

"Robust Regresyon ve Uygulamaları"

"Barış ERGÜL"

YÜKSEK LİSANS TEZİ

İSTATİSTİK Anabilim Dalı

TEMMUZ-2006

"Robust Regression and Applications"

"Barış ERGÜL"

MASTER OF SCIENCE THESIS

Department of STATISTICS

JULY-2006

"Robust Regresyon ve Uygulamaları"

"Barış ERGÜL"

Eskişehir Osmangazi Üniversitesi

Fen Bilimleri Enstitüsü

Lisansüstü Yönetmeliği Uyarınca

İSTATİSTİK Anabilim Dalı

"UYGULAMALI İSTATİSTİK" Bilim Dalında

YÜKSEK LİSANS TEZİ

Olarak Hazırlanmıştır

Danışman: "Doç.Dr. Birdal ŞENOĞLU"

TEMMUZ-2006

"Barış ERGÜL" 'ün YÜKSEK LİSANS tezi olarak hazırladığı "Robust Regresyon ve Uygulamaları" başlıklı bu çalışma, jürimizce lisansüstü yönetmeliğinin ilgili maddeleri uyarınca değerlendirilerek kabul edilmiştir.

Üye : "Doç. Dr. Birdal ŞENOĞLU-DANIŞMAN"

Üye : "Doç. Dr. Şenol ERDOĞMUŞ-JÜRİ ÜYESİ"

Üye : "Yrd. Doç. Dr. Berna YAZICI-JÜRİ ÜYESİ"

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

Prof. Dr. Abdurrahman KARAMANCIOĞLU

Enstitü Müdürü

İÇİNDEKİLER

ÖZET.....	V
SUMMARY.....	VI
TEŞEKKUR.....	VII
TABLolar DİZİNİ.....	VIII
ŞEKİLLER DİZİNİ.....	IX
1. GİRİŞ.....	1
2. REGRESYON TEKNİKLERİ.....	3
2.1. En Küçük Kareler Tekniđi.....	3
2.2. Üç Grup için Dayanıklı Doğru Tekniđi.....	5
2.3. Wald Tekniđi.....	7
2.4. Nair ve Shrivastava Tekniđi.....	8
2.5. Bartlett Tekniđi.....	9
2.6. En Küçük Mutlak Sapma (LAD) Tekniđi.....	10
2.7. M-Regresyon Tekniđi.....	12
2.8. En Küçük Medyan Kare (LMS) Tekniđi.....	18
2.9. Winsorized En Küçük Kareler (WLS) Tekniđi.....	20
2.10. Budanmış En Küçük Kareler (LTS) Tekniđi.....	21
2.11. Theil Tekniđi.....	22
2.12. R-Regresyon Tekniđi.....	23
3. UYGULAMA.....	25
3.1. Göl Dibi Sıcaklığı ile Göl Geometri Oranı Arasındaki İlişkinin Regresyon Denklemleri Yardımıyla Belirlenmesi.....	25
3.2. Balık Yakalamak için Sarfedilen Çaba ile Klorofil α Arasındaki İlişkinin Regresyon Denklemleri Yardımıyla Belirlenmesi.....	35
3.3. Beton Kırma Deneyi Sırasında Açığa Çıkan Enerji Miktarı ile Darbe Deđeri Arasındaki İlişkinin Regresyon Denklemleri Yardımıyla Belirlenmesi.....	44
4. SONUÇ VE ÖNERİLER.....	52

KAYNAKLAR DİZİNİ.....54

EKLER

EK-1 Minnessota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum

Göl Dibi Sıcaklıkları.....58

EK-2 Iowa Göller Bölgesindeki Göllere ait CPUE_w ve Chlorophyll α Değerleri...59

EK-3 Beton Kırma Denemesine ait Enerji ve be3 değerleri.....60

EK-4 Minnessota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum

Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen Analiz Sonuçları (EKK, LAD, M-Regresyonu, LMS, WLS, LTS, R-Regresyonu için Bilgisayar Çıktıları).....61

ÖZET

Bu çalışmada, basit doğrusal regresyon analizinde yaygın olarak kullanılan EKK tekniğinin varsayımlarının sağlanmaması durumunda kullanılan robust regresyon teknikleri tanıtılmıştır ve gerçek hayat verileri için uygulamalar yapılmıştır.

Bu amaçla önce basit doğrusal regresyon analizine kısa bir giriş yapılmıştır. EKK tekniğinin varsayımları kısaca belirtildikten sonra robust regresyon teknikleri ayrıntılı olarak sunulmuştur. Robust regresyon teknikleri için parametre kestirim algoritmaları incelenmiştir.

Uygulama bölümünde, robust regresyon tekniklerinden Üç Grup Dayanımlı Doğru, Wald, Nair-Shrivastava, Bartlett, Brown-Mood, LAD, M-Regresyonu, LMS, WLS, LTS, Theil Regresyonu ve R-Regresyonuna ait regresyon denklemleri hesaplanmıştır.

Sonuç ve öneriler kısmında ise uygulama sonuçları tartışılmıştır.

SUMMARY

In this study, robust regression techniques are defined when the usual assumptions of the LS technique, used commonly in the analysis of simple regression, are not satisfied and an application of these techniques are given for the real life problems.

For this purpose, simple linear regression analysis is defined briefly and then the details of the robust regression techniques are given after determining the usual assumptions of LS technique. Algorithms for estimation of the model parameters are given for the robust regression techniques.

In the application part, regression equations for the robust techniques (Three Group Resistant Line, Wald, Nair and Shritastava, Bartlett, LAD, M-Regression, LMS, WLS, LTS, Theil and R-Regression) are given.

In the conclusion and suggestions part of the study, the results of the applications are discussed.

TEŐEKKÜR

Bu alıőmanın ortaya ıkmasında her tŒrlŒ olanađı sađlayan ve beni yŒnlendiren danıőman hocam Do. Dr. Birdal ŐENOĐLU' na sonsuz teőekkŒrlerimi sunarım.

alıőmanın her aőamasında bana destek veren Yrd. Do. Dr. Hatice FİDAN' a, Dr. Kamile ŐANLI' ya, Araő. GŒr. Arzu ALTIN' a ve diđer alıőma arkadaőlarıma teőekkŒrŒ bor bilirim.

