YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

HATA PAYLARININ NORMAL DAĞILMAMASI DURUMUNDA TEK DEĞİŞKENLİ KLASİK REGRESYONDA ALTERNATİF TAHMİN YÖNTEMLERİ

İstatistikçi Sevil TEMİZ

FBE İstatistik Anabilim Dalında Hazırlanan

YÜKSEK LİSANS TEZİ

Tez Danışmanı: Prof. Dr. Mehmet GENCELİ

İSTANBUL,2006

İÇİNDEKİLER

Sayfa

SİMGE LİSTESİ iv

KISALTMA LİSTESİ v

ŞEKİL LİSTESİ vi

ÇİZELGE LİSTESİ vii

ÖZET viii

ABSTRACT ix

1. GİRİŞ 1

2. HATA PAYLARININ NORMAL DAĞILMAMASI DURUMUNDA TEK DEĞİŞKENLİ KLASİK REGRESYONDA ALTERNATİF TAHMİN YÖNTEMLERİ 2

2.1 Hata Paylarının Normal Dağılmaması 2
2.2 Robust Regresyon Yöntemlerinin Gerekliliği 3

3. EN KÜÇÜK MUTLAK SAPMALAR YÖNTEMİ (L1 TİPİ TAHMİN EDİCİLER) 4

3.1 Regresyon Doğrusunun (L1 tipi tahmin edicilerle) Tahmin Edilmesi 4
3.1.1 Algoritma 5
3.1.2 β **İçin Hipotez Testi** 8
3.1.3 τ **Parametresi** 9
3.1.4 Yöntemin Kısıtları 12
3.2 En Küçük Mutlak Sapmalar Yöntemiyle Uygulama 13
3.2.1 Veri Kümesinin Normal Dağılıma Uygunluğunun İncelenmesi 16
3.2.2 Regresyon Modelinin En Küçük Mutlak Sapmalar Yöntemi ile Bulunması 18
3.2.3 β**'nın Testi** 20

4. M-TİPİ TAHMİN YÖNTEMLERİ (M TİPİ TAHMİN EDİCİLER) 22

4.1 M- Tahmin Edicisinin L1-Tahmin Edicisinden Farkı 22
4.2 Regresyon Doğrusunun (Huber-M tipi tahmin edicilerle) Tahmin Edilmesi 23
4.2.1 $\rho(e)$ **Fonksiyonu** 24
4.2.2 Etki Fonksiyonu 25
4.2.3 Tartı Fonksiyonu 26
4.3 Algoritma 27
4.4 β **İçin Hipotez Testi** 28
4.5 M-Tahmin Edicilerine Farklı Yaklaşımlar 28
4.5.1 Tukey'in Yaklaşımı 28

4.5.2 Hampel'in Yaklaşımı 30
4.5.3 Andrews'in Yaklaşımı 32
4.6 M-Tipi Tahmin Yöntemleriyle Uygulama 33
4.6.1 Huber M-tahmin edicileri 33
4.6.2 Tukey'in M-tahmin edicileri 39
4.6.3 Hampel'ın M-tahmin edicileri 40
4.7 Andrews'in M-tahmin edicileri 41

5. SONUÇ 43

6. KAYNAKLAR 45

7. EKLER 47

Ek 1 İki Noktadan Geçen Doğrunun Eğimi Formülü Kullanılarak Türkiye noktasının diğer noktalar ile bir doğru oluşturması durumunda oluşacak eğimlerin hesaplanması 48
Ek 2 Kalıntıların Hesaplanması 50
Ek 3 Mutlak Sapmaların Hesaplanması 52
Ek 4 Daraltılmış sapmaların hesaplanması ve bu sapmalardan yola çıkılarak öngörülen y_i^* değerleri 54
Ek 5 Öngörülen y Değerleri ve $\hat{e}_i$ Kalıntıları 56
Ek 6 $\rho(ei)$ Değerleri ve daraltılmış kalıntıların karelerinin hesaplanması 58
Ek 7 Daraltılmış hata payları ve daraltılmış kalıntılar kullanılarak öngörülen y* değerleri 60
Ek 8 Daraltılmış kalıntılar ve $\rho(ei)$ değerleri 62
Ek 9 R Project'te Komut Yazımı 64

8. ÖZGEÇMİŞ 65

SİMGE LİSTESİ

L1	**En Küçük Mutlak Sapmalar**
L2	**En Küçük Kareler**
α	**Regresyon sabiti**
β	**Regresyon katsayısı**
$\sqrt{\beta_1}$	**Çarpıklık katsayısı**
β_2	**Basıklık katsayısı**
e	**Hata payı**
$\bar{e}$	**Hata paylarının ortalaması**
$\tilde{e}$	**Hata paylarının medyanı**
m	**Sıfırdan farklı kalıntı sayısı**
$\hat{e}_1$	**Küçükten büyüğe sıralanmış kalıntıların en küçüğü**
$\hat{e}_m$	**Küçükten büyüğe sıralanmış kalıntıların en büyüğü**
σ	**Rasgele hata paylarının büyüklüklerinin ölçüsü**
τ	**Rasgele hata paylarının büyüklüklerinin ölçüsü**
$\mathbf{s_e}$	**Katsayıların standart hatası**
$\mathbf{S_e}$	**Kalıntıların standart hatası**
$\mathbf{R^2}$	**Belirginlik Katsayısı**
$\rho(e)$	**e'nin bir fonksiyonu**
$\Psi(e)$	**Etki fonksiyonu**
$w(e)$	**Tartı fonksiyonu**
$\mathbf{e_i^*}$	**e_i^0'ın daraltılmasıyla elde edilen sapma**

KISALTMA LİSTESİ

E.K.K.	**En Küçük Kareler**
E.K.M.S.	**En Küçük Mutlak Sapmalar**
L.A.D.	**Least Absolute Deviation**
M.S.M.	**Mutlak Sapmaların Medyanı**
M.A.D.	**Mean Absolute Deviation**

ŞEKİL LİSTESİ

Şekil 3.1 Ülkelere göre "Beş Yaş Altı Ölüm Oranı" ve "Doğuşta Beklenen Ömür Yılı" 15
Şekil 3.2 Veri kümesinden geçen En Küçük Kareler Regresyon Doğrusu 15
Şekil 3.3 Kalıntılar 16
Şekil 4.1 Huber'in $\rho(e)$ fonksiyonunun grafiği 25
Şekil 4.2 Huber'in $\Psi(e)$ fonksiyonunun grafiği 26
Şekil 4.3 Huber'in $w(e)$ fonksiyonunun grafiği 26
Şekil 4.4 Tukey'in $\rho(e)$ fonksiyonunun grafiği 29
Şekil 4.5 Tukey'in $\Psi(e)$ fonksiyonunun grafiği 29
Şekil 4.6 Tukey'in $w(e)$ fonksiyonunun grafiği 30
Şekil 4.7 Hampel'in $\rho(e)$ fonksiyonunun grafiği 31
Şekil 4.8 Hampel'in $\Psi(e)$ fonksiyonunun grafiği 31
Şekil 4.9 Hampel'in $w(e)$ fonksiyonunun grafiği 32
Şekil 4.10 Andrews'in $\rho(e)$ fonksiyonunun grafiği 32
Şekil 4.11 Andrews'in $\Psi(e)$ fonksiyonunun grafiği 33
Şekil 4.12 Andrews'in $w(e)$ fonksiyonunun grafiği 33

ÇİZELGE LİSTESİ

Çizelge 3.1 Ülkelere göre "Beş Yaş Altı Ölüm Oranı" ve"Doğuşta Beklenen Ömür Yılı"13
Çizelge 4.1 İterasyonlar sonucunda bulunan $\bar{y}$ değerleri 36
Çizelge 4.2 R Project kullanılarak elde edilen Huber-M tamin edicileri 38
Çizelge 4.3 R Project kullanılarak elde edilen Tukey-M tahmin edicileri 39
Çizelge 4.4 R Project kullanılarak elde edilen Hampel-M tamin edicileri 40
Çizelge 4.5 R Project kullanılarak elde edilen Andrews-M tamin edicileri 41
Çizelge Ek 1.1. Türkiye noktasının diğer noktalar ile bir doğru oluşturması durumunda oluşacak eğimlerin hesaplanması 48
Çizelge Ek 2.1 Kalıntılar 50
Çizelge Ek 3.1 Mutlak Sapmalar 52
Çizelge Ek 4.1 Daraltılmış sapmalar ve Öngörülen y_i^* değerleri 54
Çizelge Ek 5.1 Öngörülen y Değerleri ve $\hat{e}_i$ Kalıntıları 56
Çizelge Ek 6.1 $\rho(ei)$ Değerleri ve daraltılmış kalıntıların kareleri 58
Çizelge Ek 7.1 Daraltılmış kalıntılar ve öngörülen y* değerleri 60
Çizelge Ek 8.1 Daraltılmış kalıntılar ve $\rho(ei)$ değerleri 62

ÖZET

Regresyon parametreleri tahmin yöntemleri içinde en çok bilinen ve kullanılan yöntem olan "En Küçük Kareler" yöntemi, tahmin edicilerin ilgili parametreleri için, "hipotez testleri yapabilmek ve güven aralığını belirleyebilmek" için hata paylarının normal dağılıma uymasını öngörmektedir.

Hata paylarının normal dağılmaması durumunda, hipotez testleri sonuçları ve bulunan güven aralıkları hatalı sonuçlar vermektedir. Ayrıca, veri kümesinde sapan değerlerin olması durumunda da, En Küçük Kareler tahmin edicileri ve hipotez testlerinin sonuçları etkilenmektedirler. Bu gibi durumlarda, sapan değerlerin etkisini azaltmak için daha robust olan, yani, varsayımlardan sapmalardan bağıl olarak daha az etkilenen, alternatif tahmin ediciler kullanılabilir. Kısacası, Robust regresyon metodolojisi, En Küçük Kareler yönteminin güvenilir sonuçlar vermediği durumlarda daha kesin sonuçlar elde edebilmek amacıyla ortaya konulmuş bir metodolojidir.

Çalışmada, Robust Regresyon yöntemlerine başvurma koşulları incelendikten sonra L1 tipi regresyon olarak da adlandırılan En Küçük Mutlak Sapmalar Yöntemi ve Huber'in M-Regresyon Yöntemi incelendikten sonra M-tipi tahmin edici için Tukey, Hampel ve Andrews yaklaşımları irdelenmiştir. Daha sonra da konular için uygulamaya yer verilmiştir.

Anahtar kelimeler: Robust Regresyon, En Küçük Mutlak Sapmalar Regresyonu, M-Regresyon, hataların normal dağılmaması.

ABSTRACT

Least Square Method is the most known and used method in the regression parameters estimate methods. Least Square Method require normaly distribute error to get the hypothesis tests and define the confidence intervals for estimators.

When the distribution of the errors don't satisfy the normal distribution, Least Square estimators and hypothesis tests give wrong results.

Beside this, when the data contain significant outliers Least Square estimators and hypothesis tests are influenced. In these situations, to reduce the effect of the outliers, more robust estimators which are influenced less from the deviations from the assumptions, can be used. Shortly, Robust Regression Methodology is the methodology which is offered to get more certain results, when Least Square Method don't give reliable results.

In this work, after the situations to use Robust Regression Methods are examined, Least Absolute Deviation Method which is named L1, and Huber's M-Regression Method are examined and Tukey's, Hampel's and Andrews's M-type estimators approaches are discussed. And an application is given for these methods.

Keywords: Robust Regression, Least Absolute Deviation, Median Absolute Deviation, M- type estimators

1. GİRİŞ

İstatistiksel analizin gittikçe önem kazandığı günümüzde, En Küçük Kareler Yöntemi, Regresyon Parametreleri Tahmin Yöntemleri içinde en bilinen ve en çok kullanılan yöntemdir. Fakat bir veri kümesinin analizi esnasında, hata paylarının normal dağılmaması veya veri kümesinin sapan değerler içermesi durumunda, sapan değerleri veri kümesinden dışlamak ya da sapan değerleri olduğu gibi alarak En Küçük Kareler Yöntemini uygulamak doğru sonuçlar vermeyebilir. Bu durumda, "sapan değerlerin etkisini azaltacak", başka bir deyişle "varsayımlardan sapmalara duyarlı olmayan" regresyon yöntemlerini kullanmak daha güvenilir sonuçlar verecektir. Tezde, hata paylarının normal dağılmaması durumunda En Küçük Kareler Yöntemine alternatif olarak kullanılan L1 ve M tipi regresyon parametrelerini tahmin yöntemleri incelenmiştir.

2. HATA PAYLARININ NORMAL DAĞILMAMASI DURUMUNDA TEK DEĞİŞKENLİ KLASİK REGRESYONDA ALTERNATİF TAHMİN YÖNTEMLERİ

2.1 Hata Paylarının Normal Dağılmaması

En Küçük Kareler yöntemi, hesaplanmasındaki kolaylık ve belirli varsayımlara sahip olması nedeniyle en yaygın olan ve en çok kullanılan regresyon parametreleri tahmin yöntemidir, fakat En Küçük Kareler yöntemi, doğrusal lineer regresyon problemlerinde kullanılan tek yöntem değildir.

En Küçük Kareler yöntemiyle elde edilen tahmin edicilerin ilgili parametreleri için, "hipotez testleri yapabilmek ve güven aralığını belirleyebilmek" hata paylarının normal dağılıma uymasını öngörmektedir. Bir istatistik işlemi, istatistiksel modelin varsayımları geçerli olmadığında yani varsayımlardan sapmalar olması halinde de geçerliliğini muhafaza ediyorsa "robust" olarak nitelendirilmektedir.[1] Hata paylarının normal dağılmaması durumunda, hipotez testleri sonuçları ve bulunan güven aralıkları hatalı sonuçlar vermektedir. Ayrıca, veri kümesinde sapan değerlerin olması durumunda da, En Küçük Kareler tahmin edicileri ve hipotez testlerinin sonuçları etkilenmektedirler. Bu gibi durumlarda, sapan değerlerin etkisini azaltmak için daha robust olan, yani, varsayımlardan sapmalardan bağıl olarak daha az etkilenen, alternatif tahmin ediciler kullanılabilir. Kısacası, Robust regresyon metodolojisi, En Küçük Kareler yönteminin güvenilir sonuçlar vermediği durumlarda daha güvenilir sonuçlar elde edebilmek amacıyla ortaya konulmuş bir metodolojidir.[2]

İstatistikte "robust" kelimesi "varsayımlardan sapmalardan fazla etkilenmeyen" anlamında kullanılmaktadır. Bir istatistiksel işlemin iyi performanslı olarak ifade edilmesi, varsayımlardan sapma olsa bile bu işlemin doğru sonuç vermesine bağlıdır. Örnek olarak $y_1,\ldots,y_n$ örnekleminin ortalamasının tahmin edilmesi verilebilir. y_i değerlerinden birinin aşırı olması ortalamayı değiştireceğinden aritmetik ortalama ($\bar{y}$) robust değildir. Fakat y_i örnek medyanı aşırı değerlerden etkilenmediği için robusttur. [3]

---

[1] Huber, P. J. , (1981), "Robust Statistics", New York: John Wiley & Sons,Inc.

[2-3] Weisberg, S. , (1985), "Applied Linear Regression", New York: Wiley

2.2 Robust Regresyon Yöntemlerinin Gerekliliği

Veri kümesinin hata paylarının normal dağılıma uyması durumunda, En Küçük Kareler tahmin edicileri, veri kümesine en uygun regresyon modelinin tahmin edicilerini vermektedirler; fakat veri kümesinde sapan değerlerin olması ya da hata paylarının dağılımının normal dağılıma uymaması durumunda En Küçük Kareler tahmin edicileri için yapılan hipotez testleri, varsayımlardan sapmalardan etkilendikleri için "robust" olmamaktadırlar.

Bir yaklaşım, sapan değerlerin veri kümesinden çıkarılarak En Küçük Kareler Yöntemi kullanılmaya devam edilmesi yönündedir; fakat bu uygulama da (özellikle birden fazla sapan değer olması durumunda) bilgi kaybına sebep olmasından dolayı, hatalı sonuçlar vermektedir.

Hata paylarının normal dağılmadığı veri setlerinde En Küçük Kareler Yönteminin yetersiz kalması, daha kesin sonuçlar verebilecek alternatif regresyon yöntemlerine olan ihtiyacı doğurmuştur.

Robust tahmin ediciler üzerine çalışmalar 1757'de Roger Joseph Boscovich tarafından En Küçük Mutlak Sapmalar (L1 tipi veya LAD) Regresyonunun bulunmasıyla başlamıştır. Fakat hesaplanmasının çok uzun ve karmaşık olması nedeniyle çok fazla kullanılamamıştır, onun yerine hesaplanması daha kısa ve kolay olan En Küçük Kareler Yöntemi tercih edilmiştir.[4] Fakat bilgisayar programcılığının, dolayısıyla bilgisayar programlarının gelişmesiyle, robust regresyon üzerine çalışmalar tekrar başlamıştır. 1960'da Tukey, 1964'de Huber robust regresyon üzerine çalışmalar yapmış,1972 ve 1973 yıllarında teorik çalışmalarını gözden geçiren Huber'i, 1973-1978 yılları arasındaki çalışmalarıyla Hampel izlemiştir. [5]

---

[4] Birkes, D. ve Dodge,Y., (1993), "Alternative Methods Of Regression", Wiley

[5] Neter, J., Kutner, M.H., Nachtsheim, C. J., Wasserman, W., (1993) "Applied Linear Statistical Methods", Wiley

3. EN KÜÇÜK MUTLAK SAPMALAR YÖNTEMİ (L1 TİPİ TAHMİN EDİCİLER)

Hata paylarının normal dağılmaması ve/veya veri kümesi içinde sapan değerlerin bulunması durumlarında L1 tipi tahmin yöntemi olarak da adlandırılan En Küçük Mutlak Sapmalar (Least Absolute Deviation) yöntemi diğer klasik tahmin yöntemlerine üstünlük göstermektedir.

