Manski, C.F., (1987), "Monte carlo evidence on adaptive maximum likelihood estimation of a regression", *The Annals of Statistics*, **15**, 541-551.

Holland, P.W., Welsch, R.E., (1977), "Robust regression using iteratively reweighted least-squares", *Communications in Statistics - Theory and Methods*, **6**, 813-827.

Huber, P.J., (1964), "Robust estimation of a location parameter", *The Annals of Mathematical Statistics*, **35**, 73-101.

Huber, P.J., Ronchetti E.M., (2009), *Robust Statistics (second edition)*, Wiley Series, Hoboken, NJ, A.B.D..

Hubert, M., (2009), *LIBRA: a MATLAB library for robust analysis*, http://wis.kuleuven.be/stat/robust/Libra.html

Hubert, M.,Rousseeuw, P.J. ve Aelst, S.V., (2008), "High breakdown robust multivariate methods", *Statistical Science*, **23**, 92–119.

Kantar, Y.M., Usta, I., Acitas, S., (2011), "A monte carlo simulation study on partially adaptive estimators of linear regression models", *Journal of Applied Statistics*, **38**, 1681-1699.

Kelly, G., (1996), "Adaptive choice of tuning constant for robust regression estimators", *The Statistician*, **45**, 35-40.

Knight, N.L. ve Wang, J., (2009), "A comparison of outlier detection procedures and robust estimation methods in GPS positioning", *Journal of Navigation*, **62**, 699-709.

Lawson, C., Keats, J.B., Montgomery D.C., (1997), "Comparison of robust and least-squares regression in computer-generated probability plots", *IEEE Transactions on Reliability*, **46**, 108-115.

Lu, S., (2004), "An experiment with experimental PROC ROBUSTREG", *Proceedings of the 2004 Meeting of the Pharmaceutical Industry SAS Users Group*, A.B.D..

Rohan, M., (2011), *Using finite mixtures to robustify statistical models*, Doktora Tezi, University of Waikato, Hamilton, Yeni Zelanda.

Maronna R., Martin D., Yohai V., (2006), *Robust Statistics Theory and Methods*, John Wiley, A.B.D..

McDonald, J.B., ve White, S.B., (1993), "A comparison of some robust, adaptive and partially adaptive estimators of regression models", *Econometric Reviews*, **12**, 103–124.

Mirza, M.J., ve Boyer, K.L., (1993), "Performance evaluation of a class of M-estimators for surface parameter estimation in noisy range data", *IEEE Transactions on Robotics and Automation*, **9**, 75-85.

Moberg, T.F., Ramberg, J.S., Randles H.R., (1978), "An adaptive m-estimator and its application to a selection problem", *Technometrics*, **20**, 255-263.

Moller S.F., Von Frese J., Bro R., (2005), "Robust methods for multivariate data analysis", *Journal of Chemometrics*, **19**, 549–563.

Montgomery, D.C., Peck, E.A., Vining, G.G., (2001), *Introduction To Linear Regression Analysis (Third Edition)*, John Wiley&Sons, New York, A.B.D..

Muthukrishnan, R., Radha, M., (2010), "M-Estimators in regression models", *Journal of Mathematics Research*, **2**, 23-27.

Nevitt, J., Tam, H.P., (1989), "A comparison of robust and nonparametric estimators under the simple linear regression model", *Multiple Linear Regression Viewpoints*, **25**, 54–69.

O'Leary, D., (1990), "Robust regression computation using iteratively reweighted least squares", *SIAM Journal on Matrix Analysis and Applications*, **11**, 466-480.

Olive, D.j., (2008), *Applied Robust Statistics*, Southern Illinois University, A.B.D..

Pires, R.C., Costa, A.S., Mili, L., (1999), "Iteratively reweighted least-squares state estimation through givens rotations", *IEEE Transactions on Power Systems*, **14**, 1499-1506.

Ramirez, O.A., ve Misra, S.K., Nelson, J., (2003), "Efficient estimation of agricultural time series models with nonnormal dependent variables", *American Journal of Agricultural Economics*, **85**, 1029-1040.

Ramsay, J.O., (1977), "A comparative study of several robust estimates of slope, intercept and scale in linear regression", *Journal of the American Statistical Association*, **72**, 608-615.