TABLolar DİZİNİ

Tablo 1. Gözlem Değerlerinin Gruplandırılması.....	6
Tablo 2. Grup Sayısına Göre Tekniğin Seçimi.....	9
Tablo 3. Skor Fonksiyonları.....	24
Tablo 4. Uygulama 1 için Sonuç Tablosu.....	34
Tablo 5. Uygulama 2 için Sonuç Tablosu.....	43
Tablo 6. Uygulama 3 için Sonuç Tablosu.....	51

ŞEKİLLER DİZİNİ

- Şekil 1. Uygulama 1' deki veriler için elde edilen hata terimlerinin Q-Q grafiği ...27
- Şekil 2. Uygulama 2' deki veriler için elde edilen hata terimlerinin Q-Q grafiği ...36
- Şekil 3. Uygulama 3' deki veriler için elde edilen hata terimlerinin Q-Q grafiği ...45

1. GİRİŞ

Regresyon analizi, değişkenler arasındaki fonksiyonel ilişkiyi belirlemede en çok kullanılan tekniklerden biridir. Bu analiz, değişkenler arasındaki ilişkiyi belirleyerek parametre kestirimini sağlayan istatistiksel bir tekniktir (David, 1981).

Regresyon analizinin temelinde, gözlenen bir olayın fonksiyonel yapısının belirlenmesi yatmaktadır. Regresyon analizi yapılırken gözlem değerlerinin ve etkilenilen olayların bir matematiksel gösterimle yani bir fonksiyon yardımıyla ifade edilmesi gerekmektedir. Kurulan bu modele “Regresyon Modeli” denilmektedir.

Regresyon analizinin ilk kullanım alanı astronomi olmuştur. Legendre (1805) ve Gauss (1809), gezegenlerin yörüngelerini belirlemek amacıyla En Küçük Kareler (EKK) olarak bilinen tekniği oluşturmuşlardır. Konu ile ilgili bilinen değişkenlerden yola çıkarak bu değişkenler için bir regresyon modeli geliştirmişlerdir.

Yule (1899), İngiltere’deki yoksulluğun nedenlerinin araştırılması adlı çalışmasında bir regresyon denklemi hesaplamıştır. Yule’ nin bu yaklaşımı özellikle sosyal bilimler ve tıp alanlarında sıklıkla kullanılmıştır.

Regresyon analizinde yapılan çalışmalar Neyman (1923) tarafından devam ettirilmiştir. Freedman (1999)’ ın yaptığı inceleme sonucunda, Neyman’ ın çalışmalarının Rubin (1974) ve Holland (1988) tarafından devam ettirildikleri söylenmiştir. Bu konudaki yeni gelişmeler Pearl (1995) ve Angrist (1996)’ in çalışmalarında mevcuttur. Regresyon analizinin önemli çalışmaları Liu (1960), Lucas (1976), Manski (1995), Abbott (1997), Goldtorpe (1998) tarafından yapılmıştır.

Genel olarak regresyon analizinin amacı, elde edilen gözlem değerlerini kullanarak bir denklem elde etmektir. Bu denklem elde edilirken, yaygın olarak EKK tekniği kullanılır.

Regresyon analizinde yaygın kullanıma sahip EKK tekniği ile parametre tahminlerinin yapılabilmesi için temel bazı varsayımlar vardır. Bu varsayımlar genel olarak hata terimleri ile ilgilidir. Yani EKK tekniğinin varsayımları ε_i değerlerinin dağılımı hakkındadır (David, 1981).

Varsayım 1: ε_i hata terimi normal dağılıma sahiptir.

Varsayım 2: Farklı gözlemlerin rassal terimleri ($\varepsilon_i, \varepsilon_j$) birbirinden bağımsızdır. Yani herhangi bir ε_i teriminin başka bir ε_j ile olan ortak varyansı sifıra eşittir.

Varsayım 3: ε_i hata terimleri x_i bağımsız değişken değerlerinden bağımsızdır. Yani $\text{cov}(x_i, \varepsilon_i) = 0$ 'dır.

Özetle, hata terimleri ε_i bağımsız ve özdeş, ortalaması sıfır, varyansı σ^2 olan normal dağılıma sahiptir. Bu varsayımların bozulması durumunda tutarlı olmayan ve yanlış tahminlere yol açılabilir (Mosteller ve Tukey, 1977).

Ancak hata terimleri dağılımı normal dağılıma sahip değilse, uygulanacak teknikler, robust regresyon teknikleridir. Robust regresyon teknikleri, aykırı değerlerden etkilenmeyen tekniklerdir.

Varsayımlara bağlı olmayan, özellikle normallik varsayımına duyarsız yaklaşımlar “robust (sağlam)” olarak adlandırılmıştır (Öztürk, 2003).

Gözlem değerlerinin normallik varsayımını bozan en önemli etken büyük hata terimi olarak nitelendirilen aykırı değerlerdir. Aykırı değerler, istatistiksel tekniğe ve kurulan regresyon modeline zarar verirler. Bununla birlikte, aykırı değer modelini açıklayan bir değer olması olasılığı çok düşüktür. Eğer model değiştirilirse, bu aykırı değer etkisi de büyük ölçüde azalmaktadır. Robust regresyon teknikleri, aykırı değere sahip gözlem değerini güncelleyerek, analizi bu yeni güncellenmiş değer ile yapmaktır.

Hata terimlerinin normal dağılmaması EKK tekniğinin uygulanmasını zorlaştırır. Bu durumda alternatif regresyon teknikleri daha uygun sonuçlar vermektedir.

Robust regresyonun kullanılma amaçlarından biri, gözlem değerlerinin önemli ve büyük bir kısmını kullanarak en iyi kestirimi yapmaktır. Aynı zamanda aykırı gözlemlerden meydana gelecek sorunları da önleyen tekniklerdir.