3.1 Regresyon Doğrusunun (L1 tipi tahmin edicilerle) Tahmin Edilmesi

En Küçük Kareler Yöntemi, $\hat{\alpha}$ **ve** $\hat{\beta}$ **tahmin edicilerini, kalıntıların karelerinin toplamını** $\left(\sum \hat{e}_i^{\,2}\right)$ **minimum yapacak şekilde hesaplamaya dayanmaktadır.[6] En Küçük Mutlak Sapmalar Yöntemi ise, kalıntıların mutlak değerlerinin toplamını minimum yapma esasına dayanmaktadır. Bu durum aşağıdaki şekilde ifade edilir:**

$$\sum |\hat{e}_i| \rightarrow \text{minimum} \qquad (3.1)$$

$y_i - (\hat{\alpha} + \hat{\beta} x_i)$ **farkı,** (x_i, y_i) **noktasının** $Y = \hat{\alpha} + \hat{\beta} X$ **doğrusundan sapması olarak tanımlanmaktadır. Bu durumda** $\sum |\hat{e}_i| = \sum \left| y_i - (\hat{\alpha} + \hat{\beta} x_i) \right|$ **olarak yazılabilmektedir. Bu ifadede yer alan y ve x değerleri belli olduğuna göre, ifadenin minimum olması** $\hat{\alpha}$ **ve** $\hat{\beta}$ **değerlerinin minimum olmasına bağlıdır. Buradan hareketle En Küçük Mutlak Sapmalar Yöntemi'nin amacının,** $\hat{\alpha}$ **ve** $\hat{\beta}$**' yı,** $\sum |\hat{e}_i|$**'yi minimum yapacak şekilde tahmin etmek olduğu söylenebilmektedir.**

En Küçük Mutlak Sapmalar Yöntemi'nin ilkeleri aslında En Küçük Kareler Yöntemi'nin ilkelerinden daha basittir. Çünkü $|\hat{e}|$**'ın hesaplanması,** $\hat{e}^2$**'nin hesaplanmasından daha kolaydır. Bunun yanında En Küçük Kareler Yöntemi'nin tahmin edicilerinin hesaplanması belirli formüllere dayanırken, En Küçük Mutlak Sapmalar Yöntemi'nin tahmin edicilerinin hesaplanması için hiçbir formül bulunmamaktadır. Bu durum ise, E.K.M.S. tahmin edicilerinin hesaplanması için bir algoritmanın oluşturulmasını gerekli kılmaktadır.**

Bu algoritma, mutlak sapmaların toplamının minimum olduğu doğruyu, yani, veri

---

[6] Genceli, M. ,(2001), Ekonometri ve İstatistik İlkeleri, Filiz Kitabevi, İstanbul.

kümesine en uygun olan doğruyu bulmayı amaçlamaktadır. Örneğin y=$\hat{\alpha} + \hat{\beta}$ x doğrusu ele alındığında, verilerin (y_i), y doğrusundan mutlak sapmalarının toplamı $\left(\sum |e_i|\right)$, verilerin doğruya olan dikey uzunluklarının toplam boyutunu vermektedir. Doğru, yukarı doğru ε kadar küçük bir miktar hareket ettirilirse, bu durumda her bir mutlak sapma (noktanın doğrunun üzerinde veya altında olmasına bağlı olarak) ε kadar artı ya da eksi yönde hareket edecektir. Buradan hareketle, $\sum \left| y_i - (\hat{\alpha} + \hat{\beta} x_i) \right|$ ifadesinin minimum değerini bulmak için, veriler arasında doğruyu aşağı/yukarı hareket ettirmek gerekmektedir.

Özetle, algoritmanın temel mantığı, verilen bir (x_0,y_0) noktasından geçen tüm doğruların arasından en iyi doğruyu bulmaya dayanmaktadır. (x_0,y_0) noktasından geçen en uygun doğru bulunduktan sonra, "bir doğru iki noktadan geçer" ilkesinden yola çıkılarak bulunan bu doğrunun geçtiği diğer nokta olan (x_1,y_1) noktası bulunmaktadır. Daha sonra (x_1,y_1) noktasından geçen doğrular arasından en uygun doğru bulunmakta ve bu doğrunun da ikinci noktası yani (x_2 ,y_2) noktası bulunmaktadır. Bu şekilde devam edilerek, her aşamada toplam mutlak sapması daha küçük bir doğru bulunarak, bulunan bir doğru, bir önceki adımda bulunan doğru ile aynı çıkana kadar işlemler devam ettirilmektedir. Aynı doğru ardı ardına iki defa elde edildiğinde ise algoritmaya son verilmektedir. Bulunan bu doğru, tüm doğrular arasında toplam mutlak sapması en küçük olanıdır dolayısıyla en iyisidir ve aranan "En Küçük Mutlak Sapmalar Regresyon Doğrusu" dur. [7]

3.1.1 Algoritma

Bu algoritma, veri kümesinin, tek olmama (nonuniqueness) ve dejenere olma (degeneracy) kavramlarına dayanmadığını varsaymaktadır, çünkü, bu kavramlar teknik problemlere sebep olabilmektedir.

(x_0,y_0) noktasından geçen tüm doğrular arasından en iyi doğruyu bulmak için tanımlanan bu algoritmanın işleyişi aşağıdaki gibidir:

Başlangıçta alınan nokta (x_0,y_0) ile her noktanın eğimi ayrı ayrı hesaplanmaktadır. Bu

---

[7] **Birkes, D. ve Dodge,Y., (1993), "Alternative Methods of Regression", Wiley**

hesaplama, "iki noktadan geçen doğrunun eğimi" formülünden $\frac{(y_i - y_0)}{(x_i - x_0)}$ yararlanılarak yapılır. Hesaplamalar sırasında, x_i, x_0'a eşit olduğunda eğim tanımlanamayacağı için $x_i=x_0$ eşitliğini sağlayan noktalar işleme tabi tutulmamaktadır. Bütün eğimler hesaplandıktan sonra

$$\frac{(y_1 - y_0)}{(x_1 - x_0)} \le \frac{(y_2 - y_0)}{(x_2 - x_0)} \le \ldots \le \frac{(y_n - y_0)}{(x_n - x_0)} \qquad \textbf{(3.2)}$$

olacak şekilde artan sırada dizilmektedir.

$$T = \sum |x_i - x_0| \qquad \textbf{(3.3)}$$

olmak üzere:.

$$|x_1 - x_0| + \ldots + |x_{k-1} - x_0| < \frac{1}{2}T \qquad \textbf{(3.4)}$$

$$|x_1 - x_0| + \ldots + |x_{k-1} - x_0| + |x_k - x_0| > \frac{1}{2}T \qquad \textbf{(3.5)}$$

Yukarıda da görüldüğü gibi $|x_k - x_0|$ ifadesinin eklenmesiyle $|x_i - x_0|$'ın kümülatif toplamı T/2'yi geçmektedir. Böylece (x_0,y_0) noktasından geçen en iyi doğrunun (x_k,y_k) noktasından geçtiği bulunmuş olup bu doğrunun eğimi

$\beta^* = \frac{(y_k - y_0)}{(x_k - x_0)}$ **dir.** **(3.6)**

$\hat{Y} = \alpha^* + \beta^* \mathbf{X}$ **olduğundan, bulunan β^* yerine konularak, α^* da hesaplanmaktadır:**

$\alpha^* = y_0 - \beta^* x_0$ **eşitliğinden** **(3.7)**

$\hat{Y} = \alpha^* + \beta^* \mathbf{X}$ **eşitliği yazılabilmektedir.** **(3.8)**

Aynı işlemler (x_k,y_k) noktası için yapılarak, (x_k,y_k) noktasından geçen tüm doğrular arasında en iyi doğru bulunur. Bulunan bu ikinci doğrunun mutlak sapması bir önceki doğrunun mutlak sapmasından daha küçüktür. Bu işlemler, ard arda birbirinin aynısı iki doğru bulunana kadar devam etmektedir. Bulunan bu doğru veri kümesini en iyi temsil eden doğrudur ve "En Küçük Mutlak Sapmalar Regresyon Doğrusu" olarak tanımlanmaktadır.

- **Algoritmanın İspatı :**

(x_0,y_0) noktası için düşünüldüğünde, $y_0 = \alpha + \beta.x_0 \rightarrow \alpha = y_0 - \beta.x_0$ **olacaktır.**

Amaç, $\sum|y_i - (\alpha + \beta.x_i)|$ **'yı minimum yapan doğruyu bulmak olduğuna göre,** α **bu formülde yerine koyulursa,**

$$\sum|y_i - (\alpha + \beta.x_i)| = \sum|y_i - (y_0 - \beta.x_0 + \beta.x_i)|$$

$$= \sum|(y_i - y_0) + \beta.(x_0 - x_i)| \text{ bulunur.}$$

Sonuç olarak, $\sum|(y_i - y_0) + \beta.(x_0 - x_i)|$ **ifadesini minimum yapacak** β **değerini bulmak,** $\sum|e_i|$ **'yı minimum yapmak anlamına gelmektedir.**

$\sum|(y_i - y_0) + \beta.(x_0 - x_i)|$ **ifadesinin,** β**'nın bir fonksiyonu olduğu düşünülürse, bu fonksiyonun minimumunu bulmak için fonksiyonun** β**'ya göre kısmi türevini almak gerekir. (** $x_0=x_i$ **olduğunda türev alınamayacağı için, bu işlem sırasında** $x_0=x_i$ **olan veriler göz ardı edilecektir.)**

$\sum|(y_i - y_0) + \beta.(x_0 - x_i)|$ **fonksiyonun türevi alındığında** $\sum|(x_i - x_0)|$ **ifadesi elde edilir. Bu ifadeye T denirse:**

$$T=\sum|(x_i - x_0)| \quad \textbf{(3.9)}$$

$$|x_1 - x_0| + \ldots + |x_{k-1} - x_0| < \frac{1}{2}T \quad \textbf{(3.10)}$$

$$|x_1 - x_0| + \ldots + |x_{k-1} - x_0| + |x_k - x_0| > \frac{1}{2}T \quad \textbf{(3.11)}$$

Bu eşitsizlikler yardımı ile, (x_0,y_0) noktasından geçen, en iyi doğru olan $\hat{Y} = \alpha^* + \beta^* X$ **için** β^* **ve** α^* **değerleri bulunur:**

$$\beta^* = \frac{y_k - y_0}{x_k - x_0} \quad \textbf{(3.12)}$$

$$\alpha^* = y_0 - \beta^* x_0 \quad \textbf{(3.13)}$$

3.1.2 β İçin Hipotez Testi

β= 0 için hipotez testi yapılırken, sıfırdan farklı olan kalıntılar ($\hat{e}_i = Y_i - \hat{Y}_i$) küçükten büyüğe sıralanmaktadır. Burada $\hat{e}_1$ en küçük kalıntıyı, $\hat{e}_2$ ikinci en küçük kalıntıyı,…, $\hat{e}_m$ de en büyük kalıntıyı simgelemektedir. m ise, sıfırdan farklı kalıntı sayısını göstermektedir.

k_1 $\frac{m+1}{2} - \sqrt{m}$ 'ye en yakın, k_2 de $\frac{m+1}{2} + \sqrt{m}$ 'ye en yakın değer olmak üzere, k_1 ve k_2 bulunur. Sıralı kalıntılarda k_1. kalıntı $\hat{e}_{(k_1)}$ ve k_2. kalıntı $\hat{e}_{(k_2)}$ olmaktadır. $\hat{e}_{(k_1)}$ ve $\hat{e}_{(k_2)}$,

$$\hat{\tau} = \frac{\sqrt{m}\left[\hat{e}_{(k_2)} - \hat{e}_{(k_1)}\right]}{4} \qquad \textbf{(3.14)}$$

formülünde yerlerine konularak τ değeri tahmin edilmektedir.

$\sqrt{\sum (x_i - \bar{x})^2}$ hesaplanıp, $\hat{\tau}$ ve $\sqrt{\sum (x_i - \bar{x})^2}$ (3.15) formülünde yerine konup $\hat{\beta}$ nın standart hatası olan tahmin edilir.

$$S_{\hat{\beta}} = \frac{\hat{\tau}}{\sqrt{\sum (x_i - \bar{x})^2}} \qquad \textbf{(3.15)}$$

Bundan sonra $|t| = \frac{|\hat{\beta}|}{S_{\hat{\beta}}}$ test istatistiği hesaplanıp kritik değer ile karşılaştırılır.

- **En Küçük Kareler ile En Küçük Mutlak Sapmalar Testlerinin Hesaplanmasındaki Farklılık:**

En Küçük Mutlak Sapmalar testi ile En Küçük Kareler testi arasında benzerlik görülmektedir. Hatta $|t|$'nin hesaplanması iki yöntemde de aynıdır. Fakat $\hat{\beta}$'nın ve $S_{\hat{\beta}}$'nin hesaplanması faklılık göstermektedir. $S_{\hat{\beta}}$'nın hesaplanmasındaki tek farklılık ise En Küçük Kareler Testinde pay kısmında $\hat{\sigma}$ bulunurken, En Küçük Mutlak Sapmalar Testinde onun yerini $\hat{\tau}$ almaktadır.

- **$S_{\hat{\beta}}$ Formülü:**

Hipotez testi sonucunda β=0 olup olmadığına karar verilir. $\hat{\beta}$ tahmin edicisi, β'ya çok

yakın beklenebilir. $\hat{\beta}$ ile β arasındaki fark bir veya iki standart hatadan daha büyük olmamalıdır. $|t|$'nin tablo değerinden büyük olması, $\hat{\beta}$ ile β=0 arasındaki farkın $S_{\hat{\beta}}$ den büyük olduğu anlamına gelir, bu durumda H_0 : β=0 reddedilerek H_1 : $\beta \neq 0$ olduğu kabul edilir.

$S_{\hat{\beta}}$ formülünde görüldüğü gibi x birimlerinin geniş bir alana yayılmış olması sonucu $S_{\hat{\beta}}$ değeri küçük çıkacaktır. Geniş bir alana yayılan x birimlerinin regresyon doğrusu hakkında bize daha geniş bilgi verir. Bunun yanında $S_{\hat{\beta}}$ değerinin küçük olması $\hat{\tau}$ değerine de bağlıdır.

En Küçük Mutlak Sapmalar Yöntemi'ndeki $\hat{\tau}$ niceliği, En Küçük Kareler Yöntemi'ndeki σ ile benzer görevdedir. En Küçük Kareler Yöntemi için $\hat{\beta}$'nın standart hatası $\frac{\sigma}{\sqrt{\sum (x_i - \bar{x})^2}}$ iken; En Küçük Mutlak Sapmalar Yöntemi için $\hat{\beta}$'nın standart hatası yaklaşık $\frac{\tau}{\sqrt{\sum (x_i - \bar{x})^2}}$ tır. Örnek hacmi büyüdükçe bu değer, parametre değeri olan $\beta_{E.K.M.S.}$'a yaklaşmaktadır.

3.1.3 τ Parametresi

Regresyon doğrusunun eğiminin tahmin edilmesinde, E.K.K. ve E.K.M.S. regresyon regresyon yöntemlerinden hangisinin daha iyi sonuç verdiği $\frac{\tau}{\sigma}$ oranına bağlıdır.

σ ve τ, rasgele hata paylarının büyüklüklerinin ölçüsüdür. ϑ medyan etrafındaki hata paylarının dağılımının olasılık yoğunluk fonksiyonu olmak üzere, $\tau = \frac{1}{2\vartheta}$ 'dır. σ büyük olduğunda, hata payları geniş bir alana yayılır. Bu durumda medyan etrafındaki olasılık yoğunluk küçüktür, bu sebeple ϑ da küçüktür. Yukarıdaki eşitlikten de görüldüğü üzere, ϑ'nın küçük olması ise, τ 'nin büyük olması anlamına gelmektedir.

Özetle;

σ büyük olduğunda τ da büyüktür ve σ küçük olduğunda τ da küçük olmaktadır. Fakat $\frac{\tau}{\sigma}$'nın gerçek oranı, anakütle hata paylarının dağılımının şekline bağlıdır. Hata

payları normal dağılıyorsa $\frac{\tau}{\sigma}$ > 1 olmaktadır. Bu durumda, en azından büyük örnekler için, En Küçük Kareler Yöntemi, En Küçük Mutlak Sapmalar Yönteminden daha iyi sonuç vermektedir. Hata paylarının normal dağılmaması halinde ise, $\frac{\tau}{\sigma}$ <1 olmaktadır.[8]

- **$\hat{\tau}$ 'nın Formülünün İspatı :**

Ortaya koyulmaya çalışılan sav, yaklaşık normal dağılıma ve aynı uzunluktaki güven sınırlarına dayanan, hata payları dağılımının medyanı için %95 yaklaşık güven sınırlarının oluşturulmasıdır.

Alt sınır < τ < Üst sınır (3.16)

Buna göre;

1. **Örnek birim mevcudu fazla,**
2. **m, sıfırdan farklı kalıntı sayısını göstermek üzere; m yaklaşık olarak n'e eşit,**
3. **Kalıntılar e_i, hata payları u_i ile aynı dağılıma sahip,**
4. **k_1 ve k_2 , $\frac{n+1}{2} \pm \sqrt{n}$ 'e en yakın rakamlar kabul edildiğinde**

τ, yaklaşık olarak $\frac{\sqrt{n}\left[\hat{e}_{k_2} - \hat{e}_{k_1}\right]}{4}$ ile ikame edilebilir.