Randal, J.A., (2008), "A reinvestigation of robust scale estimation in finite samples", *Computational Statistics & Data Analysis*, **52**, 5014-5021

Rousseeuw, P.J., Croux, C., (1992), "Time-efficient algorithms for two highly robust estimators of scale", *Computational Statistics*, **1**, 411-428.

Rousseeuw, P.J., Croux, C., (1993), "Alternatives to the median absolute deviation", *Journal of the American Statistical Association*, **88**, 1273-1283.

Rousseeuw, P.J., ve Van Driessen, K., (1999), "Computing LTS regression for large data sets", *Data Mining and Knowledge Discovery*, **12**, 29-45.

Rousseeuw, P.J., ve Leroy, A.M., (1987), *Robust Regression and Outlier Detection*, John Wiley, New York, A.B.D..

Rousseuw, P.J., Leroy, A., Daniels, B., (1984), "Resistant line fitting in actuarial science", *Advanced Science Institutes Series C*, **121**, 315-332.

Rousseeuw, P.J. ve Van Driessen, K., (2002), "Computing LTS regression for large data sets", *Estadistica*, **54**, 163-190.

Sarkar, S.K., Midi, H. ve Rana, S., (2011), "Detection of outliers and influential observations in binary logistic regression: an empirical study", *Journal of Applied Sciences* **11**, 26-35.

Serneels, S., Croux, C., Filzmoser, P., Van Espen, P.J., (2005), "Partial robust m-regression", *Chemometrics and Intelligent Laboratory Systems*, **79**, 55–64.

Shevlyakov, G.,Morgenthaler, S., Shurygin, A., (2008), "Redescending m-estimators", *Journal of Statistical Planning and Inference*, **138**, 2906-2917.

Siegel, A.F., (1982), "Robust regression using repeated medians", *Biometrika*, **69**, 242-244.

Stahel, W., (1996), "Robust alternatives to least squares", *ETH Seminar fur Statistik*, İsviçre.

Street, J.O.,Carroll, R.J., Ruppert, D., (1988), "A note on computing robust regression estimates via iteratively reweighted least squares", *The American Statistician*, **42**, 152-154.

Türkay, H., (2004), "Doğrusal regresyon analizinde m tahminciler ve ekonometrik bir uygulama", *Fırat Üniversitesi DAUM Dergisi*, **3**, 106-115.

Ullah, I., Qadir, M.F., Ali, A., (2006), "Insha's redescending m-estimator for robust regression: a comparative study", *Pakistan Journal of Statistics and Operation Research*, **2**, 135-144 .

Usta I., Kantar Y.M., (2011), "On the performance of the flexible maximum entropy distributions within partially adaptive estimation", *Computational Statistics and Data Analysis*, **55**, 2172–2182.

Verardi, V. ve Croux, C., (2009), "Robust regression in Stata", *Stata Journal*, **9**, 439-453.

Wadsworth, H.M., (1998), *Handbook of Statistical Methods for Engineers and Scientists (Second Edition)*, McGraw-Hill, New York, A.B.D..

Wang, J.L., (1999), "Asymptotic properties of M-estimators based on estimating equations and censored data", *Scandinavian Journal of Statistics*, **26**, 297–318.

Wang, J., Wang, J., (2007), "Mitigating the effect of multiple outliers on gnss navigation with m-estimation schemes", *International Global Navigation Satellite Systems Society - IGNSS Symposium*, Sydney, Avustralya, 1-9.

Western, B., (1995), "Concepts and suggestions for robust regression analysis", *Midwest Political Science Association*, **39**, pp. 786-817.

Wilcox, R.R., (2005), *Introduction to Robust Estimation and Hypothesis Testing (Second Edition)*, Elsevier Academic Press, A.B.D..

Wilcox, R.R., (2010), *Fundamentals of Modern Statistical Methods: Substantially Improving Power and Accuracy*, Springer, New York, A.B.D..

Wisnowski, J.W., Montgomery, D.C., Simpson J.R., (2001), "A comparative analysis of multiple outlier detection procedures in the linear regression model", *Computational Statistics & Data Analysis*, **36**, 351-382.

Wu, L.L., (1985), "Robust m-estimation of location and regression", *Sociological Methodology*, **15**, 316-388.