Bu tez çalışmasında günlük yaşamda karşılaşılan ve regresyon tekniği ile analizi yapılan gerçek yaşam problemleri, klasik EKK ve alternatif regresyon teknikleri ile çözüme ulaştırılacaktır. İkinci bölümde, çeşitli alternatif regresyon teknikleri ayrıntılı olarak incelenmiştir. Ayrıca bu bölümde parametre kestirimlerinin yapılabilmesi için gerekli algoritmalar üzerinde durulmuştur. Üçüncü bölümde, çalışmada ele alınan regresyon tekniklerine ait uygulamalara yer verilmiştir. Dördüncü ve son bölümde ise uygulama sonuçları tartışılmıştır.

1. REGRESYON TEKNİKLERİ

Değişkenler arasındaki sebep-sonuç ilişkisi, bağımlı ve bağımsız değişken(ler) yardımıyla açıklanmaktadır. Regresyon analizi yapılırken kurulan matematiksel modelde yer alan değişkenler bir bağımlı değişken ve bir veya birden çok bağımsız değişkenden oluşmaktadır. Kurulacak modelde bağımsız değişken sayısı bir adet ise, basit doğrusal regresyonun, bağımsız değişken sayısı bir adetten fazla ise, çoklu doğrusal regresyonun konusunu oluşturmaktadır. Bu matematiksel modeli çözmeye yarayan çeşitli teknikler vardır. Bu teknikler aşağıda izleyen bölümlerde ayrıntılı biçimde incelenecektir.

2.1. En Küçük Kareler Tekniği

Bu çalışmada basit doğrusal regresyon inceleneceğinden, ana kütle için basit doğrusal regresyon denklemi,

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, i = 1, 2, \dots, n \quad (2.1.1)$$

şeklinde belirtilir.

- n ; gözlem sayısını,
- y_i ; bağımlı değişken için gözlem değerlerini,
- x_i ; bağımsız değişken için gözlem değerlerini,
- β_0 ; doğrunun y eksenini kestiği noktayı,
- β_1 ; doğrunun eğimini,
- ε_i ; hata değerlerini göstermektedir.

Uygulamada β_0 ve β_1 değerleri bilinmiyorsa, evrenden çekilecek örneklem için bunların tahmini değerleri olan $\hat{\beta}_0$ ve $\hat{\beta}_1$ kullanılarak (2.1.1) eşitliği aşağıdaki gibi yazılır.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i, i = 1, 2, \dots, n \quad (2.1.2)$$

Kısaca regresyon analizinin varsayımlarına geri dönecek olursak, hata terimleri ε_i bağımsız ve özdeş dağılıma sahip ortalaması sıfır, varyansı σ^2 olan normal dağılıma sahiptir. Bu bilgi kullanılarak EKK parametre kestirimleri aşağıdaki gibi yapılır.

Anakütle için regresyon denklemi (2.1.1) ve örneklem için regresyon denklemi tahmini (2.1.2)' de gösterildiği gibidir.

Burada hata terimleri, gerçek gözlem değerleri ile tahmin edilen regresyon denklemine göre oluşturulan gözlem değerleri arasındaki farka eşittir. Yani $\varepsilon_i = y_i - \hat{y}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$ ' dir.

Bu farktan yararlanarak,

$$f(\varepsilon_i) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(y_i - \beta_0 - \beta_1 x_i)^2\right\} \quad (2.1.3)$$

yazarak ve

$$L = \frac{1}{\sqrt{2\pi}} \prod_{i=1}^n \exp\left\{-\frac{1}{2}(y_i - \beta_0 - \beta_1 x_i)^2\right\}$$

olabilirlik fonksiyonu tanımı kullanılarak bu fonksiyon minimize edilmeye çalışıldığında hata terimlerinin kareleri minimize edilmiş olur.

$$\ln L = -\frac{n}{2} \ln 2\pi - \frac{1}{2} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 \quad (2.1.4)$$

Bu denklemleri çözmek için $\ln L$ fonksiyonunun birinci türevleri alınarak sıfıra eşitlenmesi gerekmektedir (Ertek, 1983).

$$\begin{aligned} \frac{\partial \ln L}{\partial \beta_0} &= \frac{\partial \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2}{\partial \beta_0} = 0 \Rightarrow \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) = 0 \\ \frac{\partial \ln L}{\partial \beta_1} &= \frac{\partial \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2}{\partial \beta_1} = 0 \Rightarrow \sum_{i=1}^n (y_i x_i - \beta_0 x_i - \beta_1 x_i^2) = 0 \end{aligned} \quad (2.1.5)$$

(2.1.5) eşitliğini çözersek,

$$\hat{\beta}_0 = \frac{\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)}{n} = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n x_i (y_i - \bar{y})}{\sum_{i=1}^n x_i (x_i - \bar{x})} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.1.6)$$

bulunur (David, 1981).

Kurulan regresyon modelindeki gözlem değerlerinin modele uyumuna belirlilik katsayısı ile bakılır.

$$R^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n \varepsilon_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.1.7)$$

Bulunan belirlilik katsayısı, bağımlı değişkendeki toplam değişimin ne kadarının (% olarak) bağımsız değişken tarafından açıklandığını gösterir. Değer 0 ile 1 arasında değişmektedir (Gujarati, 1999).

1.2. Üç Grup için Dayanıklı Doğru Tekniği

Veri setinin özetlenmesinde en basit dayanıklı yaklaşım tekniği, örneklem medyanının hesaplanmasıdır. Bu temel fikirden yola çıkılarak, veri seti üç gruba bölünür. Bu üç grubun her biri için özetleyici veri noktası belirlenir. Daha sonra bu özetleyici veri noktaları kullanılarak regresyon doğrusu belirlenir.

Öncelikle veri setindeki bağımsız değişken değerleri sıralanır. Bu sıralanan değerler n adet gözlem değerine sahiptir. Bu n birimlik gözlem seti üç gruba bölünür (sol grup, merkez grup, sağ grup). Aşağıdaki tablo gözlem değerlerinin nasıl bölüneceği hakkında yol gösterici bilgiye sahiptir. n , gözlem sayısını ve k , gruplardaki gözlem sayısını belirtir.