Daha önce de değinildiği gibi E.K.M.S. regresyonundaki τ, E.K.K. regresyonundaki σ ile benzer işlevi yapmaktadır. E.K.K. regresyon tahmin edicisinin standart hatası σc biçimindedir. E.K.M.S. regresyon tahmin edicisinin standart hatası ise, yaklaşık olarak τc biçimindedir. Ele alınan regresyon modelinde, p=0 olduğu yani açıklayıcı değişkenin olmadığı, sadece aynı anakütleden çekilen e_1, e_2, …, e_n şeklinde n tane bağımsız gözlem olduğu varsayılsın. Bu durumda,

- **E.K.K. regresyon tahmin edicisi, örnek ortalaması $\bar{e}$'ye eşit olur. Çünkü a=$\bar{e}$ eşitliğini sağlayan a, kareli sapmaların toplamının $\left(\sum (e_i - a)^2\right)$ en küçük değerini verir. a'nın standart sapması $\frac{\sigma}{\sqrt{n}}$ 'dir.**

---

[8] **Birkes, D. ve Dodge,Y., (1993), "Alternative Methods Of Regression", Wiley**

- **E.K.M.S. regresyon tahmin edicisi ise, örnek medyanı $\tilde{e}$'ye eşit olur. Çünkü a=$\tilde{e}$ eşitliği, mutlak sapmaların toplamının $\left(\sum|e_i - a|\right)$ en küçük değerini verir. a'nın standart sapması ise yaklaşık olarak $\frac{\tau}{\sqrt{n}}$ dir.**

Merkezi limit teoremi, büyük örnekler için örnek ortalamalarının dağılımının normal dağılıma yaklaştığını ifade etmektedir. Bunun sonucu olarak da, büyük örneklerde, örnek meydanlarının dağılımı da normal dağılıma yaklaşır. Bunun yanında örnek medyanının beklenen değeri anakütle medyanına eşit olmaktadır.

ν'nin anakütle hata paylarının medyanı olduğu varsayılsın. Büyük örnekler için, $\tilde{e}$, ν ortalamalı, $\frac{\tau}{\sqrt{n}}$ standart hatalı bir normal dağılıma yaklaşır. Buradan hareketle; $\hat{\tau}$, τ'nın tahmin edicisi olmak üzere, ν'nin %95 yaklaşık güven sınırları $\tilde{e} \pm \frac{2\hat{\tau}}{\sqrt{n}}$ şeklinde yazılır. Aslında burada ilgilenilen ne ν'dir ne de güven sınırlarıdır; sadece güven sınırlarının hacmi olan $\frac{4\hat{\tau}}{\sqrt{n}}$'dir.

$e_{(k1)}$'den $e_{(k2)}$'ye kadar olan aralık, ν için %95 yaklaşık güven sınırıyla da gösterilebilir. Bunun sonucu olarak, en azından büyük örnekler için, bu iki aralık ve aralıkların uzunlukları birbirine yaklaşık olarak eşittir:

$\frac{4\hat{\tau}}{\sqrt{n}} \approx e_{(k2)} - e_{(k1)}$

Bu formülde e'lerin yerine $\hat{e}$ler ve n'in yerine m (n-sıfırdan farklı kalıntılar) koyulduğunda $\frac{4\hat{\tau}}{\sqrt{m}} \approx \hat{e}_{(k_2)} - \hat{e}_{(k_1)}$ eşitliği bulunur.

Bu eşitlikte $\hat{\tau}$ çekildiğinde, $\hat{\tau} \approx \frac{\sqrt{m}\left[\hat{e}_{k_2} - \hat{e}_{k_1}\right]}{4}$ formülü bulunmaktadır.

Küçük örnekler için sadece "m" sıfırdan farklı kalıntıları kullanmak, performansı yükseltmektedir.

3.1.4 Yöntemin Kısıtları

(Tek Olmama Durumu (Nonuniqueness) Ve Dejenere Olma Durumu (Degeneracy)

Bir noktadan geçen birden fazla farklı iyi/uygun doğrunun olması durumu, "tek olmama durumu" (nonuniqueness) olarak tanımlanmaktadır. Bir noktadan geçen en iyi doğrunun aynı zamanda iki veya daha fazla noktadan geçmesi durumu ise "dejenere olma durumu" (degeneracy) olarak tanımlanmaktadır.

Bu tür problemler $\beta^* = \frac{y_k - y_0}{x_k - x_0}$ **eğiminin,** $\frac{y_{k-1} - y_0}{x_{k-1} - x_0}$ **eğimine veya** $\frac{y_{k+1} - y_0}{x_{k+1} - x_0}$ **eğimine eşit olması durumunda ortaya çıkmaktadırlar. Yukarıda anlatılan algoritma ile bir çok veri kümesinin En Küçük Mutlak Sapmalar Regresyon doğrusu bulunabilmektedir fakat "tek olmama" ve "dejenere olma" durumlarında iş biraz karışmaktadır. Çünkü, algoritmanın temel mantığı, verilen bir (x_0,y_0) noktasından geçen tüm doğruların arasından en iyi doğruyu ve bu doğrunun geçtiği diğer noktayı bularak, bulunan noktayı bir sonraki adımda kullanmaktır. Fakat tek olmama durumu mevcut olduğunda, veri noktasından geçen birden çok doğru var olduğu için doğrulardan hangisinin seçilip işlemlere devam edileceği sıkıntı yaratmaktadır.**

Dejenere olma durumunda ise, bulunan en iyi doğru birden çok noktadan geçtiği için, bir sonraki adımda kullanılmak üzere birden çok nokta içinden birini seçme sorunu yaşanmaktadır. Doğru seçim yapılmadığında, algoritma döngüye girebilmekte veya gerçekte "En Küçük Mutlak Sapmalar Regresyon Doğrusu" olmayan bir doğru En Küçük Mutlak Sapmalar Regresyon doğrusuymuş gibi kabul edilebilmektedir.

"Tek olmama" ve "Dejenere olma" durumlarıyla karşılaşıldığında, yukarıda verilen algoritma kullanılamamaktadır. Veri kümesinde tek olmama durumu mevcutsa, yani bir noktadan geçen birden fazla uygun doğru varsa, bu doğruların her biri için mutlak sapmaların toplamı hesaplanmakta ve en küçük mutlak sapmaya sahip olan doğru seçilmektedir. Bu algoritmanın uygulanabilirliği, n örnek boyutuna bağlıdır. Tek olmama durumunda yani birden fazla En Küçük Mutlak Sapmalar Regresyon Doğrusu bulunması durumunda, bunlardan biri keyfi olarak seçilebilmekte veya bunların ortalaması alınabilmektedir. Fakat, dejenere olma durumunda hiçbir algoritma

kullanılamamaktadır.[9]

3.2 En Küçük Mutlak Sapmalar Yöntemiyle Uygulama

Aşağıdaki tabloda, Y değişkeni "doğuşta beklenen ömür yılı"nı ifade etmektedir. X değişkeni ise "beş yaş altı ölüm oranı(binde)"nı ifade etmektedir. Araştırma yılı 1999 dur.[10]

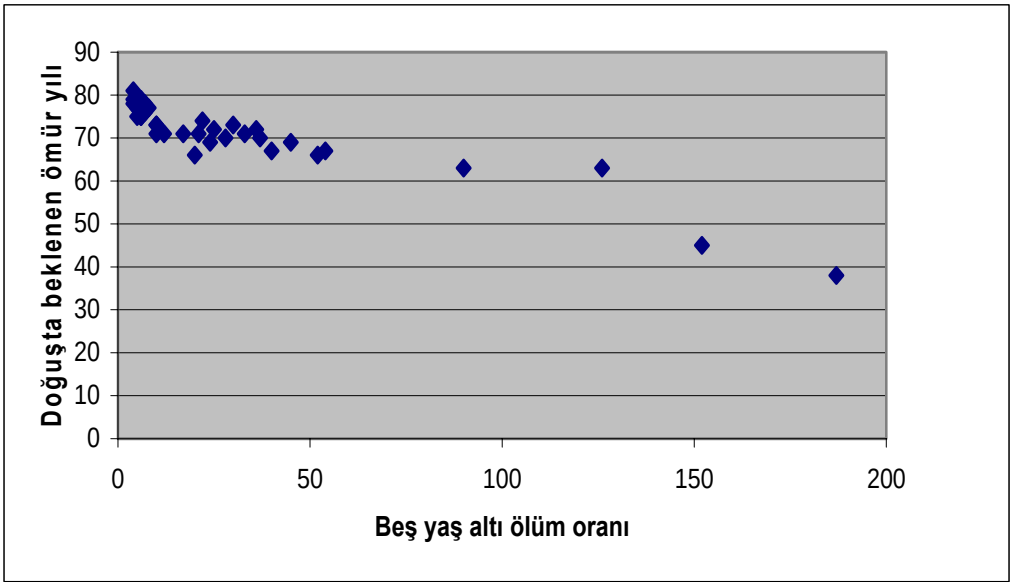
Çizelge 3.1 Ülkelere göre "Beş Yaş Altı Ölüm Oranı" ve"Doğuşta Beklenen Ömür Yılı"

Ülkeler	X	Y
Türkiye	45	69
ABD	8	77
Almanya	5	77
Arjantin	22	74
Azerbaycan	21	71
Belçika	6	78
Brezilya	40	67
Bulgaristan	17	71
Çek Cumhuriyeti	5	75
Çin	37	70
Danimarka	6	76
Endonezya	52	66
Estonya	12	71
Finlandiya	5	77
Fransa	5	79
Hindistan	90	63
İngiltere	6	77
İran	33	71

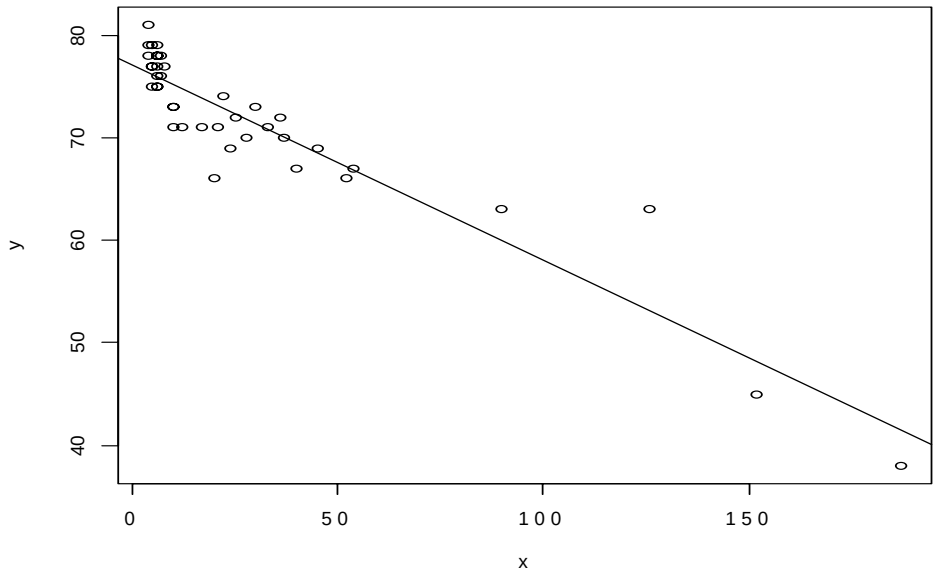
[9] Birkes, D. ve Dodge,Y., (1993), "Alternative Methods Of Regression", Wiley

[10] Kaynak: Devlet İstatistik Enstitüsü, Uluslararası Istatistikler/Temel göstergeler

İrlanda	**7**	**76**
İspanya	**6**	**78**
İsveç	**4**	**79**
İtalya	**6**	**78**
Japonya	**4**	**81**
Kanada	**6**	**79**
Kolombiya	**28**	**70**
Macaristan	**10**	**71**
Meksika	**36**	**72**
Mısır	**54**	**67**
Norveç	**4**	**78**
Pakistan	**126**	**63**
Polonya	**10**	**73**
Portekiz	**6**	**75**
Romanya	**24**	**69**
Rusya Federasyonu	**20**	**66**
Slovakya	**10**	**73**
Slovenya	**6**	**75**
Suudi Arabistan	**25**	**72**
Tanzanya	**152**	**45**
Tunus	**30**	**73**
Yunanıstan	**7**	**78**
Zambiya	**187**	**38**



Şekil 3.1 Ülkelere göre "Beş Yaş Altı Ölüm Oranı" ve "Doğuşta Beklenen Ömür Yılı"



Şekil 3.2 Veri kümesinden geçen En Küçük Kareler Regresyon Doğrusu

- **En Küçük Kareler Yöntemi uygulandığında regresyon denklemi :**

$\hat{Y}$ **= 77,17355 - 0,19114 X**

s_e : (0,568472) (0,01147)

t : (135,756) (-16,654)

R^2=0,876721607

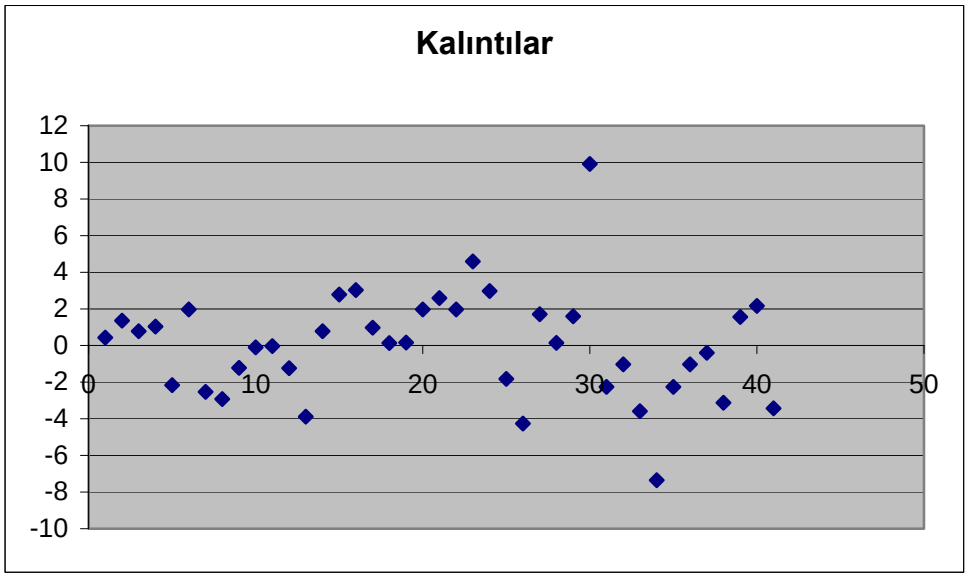
n=41

S_e=2,958625

olarak bulunmaktadır.

3.2.1 Veri Kümesinin Normal Dağılıma Uygunluğunun İncelenmesi

Kalıntıların grafiği aşağıdaki gibidir:



Şekil 3.3 Kalıntılar

H_0: Kalıntıların dağılımı normal dağılıma uygundur

Hipotezi Kolmogorov-Simirnov Normallik testi ile test edilerek, anlamlılık 0,796 > p=0,05 olduğundan dolayı H_0 kabul edilir.

Fakat Kolmogorov Simirnov ve χ^2 normallik testleri güvenirlilik anlamında zayıf testler olduğundan dolayı, çarpıklık-simetriklik katsayılarını kullanarak normalliğin tekrar ölçülmesinde fayda vardır.

Çarpıklık katsayısı simetrik durumdan uzaklaşma derecesini belirlemek için kullanılmaktadır. Bu katsayının değeri (-1;1) arasında değişmekte ve simetrik bir dağılım için sıfıra eşit olmaktadır. Çarpıklık katsayısı sağa eğik dağılımlarda pozitif, sola eğik dağılımlarda negatif değerler almaktadır.

$$\sqrt{b_1} = \frac{\sum(X_i - \bar{X})^3 / n}{\left[\sum(X_i - \bar{X})^2 / n\right]^{3/2}} \qquad \textbf{(3.17)}$$

H_0**:** $\sqrt{\beta_1}$ **=0 Dağılım simetriktir.**

H_1**:** $\sqrt{\beta_1} \neq$ **0 Dağılım simetrik değildir.**

Formülde değerler yerine koyulursa;

$\sqrt{b_1}$ **=0,485447 olarak hesaplanmaktadır. n=45 için %95 güven aralığında alt ve üst limitlerin tablo değerleri (-0,679; +0,679) dur. 0,485447<0,679 olduğu için** H_0**'ın reddi için yeterli kanıt bulunamamıştır. Dağılım simetriktir.**

Basıklık katsayısı ise dağılımın sivri ya da basıklığını belirlemektedir. Normal bir dağılımda basıklık katsayısının değeri 3'tür. Dolayısıyla, hipotez testi 3'e göre yapılır.

$$b_2 = \frac{\sum(X_i - \bar{X})^4 / n}{\left[\sum(X_i - \bar{X})^2 / n\right]^2} \qquad \textbf{(3.18)}$$

H_0**:** β_2**= 3 Dağılım basık ve ya sivri değildir.**

H_1 **:** $\beta_2 \neq$ **3 Dağılım sivri ve ya basıktır.**

Formülde değerler yerine koyulursa;

b_2**= 5,111947 olarak hesaplanır. n=45 için %95 güven aralığında alt ve üst limitlerin tablo değerleri (2,02; 4,41) dir. 5,111947>4,41 olduğu için** H_0 **reddedilir. Dağılım sivridir.**

Bu iki test sonucunda dağılımın simetrik fakat sivri olduğu görülmektedir.

Sapan değer olarak belirlenen iki nokta (Pakistan ve Rusya Federasyonu) çıkarılıp tekrar normallik testi yapılırsa :

$\sqrt{b_1}$ **= - 0,17765 olarak hesaplanır. n=45 için %95 güven aralığında alt ve üst limitlerin**

tablo değerleri (-0,679; +0,679) dur. - 0,17765 < - 0,679 olduğu için H_0'ın reddi için yeterli kanıt bulunamamıştır. Dağılım simetriktir.

b_2= 2,56892 ise olarak hesaplanır. n=45 için %95 güven aralığında alt ve üst limitlerin tablo değerleri (2,02; 4,41) dir. 2,02 < 2,56892 < 4,41 olduğu için H_0'ın reddi için yeterli kanıt bulunamamıştır. Dağılım sivri değildir.

Buradan da görüldüğü gibi, sapan değerler çıkarıldığında kalıntıların dağılımı normal dağılıma uymaktadır.

- **Sapan değerler çıkarılıp tekrar En Küçük Kareler Yöntemi uygulanıldığında:**

Pakistan (126,63) ve Rusya Federasyonu (20,66) veri kümesinden çıkarılarak tekrar En Küçük Kareler Yöntemi uygulandığında regresyon modeli :

$\hat{Y}$ =77,59229 - 0,20935 X

s_e : (0,418591) (0,009003) bulunmaktadır.

t : (185,3653) (-23,2544)

R^2=0,93596

n=39

S_e=2,144459

olarak bulunmaktadır.