Tablo 1. Gözlem Değerlerinin Gruplandırılması

Grup	$n = 3k$	$n = 3k + 1$	$n = 3k + 2$
Sol(Left)	k	k	$k + 1$
Merkez(Middle)	k	$k + 1$	k
Sağ(Right)	k	k	$k + 1$

1. Özetleyici noktaların seçimi ise aşağıdaki gibi yapılır. Öncelikle veri setindeki bağımsız değişken değerleri sıralanır. Üç gruba ait özetleyici noktalar $(x_L, y_L), (x_M, y_M), (x_R, y_R)$ olarak tanımlanır.

2. Daha sonraki adımda üç gruba bölünen veri setinde, sol bölümdeki x değerleri için medyan değeri bulunur ve aynı şekilde sol bölümdeki y değerleri için medyan değeri bulunur. Böylece (x_L, y_L) özetleyici noktası bulunmuş olur. Aynı biçimde $(x_M, y_M), (x_R, y_R)$ noktaları da hesaplanır.

3. Daha sonraki aşamada ise özetleyici noktalardan yararlanılarak regresyon parametreleri hesaplanır. $\hat{\beta}_0$ ve $\hat{\beta}_1$ parametre değerlerinin hesaplanması sırasında medyan değeri kullanıldığı için parametreler EKK tekniğine göre daha dayanıklı hale gelir.

Regresyon doğrusunun başlangıç eğimi ve sabit terimi

$$\hat{\beta}_1 = \frac{y_R - y_L}{x_R - x_L} \quad \text{ve} \quad (2.2.1)$$

$$\hat{\beta}_0^* = \frac{1}{3} [(y_L - \beta_1(x_L - x_M)) + y_M + (y_R - \beta_1(x_R - x_M))] \quad (2.2.2)$$

formülleri yardımıyla bulunur.

4. Sonraki adımda, her bir gözlem noktası için hata terimi değerleri aşağıdaki formül yardımıyla hesaplanır.

$$\varepsilon_i = y_i - [\beta_0^* + \beta_1(x_i - x_M)] \quad , \quad i = 1, 2, \dots, n$$

5. Hata terimleri değerlerinden yararlanılarak iterasyon yapılır. Başlangıçtaki hata terimleri aşağıdaki formül yardımıyla hesaplanır:

$$\varepsilon_i^{(0)} = y_i - [\beta_0^* + \beta_1(x_i - x_M)] \quad , \quad i = 1, 2, \dots, n$$

$$\text{Buradan } \delta_1 = \frac{\varepsilon_R - \varepsilon_L}{x_R - x_L} \quad \text{ve} \quad \gamma_1 = \frac{1}{3} [(\varepsilon_L - \delta_1(x_L - x_M)) + \varepsilon_M + (\varepsilon_R - \delta_1(x_R - x_M))]$$

formülleri kullanılarak regresyon doğrusu için yeni eğim ve sabit terim parametreleri

$\hat{\beta}_1^{(1)} = \hat{\beta}_1 + \delta_1$ ve $\hat{\beta}_0^{(1)} = \hat{\beta}_0 + \gamma_1$ eşitlikleri yardımıyla hesaplanır. Daha sonra yeni hata terimleri hesaplanır,

$$\varepsilon_i^{(1)} = \varepsilon_i^{(0)} - [\gamma_1 + \delta_1(x_i - x_M)] \quad , \quad i = 1, 2, \dots, n.$$

Bulunan yeni hata terimleri kullanılarak δ_2 ve γ_2 değerleri hesaplanarak yeni eğim ve sabit terim parametreleri $\hat{\beta}_1^{(2)} = \hat{\beta}_1 + \delta_2$ ve $\hat{\beta}_0^{(2)} = \hat{\beta}_0 + \gamma_2$ eşitlikleri yardımıyla revize edilir.

6. Parametre değerleri birbirine yaklaşıncaya dek iterasyon sürer (Hoaglin v.d., 1983).

2.3. Wald Tekniği

Wald (1940) tarafından önerilen bu teknikte, n birimlik gözleme sahip veri seti iki gruba ayrılır. İlk grubun gözlem sayısı, ikinci gruptan küçük olmalıdır. $m = n/2$ seçilerek n birimlik gözlem seti iki parçaya ayrılır. Bağımsız değişken değerleri sıralanır.

$(x_1, y_1), \dots, (x_m, y_m)$ ve $(x_{m+1}, y_{m+1}), \dots, (x_n, y_n)$ şeklinde iki parçaya ayrılan veri seti için parametre kestirim değerleri aşağıdaki formüller yardımıyla bulunur (Hoaglin v.d., 1983);

$$\hat{\beta}_1 = \frac{(y_{m+1} + \dots + y_n) - (y_1 + \dots + y_m)}{(x_{m+1} + \dots + x_n) - (x_1 + \dots + x_m)}, \quad (2.3.1)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} . \quad (2.3.2)$$

Burada $\bar{y} = \frac{(y_1 + \dots + y_n)}{n}$ ve $\bar{x} = \frac{(x_1 + \dots + x_n)}{n}$ ile hesaplanır. Bu teknikte, eğer $x_{m+1} = x_m$ ise bu değerlerden bir tanesi atıldıktan sonra veri seti gruplara ayrılır. Bu durumda n değerinin tek veya çift olmasından etkilenmez.

2.4. Nair ve Shrivastava Tekniği

EKK tekniğine bir diğer alternatif de Nair ve Shrivastava (1942)' nin grup ortalamaları olarak bilinen tekniğidir.

Bu teknikte, öncelikle bağımsız değişken değerleri sıralanır. Gözlem değerleri üç gruba ayrılır. Gruplar “Üç grup için dayanıklı doğru tekniği” kısmında verilen Tablo.1' e bakılarak oluşturulur. İlk n_L noktasına sahip gözlem çiftleri ilk grubu ve son n_U noktasına sahip gözlem çiftleri de ikinci grubu oluşturur. Diğer gözlem çiftleri ise merkezdeki grubu oluşturmak üzere, parametre kestirim değerleri aşağıdaki formüller yardımıyla hesaplanır (Hoaglin v.d., 1983).

$$\bar{x}_L = \frac{x_1 + \dots + x_{n_L}}{n_L}, \quad \bar{y}_L = \frac{y_1 + \dots + y_{n_L}}{n_L}$$

$$\bar{x}_U = \frac{x_{n-n_U+1} + \dots + x_n}{n_U}, \quad \bar{y}_U = \frac{y_{n-n_U+1} + \dots + y_n}{n_U}$$

Bu değerler hesaplandıktan sonra parametre kestirim değerleri aşağıdaki formül yardımıyla bulunur,

$$\hat{\beta}_1 = \frac{\bar{y}_U - \bar{y}_L}{\bar{x}_U - \bar{x}_L} \text{ ve} \quad (2.4.1)$$

$$\hat{\beta}_0 = \bar{y}_L - \hat{\beta}_1 \bar{x}_L = \bar{y}_U - \hat{\beta}_1 \bar{x}_U . \quad (2.4.2)$$

2.5. Bartlett Tekniđi

Wald tekniđine gre daha etkin sonular elde etmek amacıyla Bartlett (1949) tarafından geliřtirilmiř bir tekniktir. Nair ve Shritastava tekniđindeki gibi u grup ile alıřılır.

Parametre kestirimleri ařađıdaki gibi bulunur:

$$\hat{\beta}_1 = \frac{\bar{y}_U - \bar{y}_L}{\bar{x}_U - \bar{x}_L} , \quad (2.5.1)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} . \quad (2.5.2)$$

Eřitlik (2.5.1) ve (2.5.2)' den anlařılacađı gibi Wald tekniđinden en nemli farkı, bu tekniđin iki grup yerine u grup kullanmasıdır.

Ařađıdaki tabloda hangi tekniđin ne zaman uygulanacađı ait bir tablo gsterilmektedir (Hoaglin v.d., 1983).

Tablo 2. Grup Sayısına Gre Tekniđin Seimi

	zetleyici	İstatistik
Grup Sayısı	Ortalama	Medyan
İki	Wald	Brown-Mood
	Nair ve Shrivastava, Bartlett	 grup iin dayanıklı dođru

2.6. En Küçük Mutlak Sapma (Least Absolute Deviation-LAD) Tekniđi

LAD tekniđi EKK tekniđinden 50 yıl önce Boscovich tarafından ortaya atılmıřtır (Birkes ve Dodge, 1993). Bu teknik daha sonra Laplace tarafından yeniden ele alınmıř ve güncellenerek kullanılmaya bařlanmıřtır (Nevitt ve Tam, 1998).

Bu teknik, EKK tekniđine alternatif olması nedeniyle sıklıkla kullanılır.

LAD, mutlak hatalar toplamını minimize eden bir tekniktir ve ařađıdaki biçimde tanımlanır;

$$\min \sum_{i=1}^n |y_i - \beta_0 - \beta_1 x_i| . \quad (2.6.1)$$

(2.1.1) eřitliđinde tanımlanan basit dođrusal regresyon denkleminde $\hat{\beta}_0$ ve $\hat{\beta}_1$ parametre tahmin deđerleri EKK tekniđi ile bulunurken hata kareler toplamını minimum yapan parametre deđerleri seçilir. LAD tekniđinde ise hata terimlerinin mutlak deđerlerinin toplamını minimum yapan parametre deđerleri seçilir (Tasker, 2000).

LAD tekniđin amacı, mutlak sapmalar toplamını minimum yapan regresyon dođrusunun seçimidir. Bu dođru ise herhangi bir (x_0, y_0) veri noktasından geçen dođrular arasından en iyi sonucu veren dođru olacaktır.

LAD tekniđi için regresyon dođrusu belirlenmesinde ařađıdaki ařamalar izlenir:

1. Gözlem çiftlerinden bir tanesi rassal olarak seçilir. Seçilen bu gözlem çifti genel olarak ilk sırada yer alan gözlem çiftidir.

2. Seçilen gözlem çifti kullanılarak tek tek her bir gözlem çifti için eđim deđerleri ve bu deđgerlere karřılık gelen $x_i - x_0$ deđerleri hesaplanır.

3. Küçükten büyüđe dođru sıralanan eđim deđerlerine karřılık gelen $x_i - x_0$ deđerlerinin mutlak deđerleri alınır.

4. Bulunan $|x_i - x_0|$ deđerlerinin birikimli toplamı bulunur.

5. Kritik deđer (P) hesaplanır. Bu deđer bir önceki ařamada elde edilen birikimli toplamın yarısına eřittir.

6. Kritik deđere karřılık gelen eđim deđerinin tespiti için 3. ařamadaki gözlem deđerlerine bakılır. Kritik deđerden büyük olan ilk gözlem deđeri aranan noktadır. Bu

değere karşılık gelen değerın eğim değerine bakılır. Bu değer 3. aşama esnasında bulunan değerdir. Böylece regresyon denkleminin eğim parametresi bulunmuş olur.

7. Bu eğim değerini veren noktanın orijinal sırası tespit edilir. Bu nokta bir sonraki aşama için yeni başlangıç noktasıdır.

8. Yeni gözlem çifti kullanılarak yukarıdaki aşamalar tekrar edilir. İterasyonlar sonucunda arka arkaya iki adet aynı eğim değeri bulununca bu aşamalar durdurulur. Böylece aranan eğim değeri bulunmuş olur.

Kritik değer,

$$P = \sum_{i=1}^n |x_i - x_0|$$

olmak üzere;

$$|x_1 - x_0| + |x_2 - x_0| + \dots + |x_{k-1} - x_0| < \frac{P}{2} \quad (2.6.2)$$

$$|x_1 - x_0| + |x_2 - x_0| + \dots + |x_{k-1} - x_0| + |x_k - x_0| > \frac{P}{2} \quad (2.6.3)$$

koşullarını sağlayan k noktası bulunur. Buna göre (x_0, y_0) noktasından geçen en iyi doğru;

$$\hat{\beta}_1 = \frac{y_k - y_0}{x_k - x_0} \quad \text{ve} \quad \hat{\beta}_0 = y_0 - \hat{\beta}_1 x_0 \quad (2.6.4)$$

olmak üzere $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ olarak belirlenir (Yorulmaz, 2003).