3.2.2 Regresyon Modelinin En Küçük Mutlak Sapmalar Yöntemi ile Bulunması

İlk nokta Türkiye olduğu için işleme Türkiye (45,69) noktası ile başlanmıştır. Bu noktadan geçen en iyi doğruyu bulmak için bu nokta ile diğer noktalar için ayrı ayrı eğimler, iki noktadan geçen doğrunun eğimi formülü $\left(\frac{y_i - 69}{x_i - 45}\right)$ ile bulunmuştur. Hesaplanan eğimler artan sırada dizilmiştir. Her bir nokta için $|x_i - 45|$ hesaplanmıştır. Ek1'de tüm noktalar için hesaplanan eğimler ve mutlak sapmalar yer almaktadır. $T = \sum |x_i - 45| = 1444$ dir. Ve T/2=722 sayısını geçen ilk değere karşılık gelen nokta İngiltere'dir, dolayısıyla İngiltere doğrunun ikinci noktasıdır. Yani "Türkiye'den geçen doğru İngiltere'den geçmektedir". İkinci adımda, İngiltere'yi temel alıp İngiltere'den

geçen doğrunun geçtiği diğer noktanın hangi nokta olduğu bulunacaktır. İngiltere'nin koordinatları (6,77) olduğundan dolayı eğimler $\left(\frac{y_i - 77}{x_i - 6}\right)$ **bulunmuştur. Hesaplanan eğimler artan sırada dizilmiş ve her bir nokta için** $|x_i - 6|$ **hesaplanmıştır. Hesaplamalar sonucunda,** $T = \sum |x_i - 6| = 957$ **bulunmuştur. T/2=478,5 sayısını geçen ilk değere karşılık gelen nokta Zambiya olduğundan doğrunun ikinci noktası Zambiya noktasıdır. Yani İngiltere'den geçen doğru Zambiya'dan geçmektedir,denilir.**

Üçüncü adımda Zambiya temel alınarak aynı hesaplamalar yapılmıştır. Zambiya'nın koordinatları (187,38) olduğundan eğimler $\left(\frac{y_i - 38}{x_i - 187}\right)$ **formülüyle hesaplanmış, artan sırada dizilerek her bir nokta için** $|x_i - 187|$ **hesaplanmıştır.** $T = \sum |x_i - 187| = 6484$ **dür. Ve T/2=3242 sayısını geçen ilk değere karşılık gelen nokta ABD'dir ve doğrunun ikinci noktasıdır. Yani Zambiya'dan geçen doğru ABD'den geçmektedir.**

Dördüncü adımda, koordinatları (8,77) olan ABD temel alınarak $\left(\frac{y_i - 77}{x_i - 8}\right)$ **eğimleri hesaplanmış ve artan sırada dizilerek her bir nokta için** $|x_i - 8|$ **hesaplanmıştır.** $T = \sum |x_i - 8| = 939$ **dur. Ve T/2=469,5 sayısını geçen ilk değere karşılık gelen nokta Zambiya'dır ve doğrunun ikinci noktasıdır. Yani ABD'den geçen doğru Zambiya'dan geçmektedir.**

Görüldüğü üzere Zambiya'dan geçen doğru ABD'den, ABD'den de geçen doğru ise Zambiya'dan geçmektedir. Bu durumda algoritma durmuştur ve veri kümesinden geçen en iyi doğru, yani regresyon doğrusu bulunmuştur.

Buna göre T/2=469,5 sayısını ilk geçen sayıya karşılık gelen eğim, bulunan regresyon doğrusunun eğimidir.

$\hat{\beta}$ **= - 0,218**

ABD'nin koordinatları (8,77) olduğuna göre

Buradan a aşağıdaki şekilde bulunur:

$\hat{\alpha}$ **= 77 – 8* (-0,218) =78,744**

ve böylece regresyon modeli: $\hat{Y}$ =78,744 - 0,218.X bulunmuştur.

3.2.3 β 'nın Testi

E.K.M.S. yöntemi ile yukarıdaki örneğin $\hat{\beta}$ sı -0,21788 hesaplanmıştır. Doğuşta beklenen ömür yılı ile beş yaş altı ölüm oranı arasındaki bu negatif ilişkinin anlamlılığı için β=0 hipotez testi yapılır.

E.K.M.S. regresyon tahmin edicileri $\hat{\alpha}$ ve $\hat{\beta}$ ve kalıntılar ($\hat{e}_i$) bulunur. m sıfır olmayan kalıntıların sayısıdır. Sıfır olmayan kalıntılar en küçükten en büyüğe doğru sıralanır. Bu veri kümesinde ABD ve Zambiya için kalıntılar 0 olduğundan m=41-2=39'dur. E.K.M.S. yöntemi ile bulunan kalıntılar Ek2'de yer almaktadır.

k_1 $\frac{m+1}{2} - \sqrt{m}$ **'ye en yakın değer, k_2** $\frac{m+1}{2} - \sqrt{m}$ **'ye en yakın değerdir.**

$k_1 = \frac{39+1}{2} - \sqrt{39}$ **= 13,755 ≈ 14**

$k_2 = \frac{39+1}{2} + \sqrt{39}$ **= 26,245 ≈ 26**

$\hat{e}_{k_1}$ **14. kalıntı,** $\hat{e}_{k_2}$ **26. kalıntı olur.**

$\hat{e}_{k_1}$ **= -1,436**

$\hat{e}_{k_2}$ **= 0,062**

$\hat{\tau} = \frac{\sqrt{m}\left[\hat{e}_{k_2} - \hat{e}_{k_1}\right]}{4}$ **ve** $S_{\hat{\beta}} = \frac{\hat{\tau}}{\sqrt{\sum (x_i - \bar{x})^2}}$ **şeklinde hesaplanır.**

$\hat{\tau} = \frac{\sqrt{39}\left[0,062 - (-1,436)\right]}{4}$ **= 2,339 bulunmuştur.**

$\sum (x_i - \bar{x})^2 = 100589$ **'dur.**

$S_{\hat{\beta}} = \frac{\hat{\tau}}{\sqrt{\sum (x_i - \bar{x})^2}} = \frac{2,339}{317,1577} = 0,007$ **bulunmuştur.**

Test istatistiği $|t| = \frac{|\hat{\beta}|}{S_{\hat{\beta}}}$ **formülü ile hesaplanır.**

Buna göre $|t| = \frac{|\hat{\beta}|}{S_{\hat{\beta}}} = \frac{0,21788}{0,007} = 31,126$ **bulunmuştur.**

Tablodan 0,001 anlamlılık seviyesi için t değeri 3,551'dir.

Hesaplanan t değeri tablo değerinden büyük olduğu için H_0 hipotezi reddedilerek $\hat{\beta}$ nın anlamlı olduğu söylenir.

- **En Küçük Kareler Yöntemi ile En Küçük Mutlak Sapmalar Yöntemi'nin Sonuçlarının Karşılaştırılması:**

En Küçük Kareler Yöntemi ile σ=2,958625 bulunmuştu, EnKüçük Mutlak Sapmalar Yöntemi ile de Γ =2,339 bulunmuştur.

$\frac{\tau}{\sigma}$**=0,791 <1 olduğundan bu örnek için "En Küçük Mutlak Sapmalar Yöntemini kullanmak, En Küçük Kareler Yöntemini kullanmaktan daha uygundur" denilebilir.**[11]

[11] **E.K.M.S. tahmin edicileri ve β nın anlamlılığının testi excel kullanarak, tüm süreçler adım adım yapılmak suretiyle hesaplanmıştır. İlk adım detaylı açıklanmış, diğer adımların ise (aynı işlemlerin tekrarlanmasından ötürü) , sadece sonuçları verilmiştir.**

4. M-TİPİ TAHMİN YÖNTEMLERİ (M TİPİ TAHMİN EDİCİLER)

M-Tipi Tahmin Yöntemleri ve E.K.M.S. Yöntemi, robust istatistiğin bölümlerindendir. Bir istatistik işlemi, istatistiksel modelin varsayımları geçerli olmadığında yani varsayımlardan sapmalar olması halinde de geçerliliğini muhafaza ediyorsa "robust" olarak nitelendirilmektedir.

Veri kümesinin hata paylarının normal dağılıma uyması durumunda, En Küçük Kareler tahmin edicileri için yapılan hipotez testleri verilere uygun sonuç göstermektedirler; fakat anakütlenin hata paylarının dağılımı normallik varsayımına uygun değilse, En Küçük Kareler tahmin edicileri için yapılan hipotez testleri, varsayımlardan sapmalardan etkilendikleri için "robust" olmamaktadırlar.

M-Tipi Tahmin Yöntemleri, En Küçük Mutlak Sapmalar Yöntemi gibi, veri kümesinde sapan değer/değerlerin olması veya hata paylarının normal dağılmaması durumunda kullanılmak amacıyla ortaya atılmıştır. "M-tahmin edicisi" fikri ilk kez Peter Huber tarafından 1964'te ortaya konulmuştur.[12] Daha sonradan Tukey[13], Hampel[14] ve Andrews de çeşitli yaklaşımlarda bulunmuşlardır.

4.1 M- Tahmin Edicisinin L1-Tahmin Edicisinden Farkı

En Küçük Kareler yönteminin, sapan değerlere karşı çok duyarlı olduğu bilinmektedir. Sapan değer olması durumunda "En Küçük Mutlak Sapmalar" yönteminin En Küçük Kareler yöntemine göre daha iyi sonuç vereceği belirtilmişti. Fakat sapan değerlerin sayısının fazla olması durumunda "En Küçük Mutlak Sapmalar" yöntemi de yetersiz kalabilmektedir ve daha robust regresyon yöntemlerine ihtiyaç duyulmaktadır. Robust-M Tahmin Edicileri bu ihtiyaç sonucunda doğmuştur.

M-tipi tahmin ediciler, ortalama ve medyana alternatif olan merkezi eğilim ölçülerinin tahmincileridir. M-tipi tahmin edicilerin hesaplama şekli, birimlerin dağılımının merkezden uzaklıklarına bağlı olarak çeşitli tartı değerleri verilerek yapılmaktadır.

---

[12] **Huber, P. J. , (1981), "Robust Statistics", New York: John Wiley & Sons,Inc.**

[13] **Hoaglin, D. C., Mosteller, F., Tukey, J. W., (2000), "Understanding Robust and Exploratory Data Analysis", New York: Wiley**

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

Değer merkezden uzaklaştıkça, tartısı (verilen nispi değer) küçülür. Çeşitli M-tipi tahmin edicilerde farklı tartılandırma sistemleri kullanılmaktadır. [http://web.uccs.edu/lbecker/SPSS/explore1.htm]

E.K.M.S. tahmin edicileri, sapan değerlere karşı E.K.K. tahmin edicileri kadar duyarlı olmamaları sebebiyle E.K.K. tahmin edicilerine göre avantajlıdırlar. Fakat sapan değer olmaması durumunda E.K.K. tahmin edicileri E.K.M.S. tahmin edicilerinden daha doğru sonuçlar vermektedirler. Kısacası, sapan değer olması durumunda E.K.M.S. tahmin edicilerini kullanmak, sapan değer olmaması durumunda ise E.K.K. tahmin edicilerini kullanmak daha avantajlı olmaktadır. M tahmin edicileri ise, E.K.K. ve E.K.M.S. yöntemlerinin her ikisinin de avantajlarını içerdiğinden dolayı önem kazanmaktadırlar.

M-Regresyonun En Küçük Mutlak Sapmalar Regresyonundan farkı, sapan değerlerin etkisini azaltmak için ağırlık kullanılmasıdır. Büyük kalıntılara daha küçük ağırlıklar vermek suretiyle, büyük kalıntıların etkisini azaltarak, sapan dğerlerlerden etkilenmeyi minimum hale getirebilmektedir, böylece En Küçük Mutlak Sapmalar Regresyonuna göre daha etkin olabilmektedir.[15]

M-tahmin edicisinin formülünün basit bir değişimi ile, tartıların verilere bağlı olduğu bir tartılı ortalama elde edilebilir. Böylece, daha kolay bir hesaplama tekniğine ulaşılabilir. Bu şekilde elde edilen w-tahmin edicileri özellikle regresyon problemlerinde bilinen en küçük kareler yöntemini değiştirerek "Yeniden Tartılanmış İteratif En Küçük Kareler" yöntemine temel teşkil eder.[16]

4.2 Regresyon Doğrusunun (Huber-M tipi tahmin edicilerle) Tahmin Edilmesi

Huber M-tahmin edicileri olarak bilinen M tahmin edicileri, e'nin bir fonksiyonu olan $\rho(e)$ fonksiyonunu kullanarak e^2 ile $|e|$ arasında bir uzlaşma/uyuşma sağlayarak

---

[15] , [16] Hoaglin, D. C., Mosteller, F., Tukey, J. W., (2000), "Understanding Robust and Exploratory Data Analysis", New York: Wiley

E.K.K. ve E.K.M.S. yöntemlerinin her ikisinin de avantajlarını birleştirmektedir.[17]

4.2.1 $\rho(e)$ Fonksiyonu

M-tipi tahmin edicileri, Huber tarafından tanımlanan, hata paylarının fonksiyonu olan $\rho(e)$ fonksiyonunu minimize etmek suretiyle hesaplanmaktadır.[18]

$$\rho(e) = \begin{cases} e^2, & -k \le e \le k \\ 2k|e| - k^2, & e < -k \cup k < e \end{cases} \qquad \textbf{(4.1)}$$

$\rho(e)$; "e, 0'a yakın olduğunda e^2 'ye, e, 0'dan uzak olduğunda ise $|e|$ ye eşit olan bir fonksiyon" olarak tanımlandığında E.K.K. ve E.K.M.S. yöntemlerinin avantajları birleştirilebilmektedir.

- **$\rho(e) = e^2$ olduğunda En Küçük Kareler Yöntemini,**
- **$\rho(e) = |e|$ olduğunda En Küçük Mutlak Sapmalar Yöntemini vermektedir.**

[http://godel.cs.bilgi.edu.tr/mirror/cran/]

Yukarıdaki eşitliklerden de $\rho(e)$ fonksiyonunun, En Küçük Kareler Yöntemi ile En Küçük Mutlak Sapmalar Yöntemlerinin avantajlarını birleştirdiği görülmektedir.

$\hat{\sigma}$, anakütle hata paylarının standart sapması olan σ'nin tahmin edicisine, k ise 1,5$\hat{\sigma}$ ya eşittir. σ'yı tahmin etmek için $\hat{\sigma}$=1,483.M.S.M. kullanılmaktadır. M.S.M. mutlak sapmaların medyanı olup 1,483 çarpanı da rasgele hata paylarının normal dağıldığı varsayımı altında $\hat{\sigma}$, σ'nın iyi bir tahmin edicisi olmasını sağlayabilmek için seçilmiştir.

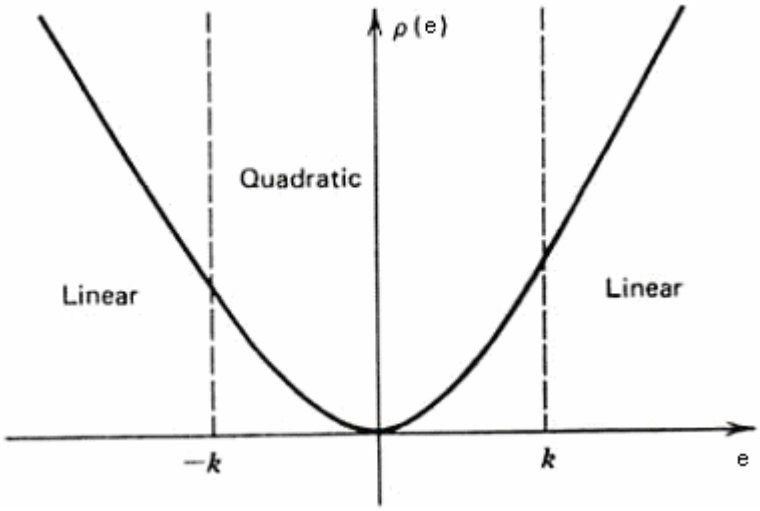
$\rho(e)$'yi keskin olmayan bir fonksiyon haline getirmek için $|e|$ yerine $2k|e| - k^2$ kullanılmıştır.

$\rho(e)$ fonksiyonun grafiği aşağıda verilmektedir.

---

[17] **Birkes, D. ve Dodge,Y., (1993), "Alternative Methods of Regression", Wiley**

[18] **Weisberg, S. , (1985), "Applied Linear Regression", New York: Wiley**



Şekil 4.1 Huber'in $\rho(e)$ fonksiyonunun grafiği

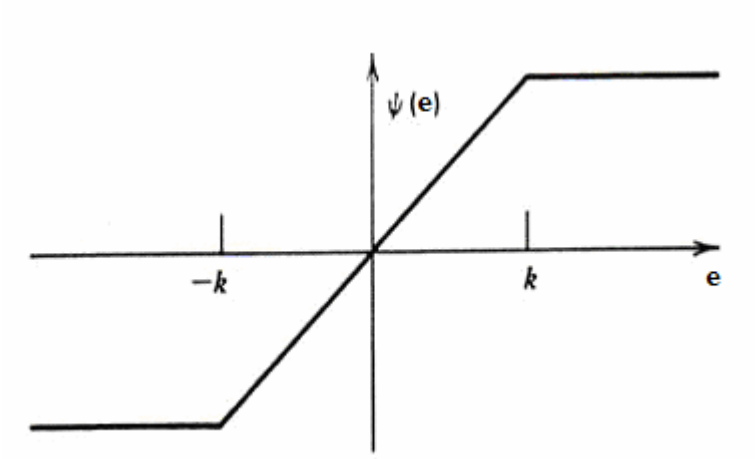
4.2.2 Etki Fonksiyonu

(4.1)'de verilen $\rho(e)$'in e'ye göre türevi olan $\Psi(e)=\frac{d\rho(e)}{dx}$ fonksiyonu, etki fonksiyonu olarak adlandırılmaktadır. Etki fonksiyonu, her verinin, parametre tahmin edicisi üzerindeki marjinal etkisini ölçmektedir. Örneğin, En Küçük Kareler Yönteminde $\rho(e)=e^2$ dir ve etki fonksiyonu da $\Psi(e)=e$ dir. bu, her verinin tahmin edici üzerindeki etkisinin hata payıyla doğru orantılı olarak arttığını göstermektedir. Bu durum da, En Küçük Kareler Yönteminin bir kez daha robust olmadığını doğrulamaktadır. Bir M tahmin edicisinin etki fonksiyonunun sınırlandırılmış olması gerekmektedir.

[http://www-sop.inria.fr/robotvis/personnel/zzhang/Publis/Tutorial-Estim/node24.html]

Huber M tahmin edicisi için etki fonksiyonu ve grafiği aşağıdaki gibidir:

$$\Psi(e)=\begin{cases} e, & |e|\le k \\ k\,sgn(e), & |e|>k \end{cases} \qquad \textbf{(4.2)}$$



Şekil 4.2 Huber'in $\Psi(e)$ fonksiyonunun grafiği

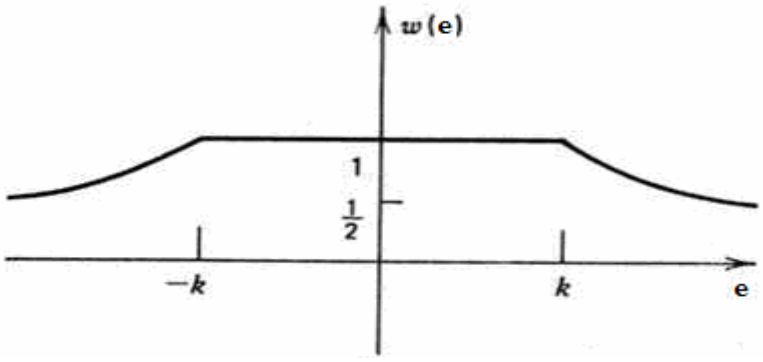
4.2.3 Tartı Fonksiyonu

Etki fonksiyonunun x'e bölünmesiyle $\left(w(e) = \frac{\Psi(e)}{e} \right)$ tartı fonksiyonu bulunmaktadır.[19]

[http://elections.fas.harvard.edu/wssmh/node11.html]

Huber M tahmin edicisi için ağırlık fonksiyonu ve grafiği aşağıdaki gibidir:

$$w(e) = \begin{cases} 1, & |e| \le k \\ \dfrac{k}{|e|}, & |e| > k \end{cases} \qquad \textbf{(4.3)}$$



Şekil 4.3 Huber'in $w(e)$ fonksiyonunun grafiği

---

[19] **Hoaglin, D. C., Mosteller, F., Tukey, J. W., (2000), "Understanding Robust and Exploratory Data Analysis", New York: Wiley**

[http://www-sop.inria.fr/robotvis/personnel/zzhang/Publis/Tutorial-Estim/node24.html]

[http://www.vision.ee.ethz.ch/~rpaget/Markov/Chapter_4/node70.html]

4.3 Algoritma

$\rho(e)$ fonksiyonunu minimum yapmak için bir algoritmaya ihtiyaç duyulmaktadır. Algoritmanın ilk aşamasında, E.K.K. Yönteminden elden edilen tahmin ediciler, α ve β'nın ilk tahmin edicileri olarak alınmaktadır. E.K.K. tahmin edicilerinden yararlanılarak, sapmalar ve σ'nın tahmin edicileri hesaplamaktadır. Bu tahmin ediciler kullanılarak α ve β'nın "iyileştirilmiş tahmin edicileri" hesaplanmaktadır. İyileştirilmiş tahmin ediciler vasıtasıyla ise yeni sapmalar ve σ'nın iyileştirilmiş tahmin edicisi hesaplanmaktadır. Daha sonra yeni sapmalar ve yeni $\hat{\sigma}$, α ve β'nın iyileştirilmiş tahmin edicilerini hesaplamada kullanılmaktadır. Algoritma, iyileştirilmiş tahmin ediciler bir önceki adımda bulunan tahmin edicilerle aynı çıkana kadar bu şekilde devam eder.

Daha açık bir şekilde belirtmek gerekirse, a^0 ve b^0, α ve β'nın mevcut tahmin edicileri olmak üzere, önce $y_i - (a^0 + b^0 x_i)$ sapmaları, sonra bu sapmalardan da $\hat{\sigma}^0$=1,483 MSM hesaplanmaktadır. Daha sonra y değerlerini, aşırı sapmalardan arındırmak amacıyla aşağıdaki düzenleme yapılmaktadır:

y_i'nin tahmin edilen başlangıç regresyon doğrusundan sapmasını gösteren $e_i^0 = y_i - (a^0 + b^0 x_i)$ ifadesi, $y_i = a^0 + b^0 x_i + e_i^0$ şeklinde yazılabilir. Buradan hareketle, e_i*, e_i^0'ın daraltılmasıyla elde edilen sapma olmak üzere, $y_i^* = a^0 + b^0 x_i + e_i^*$ şeklinde yazılır. Bu sayede sapmalardan hiçbirinin mutlak değerinin $1{,}5\hat{\sigma}^0$'nin dışında olmaması sağlanmış olur. Bu durum aşağıdaki gibi ifade edilebilir:

$$e_i^* = \begin{cases} e_i^0, & -1{,}5\hat{\sigma}^0 < e_i^0 < 1{,}5\hat{\sigma}^0 \\ -1{,}5\hat{\sigma}^0, & e_i^0 < -1{,}5\hat{\sigma}^0 \\ 1{,}5\hat{\sigma}^0, & e_i^0 > 1{,}5\hat{\sigma}^0 \end{cases} \qquad \textbf{(4.4)}$$

Ayarlanmış y_1^*, … , y_n^* veri kümesinden elde edilen EKK tahmin edicileri, α ve β'nın

iyileştirilmiş tahmin edicileridir.[20] [http://godel.cs.bilgi.edu.tr/mirror/cran/]

4.4 β İçin Hipotez Testi

Çok değişkenli regresyon modelinde β=0 hipotez testi β_{q+1} = …= β_p =0 şeklinde yazılır fakat p'nin 1 ve q'nun da 0 olduğu özel durumda β=0 test edilir. STR (Sum of Transformed Residuals) "Dönüştürülmüş Kalıntıların Toplamı" olmak üzere, test istatistiği aşağıdaki gibidir.

$$F_M = \frac{STR_{kısıtlanmış} - STR_{tam}}{(p-q)\hat{\lambda}} \qquad \textbf{(4.5)}$$

m, +/- k=1,5$\hat{\sigma}$ değerleri arasındaki kalıntı sayısı olmak üzere,

$$\lambda = \left(\frac{n}{m}\right)\frac{\sum e_i^{*2}}{(n-p-1)} \text{ dir.} \qquad \textbf{(4.6)}$$

k=1,5$\hat{\sigma}$ eşitliğinde, 1,5 yerine ∞ geldiğinde ve $\hat{e}_i^* = \hat{e}_i$ olduğunda, +/- k=1,5$\hat{\sigma}$ değerleri arasındaki kalıntı sayısını veren "m", "n" ye eşit olmaktadır. Bu durumda, λ, $\hat{\sigma}^2_{EKK}$'ye eşit olmaktadır. Hatta, 1,5 yerine ∞ geldiğinde, $\rho(e) = e^2$ olmakta ve böylece STR=SSR [21]olmaktadır ki bu da, "$F_M = F_{EKK}$" demektir.

4.5 M-Tahmin Edicilerine Farklı Yaklaşımlar

4.5.1 Tukey'in Yaklaşımı

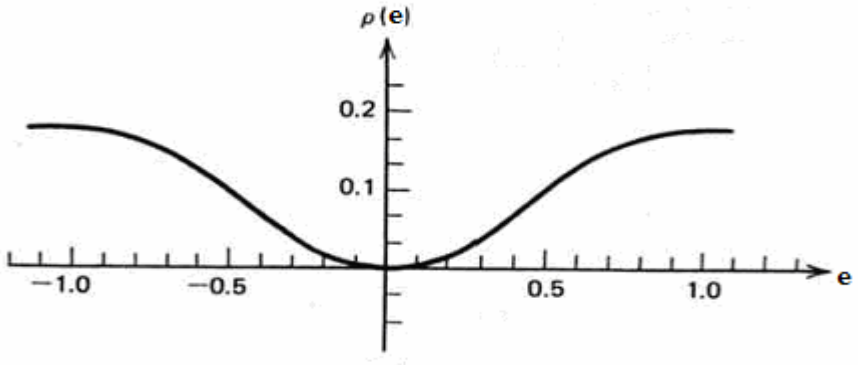
Tukey'in tanımladığı $p(e)$ fonksiyonu ve fonksiyonun grafiği aşağıdaki gibidir.

$$p(e) = \begin{cases} \frac{1}{6}\left[1-(1-e^2)^3\right] & -1 \le e \le 1 \\ \frac{1}{6} & |e| > 1 \end{cases} \qquad \textbf{(4.7)}$$

---

[20] Birkes, D. ve Dodge,Y., (1993), "Alternative Methods Of Regression", Wiley

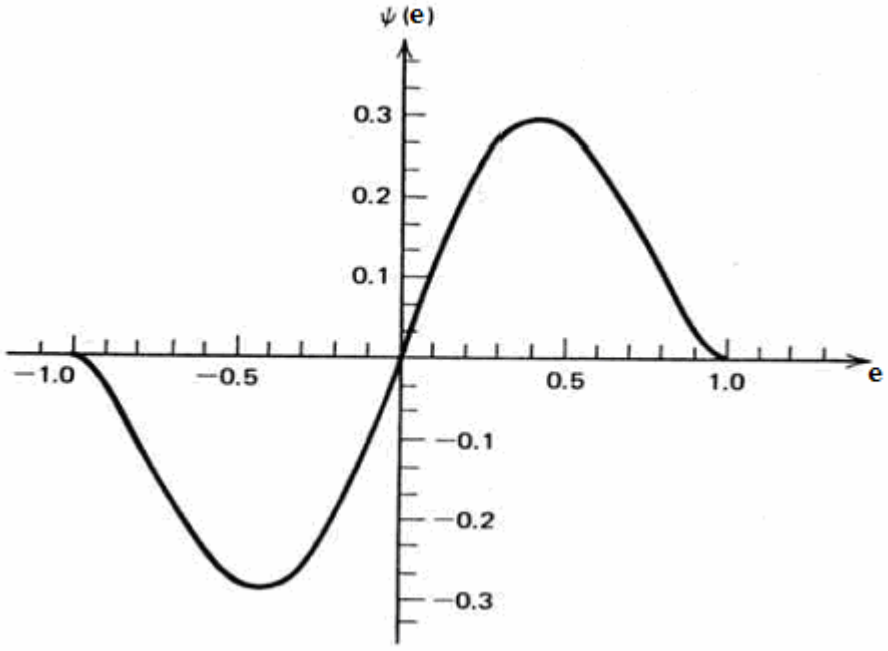
[21] SSR=Sum of Square Residuals (Kalıntıların Karelerinin Toplamı)



Şekil 4.4 Tukey'in $\rho(e)$ fonksiyonunun grafiği

Etki fonksiyonu:

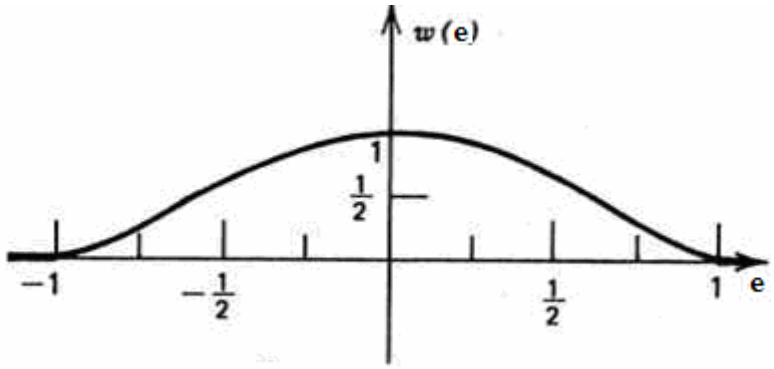
$$\Psi(e)=\begin{cases} e\left(1-e^{2}\right)^{2} & |e|\le k \\ 0 & |e|>k \end{cases} \qquad \textbf{(4.8)}$$



Şekil 4.5 Tukey'in $\Psi(e)$ fonksiyonunun grafiği

Ağırlık fonksiyonu:

$$w(e)=\begin{cases}\left(1-e^2\right)^2 & |e|\le k\\ 0 & |e|>k\end{cases} \qquad \textbf{(4.9)}$$



Şekil 4.6 Tukey'in $w(e)$ fonksiyonunun grafiği

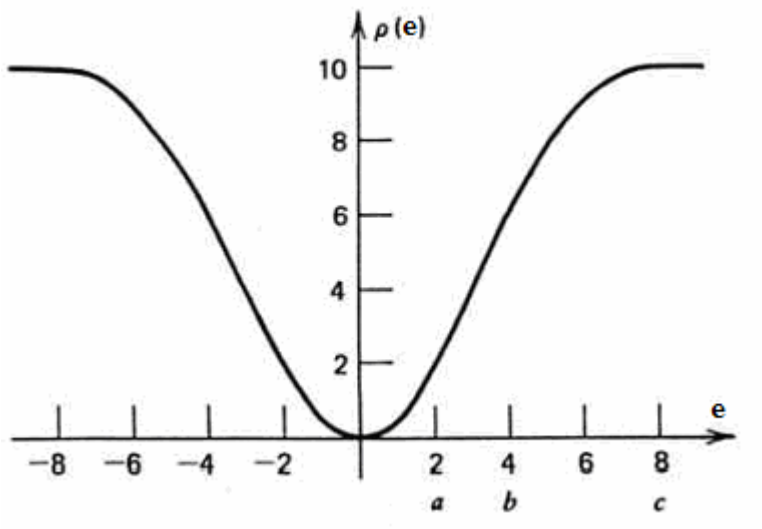
4.5.2 Hampel'in Yaklaşımı

Hampel'in tanımladığı $p(e)$ fonksiyonu ve fonksiyonun grafiği aşağıdaki gibidir.[22]

$$p(e)=\begin{cases}\frac{1}{2}e^2 & |e|\le a\\ a|e|-\frac{1}{2}a^2 & a<|e|\le b\\ ab-\frac{1}{2}a^2+(c-b)\frac{a}{2}\left[1-\left(\frac{c-|e|}{c-b}\right)\right] & b<|e|\le c\\ ab-\frac{1}{2}a^2+(c-b)\frac{a}{2} & |e|>c\end{cases} \qquad \textbf{(4.10)}$$

---

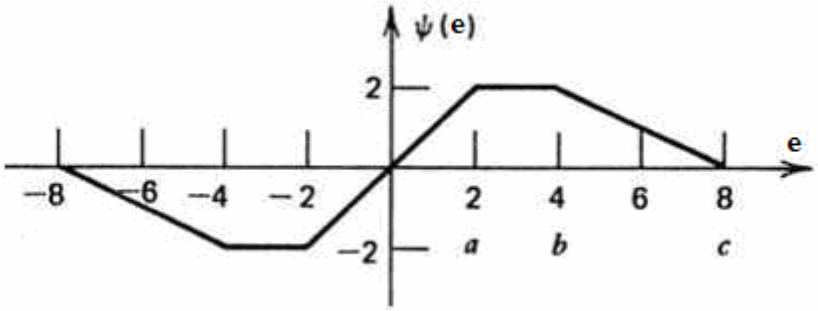
[22] **Hampel, F.R., (1975), "Beyond location parameters: robust concepts and methods. In International Statistical Institute, Proceedings of the 40th Session", 46: 375-391**



Şekil 4.7 Hampel'in $\rho(e)$ fonksiyonunun grafiği

Etki fonksiyonu:

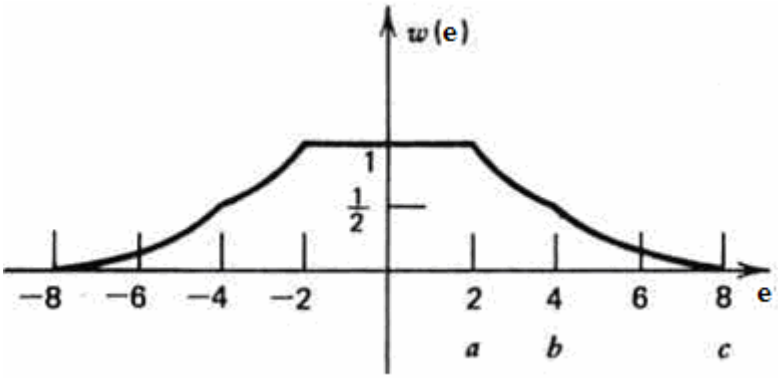
$$\Psi(e)=\begin{cases} e & |e|\le a \\ a.sgn(e) & a<|e|\le b \\ a\dfrac{c-|e|}{c-b}sgn(e) & b<|e|\le c \\ 0 & |e|>c \end{cases} \qquad (4.11)$$



Şekil 4.8 Hampel'in $\Psi(e)$ fonksiyonunun grafiği

Ağırlık fonksiyonu:

$$\Psi(e)=\begin{cases}1 & |e|\le a\\ \dfrac{a.sgn(e)}{e} & a<|e|\le b\\ a\dfrac{c-|e|}{c-b}\dfrac{sgn(e)}{e} & b<|e|\le c\\ 0 & |e|>c\end{cases} \qquad (4.12)$$

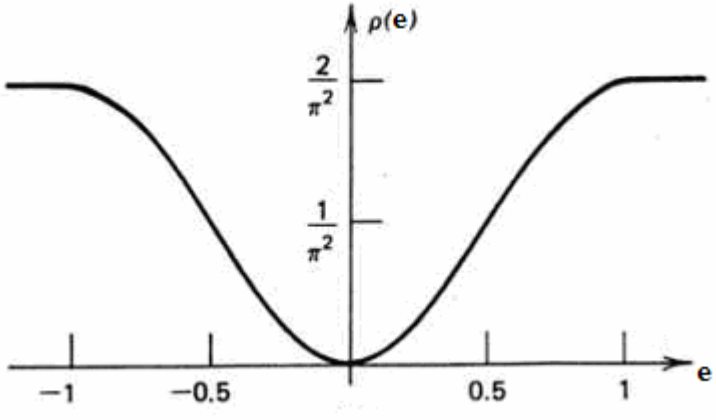


Şekil 4.9 Hampel'in $w(e)$ fonksiyonunun grafiği

4.5.3 Andrews'in Yaklaşımı

Andrews'in tanımladığı $p(e)$ fonksiyonu ve fonksiyonun grafiği aşağıdaki gibidir.