LAD tekniği, EKK tekniğinden çok daha önce bulunmasına karşın LAD tekniğinin yapısı ve algoritması uzun süren işlemler gerektirdiğinden dolayı uygulama alanı bulamamıştır. Veri seti içerisinde x-yönünde aykırı değerlerin varlığında tekniğin performansı oldukça azalmaktadır (Öztürk, 2003).

2.7. M-Regresyon Tekniđi

Bu bölümde, yaygın olarak kullanılan Huber, Hampel, Andrews ve Tukey M teknikleri üzerinde durulacaktır.

Hata terimleri ε_i bağımsız ve özdeş, ortalaması sıfır, varyansı σ^2 olan normal dağılıma sahip olduğunda EKK tekniđi oldukça iyi sonuçlar vermektedir. Ancak hata terimlerinin anakütle için normallik varsayımı geçersiz olduğunda EKK tekniđinin uygulanması doğru olamaz. M-Regresyon bu varsayıma uygun olarak geliştirilmiştir.

EKK tekniđinde ε_i^2 değerlerinin toplamı minimize edilirken, LAD tekniđinde hata terimlerinin mutlak değerler toplamı minimize edilir. Huber'ın M-Regresyon tekniđinde ise bir orta yol bulunarak hata terimlerinin fonksiyonu olan $\rho(z)$ minimize edilir.

Huber'ın oluşturduğu modelde, ε_i hata terimini, sıfıra uzak değer aldığıında $\rho(z)$ fonksiyonunun değeri $|\varepsilon_i|$ 'ye; ε_i hata terimi sıfıra yakın değer aldığıında ise ε_i^2 'ye eşitleyerek daha uzlaştıracı bir teknik öne sürmüştür (Tasker, 2000).

Buna göre Huber (1973)'in önerdiği teknikte $\rho(z)$ fonksiyonu hata terimleri için tanımlandığıında,

$$\rho(\varepsilon) = \begin{cases} \varepsilon^2 & , \quad -k \leq \varepsilon \leq k \text{ ise} \\ 2k|\varepsilon| - k^2 & , \quad \varepsilon < -k \text{ veya } k < \varepsilon \text{ ise} \end{cases} \quad (2.7.1)$$

biçiminde dönüşür (Jabr, 2005). Burada $k=1,5*MSM$ alınır ve aşağıdaki formül yardımıyla hesaplanır.

$$MSM = \frac{\text{median}\{|\varepsilon_i - \text{median}(\varepsilon_i)|\}}{0,6745} .$$

Aşağıdaki adımları izleyerek Huber'in M-Regresyonunua ait parametre tahminleri bulunur (Yorulmaz, 2003).

1. $\hat{\beta}_0$ ve $\hat{\beta}_1$ tahmin değerleri EKK tekniği yardımı ile hesaplanır.
2. Bu tahmin değerleri kullanılarak MSM ve ε_i değerleri bulunur.
3. Bulunan ε_i değerleri üzerinden $\rho(\varepsilon)$ fonksiyonu kullanılarak ε_i^* düzeltilmiş değerleri bulunur.
4. ε_i^* değerleri kullanılarak y_i^* değerleri bulunur ve EKK ile $\hat{\beta}_{00}$ ve $\hat{\beta}_{10}$ tahmin değerlerini hesaplanır.
5. $\hat{\beta}_0$ ve $\hat{\beta}_1$ için güncellenen tahminler ile bir önceki tahmin değerleri karşılaştırılır.
6. Tahminler arasındaki fark 0,001' den küçükse işlem bitirilir.
7. Değilse $\hat{\beta}_{00}$ ve $\hat{\beta}_{10}$ değerleri $\hat{\beta}_0$ ve $\hat{\beta}_1$ olarak atanır ve Adım 2' ye dönlür.

Huber'in M-Regresyon tekniğinde parametre tahminleri yukarıdaki algoritma kullanılarak bulunabileceği gibi Huber ağırlık fonksiyonu yardımıyla da bulunabilir.

Ağırlıklandırılmış EKK tekniği, regresyon analizinde kullanılan bazı gözlem değerlerinin diğerlerine göre oldukça farklı olmasından yola çıkılarak ortaya atılmıştır. Bu teknikte her bir gözlem değerine farklı ağırlıklar vermek suretiyle, aykırı değerlerin model üzerindeki olumsuz etkisinin yok edilmesi amaçlanır. Böyle bir durumda varyansların sabit olmasını sağlamak için bu duruma neden olan değişken/değişkenlere çeşitli dönüşümler uygulanır. Bu dönüşüm değerleri ağırlık olarak adlandırılır (Gujarati, 1999).

Robust regresyon tekniklerinden bazıları ağırlıklandırılmış EKK yöntemini kullanırlar. Robust regresyon tekniklerinde, ağırlıklandırılmış EKK kullanıldığında ağırlıklar hata teriminin ters bir fonksiyonudur ki bu da şu özellikleri sağlar:

1. Gözlem değerleri ağırlık fonksiyonu tarafından belirlenen aralık içindeyse ağırlık değeri 1 olacaktır.
2. Gözlem değerleri ağırlık fonksiyonu tarafından belirlenen aralık dışındaysa ağırlık değeri 1 değerinden farklı olacaktır.
3. $\sum_{i=1}^n \varepsilon_i^2$ ($i = 1, 2, \dots, n$) değerlerinin toplamı EKK ile minimize edilecektir.

w_i ağırlıkları hata teriminin bir fonksiyonudur. Bu fonksiyon aşağıdaki özelliği sağlamış olur.

$$\sum_{i=1}^n w_i (y_i - \beta_0 - \beta_1 x_i)^2 = \min$$

Parametre tahmin değerleri ise aşağıdaki gibi bulunur (Gujarati, 1999);

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n w_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n w_i (x_i - \bar{x})^2} \text{ ve} \quad (2.7.2)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} . \quad (2.7.3)$$

Burada,

$$\bar{x} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} \text{ ve } \bar{y} = \frac{\sum_{i=1}^n w_i y_i}{\sum_{i=1}^n w_i} \quad (2.7.4)$$

olarak alınır.