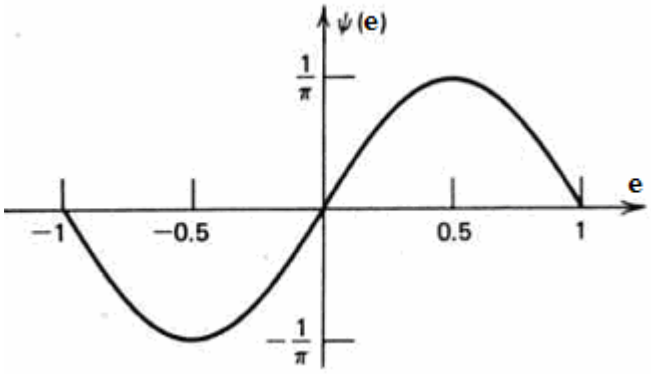
$$p(e)=\begin{cases}\dfrac{1}{\pi^2}\left[1-\cos(\pi e)\right] & -1\le e\le 1\\ \dfrac{2}{\pi^2} & |e|>1\end{cases} \qquad (4.13)$$



Şekil 4.10 Andrews'in $\rho(e)$ fonksiyonunun grafiği

Etki fonksiyonu:

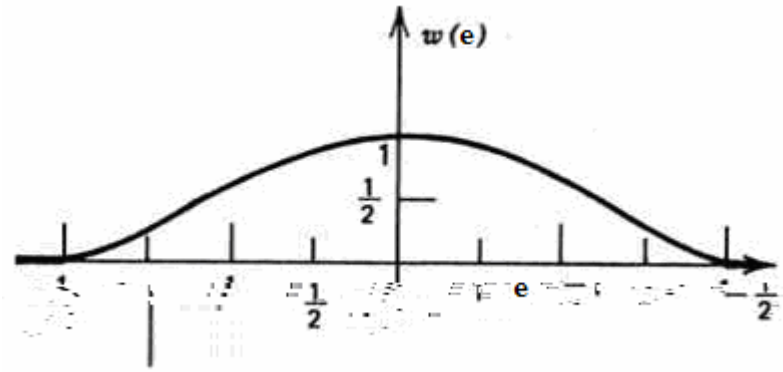
$$\Psi(e)=\begin{cases}\dfrac{1}{\pi}\sin\pi.e & |e|\le k\\ 0 & |e|>k\end{cases} \qquad \textbf{(4.14)}$$



Şekil 4.11 Andrews'in $\Psi(e)$ fonksiyonunun grafiği

Ağırlık fonksiyonu:

$$w(e)=\begin{cases}\dfrac{1}{\pi.e}\sin\pi.e & |e|\le k\\ 0 & |e|>k\end{cases} \qquad \textbf{(4.15)}$$



Şekil 4.12 Andrews'in $w(e)$ fonksiyonunun grafiği

4.6 M-Tipi Tahmin Yöntemleriyle Uygulama

4.6.1 Huber M-tahmin edicileri

Daha önce de anlatıldığı üzere algoritmaya En Küçük Kareler tahmin edicileri

hesaplanarak başlanmaktadır. Veri kümesi için α'nın tahmin edicisini simgeleyen a^0 =77,17355 ve β'nın tahmin edicisini simgeleyen b^0= -0,19114 olarak bulunmuştu. Bu tahmin ediciler vasıtası ile öngörülen y değerleri ($\hat{y}_i^0$) ve e_i^0 sapmaları Ek3'te verilmiştir.

Elde edilen mutlak sapmalar sıraya konularak mutlak sapmaların medyan değeri olan MSM(MAD) = 1,97329

$\hat{\sigma}$ =1,483 MAD =2,92638907

k = 1,5 . $\hat{\sigma}$ =4,3895836 bulunmuştur.

Bundan sonra y değerlerini, aşırı sapmalardan arındıracak olan düzenleme yapmak amacıyla mutlak değeri 1,5.$\hat{\sigma}$ yı aşan değerler, yani + 4,3895836 değerinden büyük ve -4,3895836 değerinden küçük olan sapmalar yerine + 4,3895836 veya - 4,3895836 değeri yazılmıştır. Mutlak değeri 1,5.$\hat{\sigma}$ yı aşmayan e_i^0 ler aynı kalır. Böylece daraltılmış sapmalar olan ei* değerleri hesaplanmıştır. e_i* değerleri hesaplandıktan sonra $y_i^* = a^0 + b^0 x_i + e_i^*$ formülünden y_i^* değerleri bulunmuştur. e_i* değerleri ve y_i^* değerleri Ek4'te verilmiştir.

Bulunan y_i^* değerleri üzerinden tekrar En Küçük Kareler tahmin edicileri hesaplanmıştır. a^0 =77,34826987 ve b^0= -0,199526487 bulunmuştur. Böylece 1. iterasyon tamamlanmıştır.

Yeni tahmin ediciler kullanılarak yeni y değerleri ve yeni sapmalar hesaplanıp, yeni mutlak sapmalar sıraya konularak mutlak sapmaların medyan değeri (MAD) = 1,84889 bulunmuştur.

$\hat{\sigma}$ =1,483 MAD =2,74190247 ve k = 1,5 $\hat{\sigma}$ =4,11285371 bulunmuştur.

Bundan sonra y değerlerini, aşırı sapmalardan arındırmak için tekrar düzenleme yapılır. Bunun için mutlak değeri 1,5$\hat{\sigma}$ yı aşan değerler, yani +4,11285371 değerinden büyük ve – 4,11285371 değerinden küçük olan sapmalar yerine + 4,11285371 veya - 4,11285371 değeri yazılmıştır. Mutlak değeri 1,5$\hat{\sigma}$ yı aşmayan e_i^0 ler aynı kalır. Böylece daraltılmış sapmalar olan e_i* değerleri hesaplanmış olur. e_i* değerleri hesaplandıktan sonra $y_i^* = a^0 + b^0 x_i + e_i^*$ formülünden yeni y_i^* değerleri bulunmuştur. Bulunan bu y_i^* değerleri üzerinden tekrar En Küçük Kareler tahmin edicileri hesaplanarak a^0 =77,37994699 ve b^0=-0,201275551 bulunmuştur. Böylece 2. iterasyon da tamamlanmış olur.

3.iterasyonda MAD= 1,78606 , $\hat{\sigma}$=1,483 MAD =2,64873185 ve k = 1,5 $\hat{\sigma}$=3,97309778 bulunmuştur. Aynı aşamalar takip edilerek a^0 =77,39060866 ve b^0= -0,201774518 bulunmuştur. Böylece 3. iterasyon da tamamlanmış olur.

4.iterasyonda MAD= 1,74092 , $\hat{\sigma}$=1,483 MAD =2,58178755 ve k = 1,5 $\hat{\sigma}$=3,87268132 , a^0 =77,39947034 ve b^0=-0,202031398 bulunmuştur. Böylece 4. iterasyon da tamamlanmış olur.

5.iterasyonda MAD= 1,74259 , $\hat{\sigma}$=1,483 MAD =2,58426274 ve k = 1,5 $\hat{\sigma}$=3,87639411 a^0 =77,40037869 ve b^0=-0,202065905 bulunmuştur. Böylece 5. iterasyon da tamamlanmış olur.

6.iterasyonda MAD= 1,74253 , $\hat{\sigma}$=1,483 MAD =2,58417693 ve k = 1,5 $\hat{\sigma}$=3,8762654 a^0 =77,40051673 ve b^0=-0,202071763 bulunmuşur. Böylece 6. iterasyon da tamamlanmış olur.

7.iterasyonda MAD= 1,74251, $\hat{\sigma}$=1,483 MAD =2,58413841 ve k = 1,5 $\hat{\sigma}$=3,87620762 a^0 =77,40054124 ve b^0= -0,202072832 bulunmuştur. Böylece 7. iterasyon da tamamlanmış olur.

7 iterasyon sonucunda hem α'nın hem de β'nın yeni tahmin edicileri ile bir önceki iterasyondaki tahmin edicileri arasında 10^{-4} ten küçük bir fark olduğundan dolayı algoritma durur ve $\hat{\alpha}$=77,40054124 ve $\hat{\beta}$= -0,202072832 elde edilmiş olur.

Sonuç olarak, Huber-M yöntemi ile $\hat{Y}$ = 77,40054124 - 0,202072832 X regresyon modeli bulunmuştur.

- **β=0'ın testi:**

Elde edilen tahmin ediciler vasıtası ile hesaplanan öngörülen y değerleri ve $\hat{e}_i$ kalıntıları Ek 5'te verilmiştir.

Mutlak kalıntılar sıralanırsa MAD=1,74250, $\hat{\sigma}$=1,483.MAD=2,58413036 ve

k=1,5 $\hat{\sigma}$=3,87619554 hesaplanır.

$p(e)=\begin{cases} e^2, & -k \le e \le k \\ 2k|e|-k^2, & e<-k \cup k<e \end{cases}$ **fonksiyonunda da tanımlandığı gibi mutlak kalıntısı**

(–k,+k) aralığında olan e'ler için e^2 , mutlak kalıntısı k'yı aşan e'ler için $2k|e|-k^2$ hesaplanır. $\rho(ei)$ değerleri, daraltılmış kalıntılar ve daraltılmış kalıntıların kareleri Ek 6'da verilmiştir.

STR_{tam}=285,36072 olarak hesaplanmıştır.

Daraltılmış kalıntıların karelerinin toplamı = 193,867210 olarak hesaplanmıştır.

$$\lambda = \left(\frac{n}{m}\right)\frac{\sum e_i^{*2}}{(n-p-1)} = \left(\frac{41}{41-5}\right)\frac{(193,867210)}{41-1-1} = 5,66136$$

β=0'ın testi için indirgenmiş model Y=α+e olarak alınır. Bu modele de M tahmin edici yöntemi uygulanır, fakat bu defa $\hat{\sigma}$ iterasyona uğratılmaz. İterasyonlar süresice $\hat{\sigma}$, tam modelin son iterasyonunda bulunan değeri alır. Bu örnek için $\hat{\sigma}$=2,58413036 dir.

1,5.k = 3,876208 dir.

Algoritmaya, Y=α+e modeli için α nın En Küçük Kareler tahmin edicisi olan $\bar{y}$ nın hesaplanması suretiyle başlanır. $a^0 = \bar{y} = 71,65854$ dir. Buradan hareketle sapmalar $e_i^0 = y_i - 71,65854$ formülünden hesaplanmıştır. Daha sonra 1,5.k = 3,876208 değerinden büyük olan sapmalar +3,876208, 1,5.k = 3,876208 değerinden küçük olan sapmalar ise -3,876208 değerini alırlar. Diğer sapmalar aynı kalır, böylece daraltılmış sapmalar hesaplanmış olunur. Ve en son olarak da y* değerleri $y^* = a^0 + e^*$ formülünden hesaplanır. Hesaplanan daraltılmış sapmalar ve y* değerleri Ek 7'de verilmiştir.

Yeni $a^0 = \bar{y}$ =72,44585 bulunur. böylece 1. iterasyon tamamlanmış olur.

14 iterasyon sonucunda bulunan $a^0 = \bar{y}$ değerleri aşağıdaki tabloda verilmiştir.

Çizelge 4.1 İterasyonlar sonucunda bulunan $\bar{y}$ değerleri

	$a^0 = \bar{y}$
1.İterasyon	**72,44585**
2. İterasyon	**72,87180**
3. İterasyon	**73,08997**

4. İterasyon	**73,21215**
5. İterasyon	**73,27206**
6. İterasyon	**73,29983**
7. İterasyon	**73,31270**
8. İterasyon	**73,31866**
9. İterasyon	**73,32142**
10. İterasyon	**73,32270**
11. İterasyon	**73,32330**
12. İterasyon	**73,32357**
13. İterasyon	**73,32370**
14. İterasyon	**73,32376**

13. iterasyon sonunda bulunan a^0 (73,323699) ile 14. iterastonda bulunan a^0 (73,32376) arasındaki fark 0,00006 dır. farkın çok küçük olması sebebiyle algoritma durmuştur. Bulunan son a^0 değeri, $\hat{\alpha}$ =73,323699 dır.

$\hat{e}_i$ **kalıntıları ve daraltılmış e* kalıntıları hesaplanır. daraltılmış e* kareleri alınır.** $p(\hat{e}_i) = \begin{cases} e^2, & -k \le e \le k \\ 2k|e| - k^2, & e < -k \cup k < e \end{cases}$ **fonksiyonunda da tanımlandığı gibi mutlak kalıntısı (–k,+k) aralığında olan e'ler için e^2 , mutlak kalıntısı k'yı aşan e'ler için** $2k|e| - k^2$ **yazılarak** $p(\hat{e}_i)$ **değerleri bulunur. Daraltılmış kalıntılar, daraltılmış kalıntıların kareleri ve** $\rho(ei)$ **değerleri Ek 8'da verilmiştir.**

$p(\hat{e}_i)$ **değerlerinin toplamı $STR_{kısıtlanmış}$ değerini oluşturur. $STR_{kısıtlanmış}$ =1149,624 dir.**

Test istatistiği $F_M = \dfrac{STR_{kısıtlanmış} - STR_{tam}}{(p-q)\hat{\lambda}} = \dfrac{(1149{,}624 - 285{,}36072)}{5{,}66136} = 152{,}66011$

bulunmuştur.

1,39 serbestlik dereceli F_{tablo} değeri 12,61'dir.

$F_M > F_{tablo}$ olduğundan dolayı H_0 reddedilerek β değerinin anlamlı olduğu söylenir.[23]

Tabii ki bir regresyon modeli bulmak için bu kadar çok işlem yapmak epey külfetlidir. M tahmin edicilerini hesaplamak için R project programı kullanılabilir. X ve Y değerlerinin vektör olarak tanımlanmasını gerekli kılan bu programda kısa komutlar yazmak gerekmektedir.[24] R project'te En Küçük Kareler ve M-tahmin edicilerin hesaplanmasına ilişkin komutlar Ek 9'da verilmiştir.

- **R project kullanılarak bulunan Huber-M tahmin edicileri aşağıdaki tablodadır.**

Çizelge 4.2 R Project kullanılarak elde edilen Huber-M tamin edicileri

Residuals:

Min	1Q	Median	3Q	Max
-7.3718	-1.5325	0.2714	1.7816	11.1810

Coefficients:

	Value	Std. Error	t value
(Intercept)	77.4384	0.4957	156.2206
x	-0.2033	0.0100	-20.3172

Residual standard error: 2.587 on 39 degrees of freedom

Correlation of Coefficients:

	(Intercept)
x	-0.5825

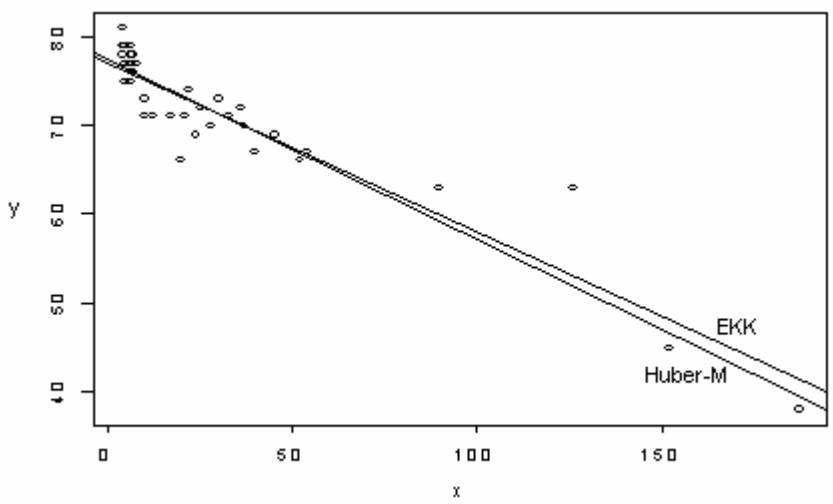
R project'te Huber-M tahmin edicileri kullanılarak bulunan regresyon doğrusu

$\hat{Y}$ = 77,4384 - 0,2033 X olmaktadır.

---

[23] **Huber-M tahmin edicileri ve β nın anlamlılığının testi excel kullanarak, tüm süreçler adım adım yapılmak suretiyle hesaplanmıştır. İlk iterasyonlar detaylı açıklanmış, diğer iterasyonların ise (aynı işlemlerin tekrarlanmasından ötürü) , sadece sonuçları verilmiştir.**

[24] **M-tahmin edicilerini hesaplamak için kullanılan bir diğer paket program S Plus'tır, fakat S Plus ücretli bir programdır. R ise http://cran.r-project.org/ adresinden indirilebilmekte ve ücretsiz olarak kullanılabilmektedir. Bu nedenle uygulamalar R üzerinde yapılmıştır. Bu iki programda da S proglamlama dili kullanılmaktadır.**



Şekil 4.13 En Küçük Kareler ve Huber-M tahmin edicileriyle bulunan regresyon doğruları

4.6.2 Tukey'in M-tahmin edicileri

R project kullanılarak bulunan Tukey-M tahmin edicileri aşağıdadır.

Çizelge 4.3 R Project kullanılarak elde edilen Tukey-M tahmin edicileri

Converged in 6 iterations

Coefficients:

(Intercept) x

77.5586995 -0.2102955

Degrees of freedom: 41 total; 39 residual

Scale estimate: 2.53

Residuals:

Min	1Q	Median	3Q	Max
-7.3528	-1.5072	0.3811	1.7031	11.9385

Coefficients:

	Value	Std. Error	t value
(Intercept)	77.5587	0.4617	167.9990

x -0.2103 0.0093 -22.5626
Residual standard error: 2.525 on 39 degrees of freedom
Correlation of Coefficients:
(Intercept)
x -0.5825

R project'te Tukey-M tahmin edicileri kullanılarak bulunan regresyon doğrusu

$\hat{Y}$ **= 77,5587 - 0,2103 X olmaktadır.**

4.6.3 Hampel'ın M-tahmin edicileri

R project kullanılarak bulunan Hampel-M tahmin edicileri aşağıdadır.