EKK genel işleyişinde hata kareler toplamını minimize ederken tüm hata değerlerine ağırlık değeri olarak 1 değeri verilir.

M-Regresyon tekniği, hata terimlerinin kareleri toplamını minimum yapmak yerine hata terimlerinin fonksiyonunu minimize eder. Standartlaştırılmış hata terimi

$r = \frac{\varepsilon_i}{MSM}$ olarak tanımlanır. Standartlaştırılmış hata terimleri için literatürde yaygın

olarak kullanılan bazı $\rho(r)$ fonksiyonları aşağıda gösterildiği gibidir.

Huber' in fonksiyonu

Huber tarafından standartlaştırılmış hata terimleri için önerilen fonksiyon,

$$\rho(r) = \begin{cases} \frac{r^2}{2} & , -k \leq r \leq k \\ k|r| - \frac{k^2}{2} & , r < -k \text{ veya } k < r \end{cases} \quad (2.7.5)$$

biçimine dönüşür. Burada $k=1,5$ değerini alır.

(2.7.5) eşitliğinin türevi alınarak,

$$\psi(r) = \begin{cases} -k & , r < -k \\ r & , |r| \leq k \\ k & , r > k \end{cases} \quad (2.7.6)$$

fonksiyonu elde edilir. Ağırlık fonksiyonu ise ψ fonksiyonunun r ' ye bölünmesiyle elde edilir. Buna göre Huber'in ağırlık fonksiyonu aşağıdaki biçimde tanımlanır (Şanlı, 2005).

$$w = \begin{cases} 1 & , |r| \leq k \\ \frac{k}{r} & , |r| > k \end{cases} \quad (2.7.7)$$

Parametre tahminlerinin yapılabilmesi için kullanılacak adımlar aşağıda gösterilmiştir (Fox, 2002).

1. $\hat{\beta}_0$ ve $\hat{\beta}_1$ tahmin değerleri EKK tekniği ile hesaplanır.
2. Bu tahminleri kullanarak ε_i değerleri bulunur.
3. MSM değeri hesaplanır.
4. Ağırlık değerleri hesaplanır.
5. Ağırlıklı EKK tekniği ile $\hat{\beta}_{00}$ ve $\hat{\beta}_{10}$ tahmin değerleri hesaplanır.
6. $\hat{\beta}_{00}$ ve $\hat{\beta}_{10}$ ile $\hat{\beta}_0$ ve $\hat{\beta}_1$ değerleri karşılaştırılır.
7. Tahminler arasındaki fark 0,001'den küçük ise işlem bitirilir.

8. Değilse $\hat{\beta}_{00}$ ve $\hat{\beta}_{10}$ değerleri $\hat{\beta}_0$ ve $\hat{\beta}_1$ olarak atanır ve Adım 2' ye dönülür.

Hampel' in fonksiyonu

Hampel (1971) tarafından öne sürülen fonksiyon,

$$\rho(r) = \begin{cases} \frac{r^2}{2} & , \quad 0 < |r| \leq a \\ a|r| - \frac{a^2}{2} & , \quad a < |r| \leq b \\ -\frac{a}{2(c-b)}(c-r)^2 + \frac{a}{2}(b+c-a) & , \quad b < |r| \leq c \\ \frac{a}{2}(b+c-a) & , \quad c < |r| \end{cases} \quad (2.7.8)$$

biçimindedir.

(2.7.8) eşitliğinin türevi alınır, bulunan fonksiyon r ' ye bölünür ve ağırlık fonksiyonu elde edilir. Buna göre Hampel'in ağırlık fonksiyonu aşağıdaki biçimde tanımlanır,

$$w = \begin{cases} 1 & , \quad 0 < |r| \leq a \\ \frac{a}{r} \text{sgn}(r) & , \quad a < |r| \leq b \\ \frac{a}{r} \left[\frac{c-|r|}{c-b} \right] \text{sgn}(r) & , \quad b < |r| \leq c \\ 0 & , \quad c < |r| \end{cases} \quad (2.7.9)$$

Genellikle sabitlerin değerleri $a = 1,7$, $b = 3,4$, $c = 8,5$ olarak alınır.

Hampel, Andrews ve Tukey fonksiyonlarını kullanarak parametre tahminleri yapılırken bir önceki bölümde Huber fonksiyonu için verilen algoritmanın aynıısı kullanıldığından bu bölümde ve diğer bölümlerde bu algoritmaya yer verilmemiştir.

Andrews' in fonksiyonu

Andrews (1974) tarafından öne sürülen fonksiyon,

$$\rho(r) = \begin{cases} k^2 \left[1 - \cos \frac{|r|}{k} \right] & , |r| \leq k\pi \\ 2k^2 & , |r| > k\pi \end{cases} \quad (2.7.10)$$

biçimindedir.

(2.7.10) eşitliğinin türevi alınır, bulunan fonksiyon r' ye bölünür ve ağırlık fonksiyonu elde edilir. Buna göre Andrews'in ağırlık fonksiyonu aşağıdaki biçimde tanımlanır,

$$w = \begin{cases} \frac{\sin\left(\frac{r}{k}\right)}{r} & , |r| \leq k\pi \\ 0 & , |r| > k\pi \end{cases} \quad (2.7.11)$$

Burada genel olarak $k = 1,5$ veya $k = 2,1$ olarak seçilir (Şanlı, 2005).

Tukey' in fonksiyonu

Tukey (1970) tarafından öne sürülen fonksiyon,

$$\rho(r) = \begin{cases} \frac{1}{6} \left(1 - \left(1 - \left(\frac{r}{k} \right)^2 \right)^3 \right) & , |r| \leq k \\ \frac{1}{6} & , |r| > k \end{cases} \quad (2.7.12)$$

biçimindedir.