Çizelge 4.4 R Project kullanılarak elde edilen Hampel-M tamin edicileri

Converged in 5 iterations
Coefficients:
(Intercept) x
77.3156593 -0.1989494
Degrees of freedom: 41 total; 39 residual
Scale estimate: 2.78
Residuals:
Min 1Q Median 3Q Max
-7.3367 -2.0754 0.2497 1.8465 10.7520
Coefficients:
Value Std. Error t value
(Intercept) 77.3157 0.5006 154.4443

x	-0.1989	0.0101	-19.6848

Residual standard error: 2.784 on 39 degrees of freedom

Correlation of Coefficients:

(Intercept)

x -0.5825

R project'te Hampel-M tahmin edicileri kullanılarak bulunan regresyon doğrusu

$\hat{Y}$ **= 77,3157 - 0,1989 X olmaktadır.**

4.7 Andrews'in M-tahmin edicileri

R project kullanılarak bulunan Andrews-M tahmin edicileri aşağıdadır.

Çizelge 4.5 R Project kullanılarak elde edilen Andrews-M tamin edicileri

Converged in 5 iterations

Coefficients:

(Intercept)	x
77.4251826	-0.2015312

Degrees of freedom: 41 total; 39 residual

Scale estimate: 2.65

Residuals:

Min	1Q	Median	3Q	Max
-7.3451	-1.8026	0.2836	1.8326	11.1268

Coefficients:

	Value	Std. Error	t value
(Intercept)	77.4252	0.4818	160.2664

x	**-0.2015**	**0.0092**	**-21.5821**

Residual standard error: 2.653 on 39 degrees of freedom

Correlation of Coefficients:

(Intercept)

x -0.5825

R project'te Andrews-M tahmin edicileri kullanılarak bulunan regresyon doğrusu

$\hat{Y}$ **= 77,4252 - 0,2015 X olmaktadır.**

5. SONUÇ

Hata payları normal dağıldığında ve/veya veri kümesinde sapan değerler olmadığında En Küçük Kareler yöntemi klasik regresyon parametreleri tahmin yöntemleri içinde en kesin / güvenilir sonucu vermektedir.

Fakat veri kümesinde sapan değerlerin olması ve/veya hata paylarının normal dağılıma uymaması durumunda, varsayımlardan etkilenmesi sebebiyle hatalı sonuçlar vermektedir.

Çalışmada, veri kümesine önce E.K.K. yöntemi uygulanmış, regresyon modeli

$\hat{Y}$ **= 77,17355 - 0,19114 X bulunmuştur.**

Fakat normallik testleri yapıldığında hata paylarının normal dağılmadığı saptanmış, hata paylarının grafiğinden de Pakistan ve Rusya Federasyonu noktalarının sapan değer olduğu tespit edilmiştir. Bu veriler çıkarılarak tekrar En Küçük Kareler Yöntemi uygulandığında,

$\hat{Y}$ **=77,59229 - 0,20935 X bulunmuştur.**

Fakat, sapan değerlerin veri kümesinden dışlanarak regresyon modeli bulmak, bilgi kaybına sebep olacağından ve hatalı sonuç vereceğinden robust regresyon yöntemlerine gerek duyulmuş ve En Küçük Mutlak Sapmalar yöntemi uygulanmıştır. En Küçük Mutlak Sapmalar yöntemi, sapan değerlere karşı En Küçük Kareler yönteminden daha dirençli bir yöntemdir. En Küçük Mutlak Sapmalar yöntemi uygulandığında

$\hat{Y}$ **=78,744 - 0,218.X bulunmuştur.**

Daha önce de bahsedildiği gibi E.K.K. ve E.K.M.S. regresyon yöntemlerinden hangisinin daha iyi sonuç verdiği $\frac{\tau}{\sigma}$ **oranına bağlıdır. Hata payları normal dağılıyorsa** $\frac{\tau}{\sigma} > 1$ **olmaktadır. Hata paylarının normal dağılmaması halinde ise,** $\frac{\tau}{\sigma} < 1$ **olmaktadır. Ele alınan veri kümesi için,** $\frac{\tau}{\sigma} = 0{,}791 < 1$ **olduğundan bu veri kümesi için "En Küçük Mutlak Sapmalar Yöntemini kullanmak, En Küçük Kareler Yöntemini kullanmaktan daha uygundur" denilebilir.**

Her ne kadar "En Küçük Mutlak Sapmalar" yöntemi En Küçük Kareler yöntemine

göre daha iyi sonuç verse de, sapan değerlerin sayısının fazla olması durumunda "En Küçük Mutlak Sapmalar" yöntemi de yetersiz kalabilmektedir ve daha robust regresyon yöntemlerine ihtiyaç duyulmaktadır. Bu durumda Robust-M tahmin edicileri kullanılmaktadır. Huber'in M-tahmin edicisi yönteminin excel kullanılarak, tüm iterasyonlar tek tek yapılarak uygulanması sonucu

$\hat{Y}$ **= 77,40054124 - 0,202072832 X modeli bulunmuştur.**

R project'te Huber-M tahmin edicileri kullanılarak bulunan regresyon doğrusu ise

$\hat{Y}$ **= 77,4384 - 0,2033 X olmaktadır. Aynı yöntem uygulandığı halde birebir aynı sonucu vermemesinin sebebi, birinin elle hesaplanmasından, diğerinin ise program çıktısı olmasından kaynaklanmaktadır.**

R project'te Tukey-M tahmin edicileri kullanıldığında $\hat{Y}$ **= 77,5587 - 0,2103 X regresyon modeli bulunmuştur.**

R project'te Hampel-M tahmin edicileri kullanıldığında $\hat{Y}$ **= 77,3157 - 0,1989 X regresyon modeli bulunmuştur.**

R project'te Andrews-M tahmin edicileri kullanıldığında $\hat{Y}$ **= 77,4252 - 0,2015 X regresyon modeli bulunmuştur.**

Hesaplanan bu 4 farklı M-tahmin edicisinin arasındaki fark, daha önce de belirtildiği gibi farklı tartılandırma fonksiyonlarının kullanılmasından ileri gelmektedir.

6. KAYNAKLAR

Birkes, D. ve Dodge,Y., (1993), "Alternative Methods of Regression", Wiley

Dodge, Y. ve Arthanari, T.S. , (1993) , "Mathematical Programming in Statistics", New York: Wiley

Faraway, J.,(2002), "Practical Regression and Anova using R", http://cran.r-project.org/, web book

Fox,J., (2002), "An S and S-Plus Companion to Applied Regression", Sage Publications

Genceli, M. ,(2001), Ekonometri ve İstatistik İlkeleri, İstanbul:Filiz Kitabevi

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

Hampel, F.R., (1975), "Beyond location parameters: robust concepts and methods. In International Statistical Institute, Proceedings of the 40th Session", 46: 375-391

Hoaglin, D. C., Mosteller, F., Tukey, J. W., (2000), "Understanding Robust and Exploratory Data Analysis", New York: Wiley

Huber, P. J. , (1981), "Robust Statistics", New York: John Wiley & Sons,Inc.

Koenker,R., ve Portnoy, S., (1987), "L-estimation for Linear Models", Journal of the American Statistical Assosiasion, 82:851-857

Neter, J., Kutner, M.H., Nachtsheim, C. J., Wasserman, W., (1993) "Applied Linear Statistical Methods", Wiley

Radhakrishna,C, (1999), "Linear models : least squares and alternatives", Springer Series in Statistics

Rousseeuw,P.,J, Leroy,A.,M, (2003), "Robust regression and outlier detection",Wiley

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

Weisberg, S. , (1985), "Applied Linear Regression", New York: Wiley

Yohai, V.J., ve Zamer, R.H., (1988)., "High breakdown point estimates of regressions by means of the minimizations of the efficient scale". Journal of the American Statistical Association, 83: 406-413

INTERNET KAYNAKLARI

[1]http://web.uccs.edu/lbecker/SPSS/explore1.htm

[2]http://godel.cs.bilgi.edu.tr/mirror/cran/

[3]http://www-sop.inria.fr/robotvis/personnel/zzhang/Publis/Tutorial-Estim/node24.html

[4]http://elections.fas.harvard.edu/wssmh/node11.html

[5]http://www.vision.ee.ethz.ch/~rpaget/Markov/Chapter_4/node70.html

[6]http://www.uni-muenster.de/ZIV/Mitarbeiter/BennoSueselbeck/s-html/shelp.html

7. EKLER

Ek 1 İki Noktadan Geçen Doğrunun Eğimi Formülü Kullanılarak Türkiye noktasının diğer noktalar ile bir doğru oluşturması durumunda oluşacak eğimlerin hesaplanması

Ek 2 Kalıntıların hesaplanması

Ek 3 Mutlak Sapmaların Hesaplanması

Ek 4 Daraltılmış sapmaların hesaplanması ve bu sapmalardan yola çıkılarak öngörülen y_i^* değerleri

Ek 5 Öngörülen y Değerleri ve $\hat{e}_i$ Kalıntıları

Ek 6 $\rho(ei)$ Değerleri ve Daraltılmış kalıntıların karelerinin hesaplanması

Ek 7 Daraltılmış kalıntılar ve daraltılmış kalıntıları kullanılarak öngörülen y* değerleri

Ek 8 Daraltılmış Kalıntılar ve $\rho(ei)$ Değerleri

Ek 9 R Project'te Komut Yazımı

Ek 1 İki Noktadan Geçen Doğrunun Eğimi Formülü Kullanılarak Türkiye noktasının diğer noktalar ile bir doğru oluşturması durumunda oluşacak eğimlerin hesaplanması

Çizelge Ek 1.1. Türkiye noktasının diğer noktalar ile bir doğru oluşturması durumunda oluşacak eğimlerin hesaplanması

Ülkeler	X	Y	$\frac{y_i - 69}{x_i - 45}$	$\lvert x_i - 45 \rvert$	$\sum \lvert x_i - 45 \rvert$
Endonezya	52	66	-0,429	7	7
Meksika	36	72	-0,333	9	16
Japonya	4	81	-0,293	41	57
Tunus	30	73	-0,267	15	72
Kanada	6	79	-0,256	39	111
Fransa	5	79	-0,250	40	151
İsveç	4	79	-0,244	41	192
Yunanıstan	7	78	-0,237	38	230
Belçika	6	78	-0,231	39	269
İspanya	6	78	-0,231	39	308
İtalya	6	78	-0,231	39	347
Tanzanya	152	45	-0,224	107	454
Mısır	54	67	-0,222	9	463
Norveç	4	78	-0,220	41	504
Zambiya	187	38	-0,218	142	646
Arjantin	22	74	-0,217	23	669
ABD	8	77	-0,216	37	706
İngiltere	6	77	-0,205	39	745
Almanya	5	77	-0,200	40	785
Finlandiya	5	77	-0,200	40	825
İrlanda	7	76	-0,184	38	863

Danimarka	**6**	**76**	**-0,179**	**39**	**902**
İran	**33**	**71**	**-0,167**	**12**	**914**
Portekiz	**6**	**75**	**-0,154**	**39**	**953**
Slovenya	**6**	**75**	**-0,154**	**39**	**992**
Çek Cumhuriyeti	**5**	**75**	**-0,150**	**40**	**1032**
Suudi Arabistan	**25**	**72**	**-0,150**	**20**	**1052**
Hindistan	**90**	**63**	**-0,133**	**45**	**1097**
Çin	**37**	**70**	**-0,125**	**8**	**1105**
Polonya	**10**	**73**	**-0,114**	**35**	**1140**
Slovakya	**10**	**73**	**-0,114**	**35**	**1175**
Azerbaycan	**21**	**71**	**-0,083**	**24**	**1199**
Pakistan	**126**	**63**	**-0,074**	**81**	**1280**
Bulgaristan	**17**	**71**	**-0,071**	**28**	**1308**
Estonya	**12**	**71**	**-0,061**	**33**	**1341**
Kolombiya	**28**	**70**	**-0,059**	**17**	**1358**
Macaristan	**10**	**71**	**-0,057**	**35**	**1393**
Romanya	**24**	**69**	**0,000**	**21**	**1414**
Rusya Federasyonu	**20**	**66**	**0,120**	**25**	**1439**
Brezilya	**40**	**67**	**0,400**	**5**	**1444**

Ek 2 Kalıntıların Hesaplanması

Çizelge Ek 2.1 Kalıntılar

	Ülkeler	Kalıntılar
1	Rusya Federasyonu	-8,385
2	Macaristan	-5,564
3	Estonya	-5,128
4	Romanya	-4,514
5	Bulgaristan	-4,039
6	Polonya	-3,564
7	Slovakya	-3,564
8	Azerbaycan	-3,168
9	Brezilya	-3,028
10	Çek Cumhuriyeti	-2,654
11	Kolombiya	-2,642
12	Portekiz	-2,436
13	Slovenya	-2,436
14	Danimarka	-1,436
15	Endonezya	-1,413
16	Suudi Arabistan	-1,296
17	İrlanda	-1,218
18	Çin	-0,681
19	Almanya	-0,654
20	Finlandiya	-0,654
21	Tanzanya	-0,625
22	İran	-0,553
23	İngiltere	-0,436

24	**Mısır**	**0,022**
25	**Arjantin**	**0,050**
26	**Türkiye**	**0,062**
27	**Norveç**	**0,128**
28	**Belçika**	**0,564**
29	**İspanya**	**0,564**
30	**İtalya**	**0,564**
31	**Yunanıstan**	**0,782**
32	**Tunus**	**0,793**
33	**Meksika**	**1,101**
34	**İsveç**	**1,128**
35	**Fransa**	**1,346**
36	**Kanada**	**1,564**
37	**Japonya**	**3,128**
38	**Hindistan**	**3,866**
39	**Pakistan**	**11,710**

Ek 3 Mutlak Sapmaların Hesaplanması

Çizelge Ek 3.1 Mutlak Sapmalar

Ülkeler	X	Y	Öngörülen $\hat{y}_i^0$	Sapmalar e_i^0	Mutlak Sapmalar
Türkiye	45	69	68,57225	0,42775	0,42775
ABD	8	77	75,64443	1,35557	1,35557
Almanya	5	77	76,21785	0,78215	0,78215
Arjantin	22	74	72,96847	1,03153	1,03153
Azerbaycan	21	71	73,15961	-2,15961	2,15961
Belçika	6	78	76,02671	1,97329	1,97329
Brezilya	40	67	69,52795	-2,52795	2,52795
Bulgaristan	17	71	73,92417	-2,92417	2,92417
Çek Cumhuriyeti	5	75	76,21785	-1,21785	1,21785
Çin	37	70	70,10137	-0,10137	0,10137
Danimarka	6	76	76,02671	-0,02671	0,02671
Endonezya	52	66	67,23427	-1,23427	1,23427
Estonya	12	71	74,87987	-3,87987	3,87987
Finlandiya	5	77	76,21785	0,78215	0,78215
Fransa	5	79	76,21785	2,78215	2,78215
Hindistan	90	63	59,97095	3,02905	3,02905
İngiltere	6	77	76,02671	0,97329	0,97329
İran	33	71	70,86593	0,13407	0,13407
İrlanda	7	76	75,83557	0,16443	0,16443
İspanya	6	78	76,02671	1,97329	1,97329
İsveç	4	79	76,40899	2,59101	2,59101
İtalya	6	78	76,02671	1,97329	1,97329
Japonya	4	81	76,40899	4,59101	4,59101

Kanada	**6**	**79**	**76,02671**	**2,97329**	**2,97329**
Kolombiya	**28**	**70**	**71,82163**	**-1,82163**	**1,82163**
Macaristan	**10**	**71**	**75,26215**	**-4,26215**	**4,26215**
Meksika	**36**	**72**	**70,29251**	**1,70749**	**1,70749**
Mısır	**54**	**67**	**66,85199**	**0,14801**	**0,14801**
Norveç	**4**	**78**	**76,40899**	**1,59101**	**1,59101**
Pakistan	**126**	**63**	**53,08991**	**9,91009**	**9,91009**
Polonya	**10**	**73**	**75,26215**	**-2,26215**	**2,26215**
Portekiz	**6**	**75**	**76,02671**	**-1,02671**	**1,02671**
Romanya	**24**	**69**	**72,58619**	**-3,58619**	**3,58619**
Rusya Federasyonu	**20**	**66**	**73,35075**	**-7,35075**	**7,35075**
Slovakya	**10**	**73**	**75,26215**	**-2,26215**	**2,26215**
Slovenya	**6**	**75**	**76,02671**	**-1,02671**	**1,02671**
Suudi Arabistan	**25**	**72**	**72,39505**	**-0,39505**	**0,39505**
Tanzanya	**152**	**45**	**48,12027**	**-3,12027**	**3,12027**
Tunus	**30**	**73**	**71,43935**	**1,56065**	**1,56065**
Yunanıstan	**7**	**78**	**75,83557**	**2,16443**	**2,16443**
Zambiya	**187**	**38**	**41,43037**	**-3,43037**	**3,43037**

Ek 4 Daraltılmış sapmaların hesaplanması ve bu sapmalardan yola çıkılarak öngörülen y_i^* değerleri

Çizelge Ek 4.1 Daraltılmış sapmalar ve Öngörülen y_i^* değerleri

Ülkeler	X	Y	$\hat{y}_i^0$	Sapmalar e_i^0	Mutlak sapmalar	Daraltılmış sapmalar e_i^*	y_i^*
Türkiye	45	69	68,572	0,428	0,428	0,428	69
ABD	8	77	75,644	1,356	1,356	1,356	77
Almanya	5	77	76,218	0,782	0,782	0,782	77
Arjantin	22	74	72,968	1,032	1,032	1,032	74
Azerbaycan	21	71	73,160	-2,160	2,160	-2,160	71
Belçika	6	78	76,027	1,973	1,973	1,973	78
Brezilya	40	67	69,528	-2,528	2,528	-2,528	67
Bulgaristan	17	71	73,924	-2,924	2,924	-2,924	71
Çek Cumhuriyeti	5	75	76,218	-1,218	1,218	-1,218	75
Çin	37	70	70,101	-0,101	0,101	-0,101	70
Danimarka	6	76	76,027	-0,027	0,027	-0,027	76
Endonezya	52	66	67,234	-1,234	1,234	-1,234	66
Estonya	12	71	74,880	-3,880	3,880	-3,880	71
Finlandiya	5	77	76,218	0,782	0,782	0,782	77
Fransa	5	79	76,218	2,782	2,782	2,782	79
Hindistan	90	63	59,971	3,029	3,029	3,029	63
İngiltere	6	77	76,027	0,973	0,973	0,973	77
İran	33	71	70,866	0,134	0,134	0,134	71
İrlanda	7	76	75,836	0,164	0,164	0,164	76
İspanya	6	78	76,027	1,973	1,973	1,973	78
İsveç	4	79	76,409	2,591	2,591	2,591	79

İtalya	**6**	**78**	**76,027**	**1,973**	**1,973**	**1,973**	**78**
Japonya	**4**	**81**	**76,409**	**4,591**	**4,591**	**4,390**	**80,799**
Kanada	**6**	**79**	**76,027**	**2,973**	**2,973**	**2,973**	**79**
Kolombiya	**28**	**70**	**71,822**	**-1,822**	**1,822**	**-1,822**	**70**
Macaristan	**10**	**71**	**75,262**	**-4,262**	**4,262**	**-4,262**	**71**
Meksika	**36**	**72**	**70,293**	**1,707**	**1,707**	**1,707**	**72**
Mısır	**54**	**67**	**66,852**	**0,148**	**0,148**	**0,148**	**67**
Norveç	**4**	**78**	**76,409**	**1,591**	**1,591**	**1,591**	**78**
Pakistan	**126**	**63**	**53,090**	**9,910**	**9,910**	**4,390**	**57,479**
Polonya	**10**	**73**	**75,262**	**-2,262**	**2,262**	**-2,262**	**73**
Portekiz	**6**	**75**	**76,027**	**-1,027**	**1,027**	**-1,027**	**75**
Romanya	**24**	**69**	**72,586**	**-3,586**	**3,586**	**-3,586**	**69**
Rusya Federasyonu	**20**	**66**	**73,351**	**-7,351**	**7,351**	**-4,390**	**68,961**
Slovakya	**10**	**73**	**75,262**	**-2,262**	**2,262**	**-2,262**	**73**
Slovenya	**6**	**75**	**76,027**	**-1,027**	**1,027**	**-1,027**	**75**
Suudi Arabistan	**25**	**72**	**72,395**	**-0,395**	**0,395**	**-0,395**	**72**
Tanzanya	**152**	**45**	**48,120**	**-3,120**	**3,120**	**-3,120**	**45**
Tunus	**30**	**73**	**71,439**	**1,561**	**1,561**	**1,561**	**73**
Yunanıstan	**7**	**78**	**75,836**	**2,164**	**2,164**	**2,164**	**78**
Zambiya	**187**	**38**	**41,430**	**-3,430**	**3,430**	**-3,430**	**38**

Ek 5 Öngörülen y Değerleri ve $\hat{e}_i$ Kalıntıları

Çizelge Ek 5.1 Öngörülen y Değerleri ve $\hat{e}_i$ Kalıntıları

Ülkeler	X	Y	Öngörülen $\hat{y}_i^0$	$\hat{e}_i$ Kalıntıları	Mutlak Kalıntılar
Türkiye	45	69	68,30726	0,69274	0,69274
ABD	8	77	75,78396	1,21604	1,21604
Almanya	5	77	76,39018	0,60982	0,60982
Arjantin	22	74	72,95494	1,04506	1,04506
Azerbaycan	21	71	73,15701	-2,15701	2,15701
Belçika	6	78	76,18810	1,81190	1,81190
Brezilya	40	67	69,31763	-2,31763	2,31763
Bulgaristan	17	71	73,96530	-2,96530	2,96530
Çek Cumhuriyeti	5	75	76,39018	-1,39018	1,39018
Çin	37	70	69,92385	0,07615	0,07615
Danimarka	6	76	76,18810	-0,18810	0,18810
Endonezya	52	66	66,89275	-0,89275	0,89275
Estonya	12	71	74,97567	-3,97567	3,97567
Finlandiya	5	77	76,39018	0,60982	0,60982
Fransa	5	79	76,39018	2,60982	2,60982
Hindistan	90	63	59,21399	3,78601	3,78601
İngiltere	6	77	76,18810	0,81190	0,81190
İran	33	71	70,73214	0,26786	0,26786
İrlanda	7	76	75,98603	0,01397	0,01397
İspanya	6	78	76,18810	1,81190	1,81190
İsveç	4	79	76,59225	2,40775	2,40775
İtalya	6	78	76,18810	1,81190	1,81190

Japonya	**4**	**81**	**76,59225**	**4,40775**	**4,40775**
Kanada	**6**	**79**	**76,18810**	**2,81190**	**2,81190**
Kolombiya	**28**	**70**	**71,74250**	**-1,74250**	**1,74250**
Macaristan	**10**	**71**	**75,37981**	**-4,37981**	**4,37981**
Meksika	**36**	**72**	**70,12592**	**1,87408**	**1,87408**
Mısır	**54**	**67**	**66,48861**	**0,51139**	**0,51139**
Norveç	**4**	**78**	**76,59225**	**1,40775**	**1,40775**
Pakistan	**126**	**63**	**51,93936**	**11,06064**	**11,06064**
Polonya	**10**	**73**	**75,37981**	**-2,37981**	**2,37981**
Portekiz	**6**	**75**	**76,18810**	**-1,18810**	**1,18810**
Romanya	**24**	**69**	**72,55079**	**-3,55079**	**3,55079**
Rusya Federasyonu	**20**	**66**	**73,35908**	**-7,35908**	**7,35908**
Slovakya	**10**	**73**	**75,37981**	**-2,37981**	**2,37981**
Slovenya	**6**	**75**	**76,18810**	**-1,18810**	**1,18810**
Suudi Arabistan	**25**	**72**	**72,34872**	**-0,34872**	**0,34872**
Tanzanya	**152**	**45**	**46,68547**	**-1,68547**	**1,68547**
Tunus	**30**	**73**	**71,33836**	**1,66164**	**1,66164**
Yunanıstan	**7**	**78**	**75,98603**	**2,01397**	**2,01397**
Zambiya	**187**	**38**	**39,61292**	**-1,61292**	**1,61292**

Ek 6 $\rho(ei)$ Değerleri ve daraltılmış kalıntıların karelerinin hesaplanması

Çizelge Ek 6.1 $\rho(ei)$ Değerleri ve daraltılmış kalıntıların kareleri

$\hat{e}_i$ Kalıntıları	Mutlak Kalıntılar	$\rho(ei)$	Daraltılmış Kalıntılar e_i*	$\left(e_i^*\right)^2$
0,69274	0,69274	0,47988	0,69274	0,479883
1,21604	1,21604	1,47876	1,21604	1,478757
0,60982	0,60982	0,37188	0,60982	0,371884
1,04506	1,04506	1,09215	1,04506	1,092153
-2,15701	2,15701	4,65270	-2,15701	4,652700
1,81190	1,81190	3,28297	1,81190	3,282966
-2,31763	2,31763	5,37140	-2,31763	5,371399
-2,96530	2,96530	8,79302	-2,96530	8,793022
-1,39018	1,39018	1,93259	-1,39018	1,932592
0,07615	0,07615	0,00580	0,07615	0,005799
-0,18810	0,18810	0,03538	-0,18810	0,035383
-0,89275	0,89275	0,79701	-0,89275	0,797010
-3,97567	3,97567	15,79604	-3,87620	15,024892
0,60982	0,60982	0,37188	0,60982	0,371884
2,60982	2,60982	6,81118	2,60982	6,811176
3,78601	3,78601	14,33390	3,78601	14,333900
0,81190	0,81190	0,65917	0,81190	0,659175
0,26786	0,26786	0,07175	0,26786	0,071750
0,01397	0,01397	0,00020	0,01397	0,000195
1,81190	1,81190	3,28297	1,81190	3,282966
2,40775	2,40775	5,79726	2,40775	5,797260

1,81190	1,81190	3,28297	1,81190	3,282966
4,40775	4,40775	19,14571	3,87620	15,024892
2,81190	2,81190	7,90676	2,81190	7,906758
-1,74250	1,74250	3,03631	-1,74250	3,036313
-4,37981	4,37981	18,92913	-3,87620	15,024892
1,87408	1,87408	3,51218	1,87408	3,512179
0,51139	0,51139	0,26152	0,51139	0,261521
1,40775	1,40775	1,98176	1,40775	1,981760
11,06064	11,06064	70,72148	3,87620	15,024892
-2,37981	2,37981	5,66351	-2,37981	5,663510
-1,18810	1,18810	1,41159	-1,18810	1,411592
-3,55079	3,55079	12,60813	-3,55079	12,608133
-7,35908	7,35908	42,02561	-3,87620	15,024892
-2,37981	2,37981	5,66351	-2,37981	5,663510
-1,18810	1,18810	1,41159	-1,18810	1,411592
-0,34872	0,34872	0,12161	-0,34872	0,121606
-1,68547	1,68547	2,84081	-1,68547	2,840811
1,66164	1,66164	2,76106	1,66164	2,761060
2,01397	2,01397	4,05607	2,01397	4,056069
-1,61292	1,61292	2,60152	-1,61292	2,601516

Ek 7 Daraltılmış hata payları ve daraltılmış kalıntılar kullanılarak öngörülen y* değerleri

Çizelge Ek 7.1 Daraltılmış kalıntılar ve öngörülen y* değerleri

Ülkeler	Y	e_i	Daraltılmış e_i*	y*
Türkiye	69	-2,65854	-2,65854	69,00000
ABD	77	5,34146	3,87621	75,53474
Almanya	77	5,34146	3,87621	75,53474
Arjantin	74	2,34146	2,34146	74,00000
Azerbaycan	71	-0,65854	-0,65854	71,00000
Belçika	78	6,34146	3,87621	75,53474
Brezilya	67	-4,65854	-3,87621	67,78233
Bulgaristan	71	-0,65854	-0,65854	71,00000
Çek Cumhuriyeti	75	3,34146	3,34146	75,00000
Çin	70	-1,65854	-1,65854	70,00000
Danimarka	76	4,34146	3,87621	75,53474
Endonezya	66	-5,65854	-3,87621	67,78233
Estonya	71	-0,65854	-0,65854	71,00000
Finlandiya	77	5,34146	3,87621	75,53474
Fransa	79	7,34146	3,87621	75,53474
Hindistan	63	-8,65854	-3,87621	67,78233
İngiltere	77	5,34146	3,87621	75,53474
İran	71	-0,65854	-0,65854	71,00000
İrlanda	76	4,34146	3,87621	75,53474
İspanya	78	6,34146	3,87621	75,53474
İsveç	79	7,34146	3,87621	75,53474
İtalya	78	6,34146	3,87621	75,53474
Japonya	81	9,34146	3,87621	75,53474
Kanada	79	7,34146	3,87621	75,53474
Kolombiya	70	-1,65854	-1,65854	70,00000

Macaristan	**71**	**-0,65854**	**-0,65854**	**71,00000**
Meksika	**72**	**0,34146**	**0,34146**	**72,00000**
Mısır	**67**	**-4,65854**	**-3,87621**	**67,78233**
Norveç	**78**	**6,34146**	**3,87621**	**75,53474**
Pakistan	**63**	**-8,65854**	**-3,87621**	**67,78233**
Polonya	**73**	**1,34146**	**1,34146**	**73,00000**
Portekiz	**75**	**3,34146**	**3,34146**	**75,00000**
Romanya	**69**	**-2,65854**	**-2,65854**	**69,00000**
Rusya Federasyonu	**66**	**-5,65854**	**-3,87621**	**67,78233**
Slovakya	**73**	**1,34146**	**1,34146**	**73,00000**
Slovenya	**75**	**3,34146**	**3,34146**	**75,00000**
Suudi Arabistan	**72**	**0,34146**	**0,34146**	**72,00000**
Tanzanya	**45**	**-26,65854**	**-3,87621**	**67,78233**
Tunus	**73**	**1,34146**	**1,34146**	**73,00000**
Yunanıstan	**78**	**6,34146**	**3,87621**	**75,53474**
Zambiya	**38**	**-33,65854**	**-3,87621**	**67,78233**

Ek 8 Daraltılmış kalıntılar ve $\rho(ei)$ değerleri

Çizelge Ek 8.1 Daraltılmış kalıntılar ve $\rho(ei)$ değerleri

Ülkeler	Y	e_i	Daraltılmış e_i*	$\left(e_i^*\right)^2$	$p(\hat{e}_i)$
Türkiye	69	-4,32376	-3,87621	15,02499	18,49458
ABD	77	3,67624	3,67624	13,51476	13,51476
Almanya	77	3,67624	3,67624	13,51476	13,51476
Arjantin	74	0,67624	0,67624	0,45730	0,457304
Azerbaycan	71	-2,32376	-2,32376	5,39985	5,39985
Belçika	78	4,67624	3,87621	15,02499	21,22719
Brezilya	67	-6,32376	-3,87621	15,02499	33,99941
Bulgaristan	71	-2,32376	-2,32376	5,39985	5,39985
Çek Cumhuriyeti	75	1,67624	1,67624	2,80979	2,809788
Çin	70	-3,32376	-3,32376	11,04737	11,04737
Danimarka	76	2,67624	2,67624	7,16227	7,162273
Endonezya	66	-7,32376	-3,87621	15,02499	41,75183
Estonya	71	-2,32376	-2,32376	5,39985	5,39985
Finlandiya	77	3,67624	3,67624	13,51476	13,51476
Fransa	79	5,67624	3,87621	15,02499	28,9796
Hindistan	63	-10,32376	-3,87621	15,02499	65,00907
İngiltere	77	3,67624	3,67624	13,51476	13,51476
İran	71	-2,32376	-2,32376	5,39985	5,39985
İrlanda	76	2,67624	2,67624	7,16227	7,162273
İspanya	78	4,67624	3,87621	15,02499	21,22719
İsveç	79	5,67624	3,87621	15,02499	28,9796
İtalya	78	4,67624	3,87621	15,02499	21,22719
Japonya	81	7,67624	3,87621	15,02499	44,48443
Kanada	79	5,67624	3,87621	15,02499	28,9796
Kolombiya	70	-3,32376	-3,32376	11,04737	11,04737

Macaristan	**71**	**-2,32376**	**-2,32376**	**5,39985**	**5,39985**
Meksika	**72**	**-1,32376**	**-1,32376**	**1,75233**	**1,752335**
Mısır	**67**	**-6,32376**	**-3,87621**	**15,02499**	**33,99941**
Norveç	**78**	**4,67624**	**3,87621**	**15,02499**	**21,22719**
Pakistan	**63**	**-10,32376**	**-3,87621**	**15,02499**	**65,00907**
Polonya	**73**	**-0,32376**	**-0,32376**	**0,10482**	**0,104819**
Portekiz	**75**	**1,67624**	**1,67624**	**2,80979**	**2,809788**
Romanya	**69**	**-4,32376**	**-3,87621**	**15,02499**	**18,49458**
Rusya Federasyonu	**66**	**-7,32376**	**-3,87621**	**15,02499**	**41,75183**
Slovakya	**73**	**-0,32376**	**-0,32376**	**0,10482**	**0,104819**
Slovenya	**75**	**1,67624**	**1,67624**	**2,80979**	**2,809788**
Suudi Arabistan	**72**	**-1,32376**	**-1,32376**	**1,75233**	**1,752335**
Tanzanya	**45**	**-28,32376**	**-3,87621**	**15,02499**	**204,5525**
Tunus	**73**	**-0,32376**	**-0,32376**	**0,10482**	**0,104819**
Yunanıstan	**78**	**4,67624**	**3,87621**	**15,02499**	**21,22719**
Zambiya	**38**	**-35,32376**	**-3,87621**	**15,02499**	**258,8195**

Ek 9 R Project'te Komut Yazımı

Veri setinin bulunduğu Excel dosyasını ".csv" uzantısıyla örneğin Z sürücüsüne "data.csv" ismiyle kaydettikten sonra,

data = read.table('z:\data.csv',header=T,sep',')

komutuyla dosya okutulur.

En Küçük Kareler yöntemini uygulamak için

> lm(y~x) komutu kullanılmaktadır.

> plot(y~x) komutu datanın x-y koordinat sistemindeki dağılımını,

> abline(rlm(y~x)) ise bu veri kümesinden geçen, En Küçük Kareler yöntemiyle bulunan doğruyu vermektedir.

M-tahmin edicileri hesaplamak için ise, robust kütüphaneye erişmek gerekmektedir. Bu işlem,

> library(MASS)

komutu yazılmak suretiyle yapılabilir.

Bundan sonra

> g <- rlm(y~x)

>summary(g)

komutları yazılırsa default'unda Huber tanımlandığından dolayı Huber-M tahmin edicileri verecektir.

> rlm(y~x, psi=psi.bisquare) komutu Tukey M-tahmin edicileri,

> rlm(y~x, psi=hampel) komutu Hampel M-tahmin edicileri,

> rlm(y~x, psi=psi.andrews) komutu Andrews M-tahmin edicileri verecektir.

8. ÖZGEÇMİŞ

Doğum Tarihi 11.01.1981

Doğum Yeri Rize

Lise	**1994-1998**	**Kadir Has Lisesi**
Lisans	**1998-2002**	**Yıldız Teknik Üniversitesi/ Fen-Edebiyat Fakültesi İstatistik Bölümü**
Lisans	**2000-2004**	**Yıldız Teknik Üniversitesi/ Kimya-Metalurji Fakültesi Matematik Mühendisliği Bölümü**
Yüksek Lisans	**2003-2006**	**Yıldız Teknik Üniversitesi/ Fen Bilimleri Enstitüsü İstatistik Yüksek Lisans**

Çalıştığı Kurumlar

2005-	**Anadolu Bankası**
2003-2004	**Genar Araştırma Eğitim Danışmanlık**
2002-2003	**Sosyal Araştırmalar Merkezi**