(2.7.12) eşitliğinin türevi alınır, bulunan fonksiyon r' ye bölünür ve ağırlık fonksiyonu elde edilir. Buna göre Tukey'in ağırlık fonksiyonu aşağıdaki biçimde tanımlanır;

$$w = \begin{cases} \left(1 - \left(\frac{r}{k}\right)^2\right)^2 & , \quad |r| \leq k \\ 0 & , \quad |r| > k \end{cases} \quad (2.7.13)$$

Burada genellikle $k = 5$ veya 6 olarak seçilir (Fox, 2002).

M-Regresyon tekniği, x -yönündeki aykırı gözlem değerlerine karşı duyarlıdır. x -yönündeki gözlem değerlerinin etkisi altında kalmaktadır.

M-Regresyon tekniği, hata terimlerinin normal dağıldığı bir modelde LAD tekniğinden daha iyidir (Öztürk, 2003).

2.8. En Küçük Medyan Kare (Least Median Square-LMS) Tekniği

LMS, EKK tekniğinde olduğu gibi hata terimlerinin kareler toplamını minimize etmek yerine, kareleri alınmış hata terimlerinin medyan değerini minimize eder (Rousseeuw, 1984).

Başka bir ifadeyle LMS tekniğinde minimize edilecek fonksiyon,

$$\min \text{med}(\varepsilon_i^2) \quad (2.8.1)$$

şeklinde tanımlanır.

LMS tekniği için geliştirilen algoritma, k adet farklı gözlem değerinin alt kümelerini elde etmekle başlar. k adet gözlem değeri için regresyon parametreleri hesaplanır. Daha sonra tüm gözlemler kullanılarak hata terimlerinin karelerinin medyan değeri hesaplanır. Sonuçta minimum medyan değeri veren parametre değeri, en iyi parametre değeri olarak kabul edilir.

Genel olarak bu işlem, tüm k sayıdaki alt küme için tekrarlanır. C_n^k tane alt küme için hesaplama gerekmektedir. Çoğu uygulamada bütün alt kümeleri ele alarak parametre değerlerini hesaplamak olanaksız hale gelmektedir. Bu durumda, ele alınacak m adet alt örneklemden en az bir tanesinin en iyi olması olasılığının bire yakın olacağı biçimde çeşitli sayıda rassal seçim yapılır.

LMS tekniđi, en iyi alt küme regresyon doğrusunun geçtiđi iki noktadan oluşmaktadır. Tüm bu noktalar ikişer ikişer alınarak alt kümeler oluşturulmakta ve tüm nokta kümelerinden regresyon doğrusu geçirilerek hata terimleri hesaplanmaktadır. Hata terimi değerlerini minimum yapan regresyon doğrusunun geçtiđi iki nokta ve bu noktaya ait parametre değeri en iyi regresyon doğrusu olarak kabul edilir.

Bu algoritmayı daha detaylı biçimde inceleyecek olursak:

1. Mümkün olan tüm nokta çiftleri için parametre değerleri hesaplanır.
2. Bulunan her bir $\hat{\beta}_0, \hat{\beta}_1$ değeri için n tane gözlem çiftine ait hata terimleri bulunur.
3. Her bir $\hat{\beta}_0, \hat{\beta}_1$ değeri için hesaplanan hata terimlerinin kareleri alınarak medyanı bulunur.
4. Bulunan medyan kare değerlerinin içerisindeki en küçük medyan kare değerine karşılık gelen değerler $\hat{\beta}_0, \hat{\beta}_1$ olarak alınır.
5. Aşağıdaki ağırlık değerleri kullanılarak ağırlıklandırılmış EKK tekniđi uygulanır. Bu aşama yeniden ağırlıklandırılmış EKK aşaması olarak da adlandırılır.

LMS tekniđinde ağırlıklar,

$$w_i = \begin{cases} 1 & , \quad |\varepsilon_i/s_0| \leq 2,5 \\ 0 & , \quad |\varepsilon_i/s_0| > 2,5 \end{cases}$$

$$\text{ve } s_0 = 1,4826 \left[1 + \frac{5}{n-p} \right] \left[\sqrt{\text{med}(\varepsilon_i^2)} \right]$$

olarak hesaplanır (Rousseeuw ve Leroy, 1987).

Çalışmada yukarıda belirtilen algoritma kullanılarak parametre tahmin değerleri elde edilmiştir.

LMS tekniđinin en büyük dezavantajı, hata terimlerinin dağılımı normal ise parametre tahmincileri etkin tahminciler değildir. Ancak LMS tahmincileri asimptotik olarak normal dağılır (Öztürk, 2003).

2.9. Winsorized En Küçük Kareler (Winsorised Least Square-WLS) Tekniđi

EKK tekniđine başka bir alternatif yöntem de WLS tekniđidir. Bu teknik hata terimlerinin büyüklüklerinde önemli deđişiklikler olduđu zaman kullanılır. Böylece parametre tahmin deđerlerini etkileyen aykırı deđerlerin etkileri bu tekniđin kullanılmasıyla azalır.

WLS tekniđinde parametre tahmini yapmak için iterasyon tekniđi, seviyelendirme tekniđi, artan seviyelerde iterasyon tekniđi olarak adlandırılan üç farklı algoritma geliştirilmiştir.

İterasyon tekniđinde hata terimi deđerleri EKK tekniđinden hesaplanır ve büyüklüklerine göre sıralanır. Daha sonra her aykırı y deđerleri için hesaplanan hata terimi deđerleri kendisine en yakın hata terimi ile deđiştirilir ve yeni y deđerleri $y_i' = \hat{y}_i + \varepsilon_i'$ formülü ile hesaplanır. Winsorise etme sürecinde hata terimlerinin küçükten büyüğe sıralanmasının ardından $g = \alpha n$ ve $1 \leq g \leq n/2$ olmak üzere ilk g adet hata terimi ve son g adet hata teriminin yerleri kendilerine en yakın hata terimi ile deđiştirilerek elde edilen yeni hata terimlerinden bulunan yeni y deđerlerine EKK tekniđinin uygulandıđı bir tekniktir. Yeni y deđerlerine, uyarlanmış y deđerleri denilmektedir (Nevitt ve Tam, 1998). Bir başka deyişle, sıralanmış hata terimlerine ait baştan ve sondan g adet hata terimi atılarak, baştaki g adet hata teriminden sonraki hata terimi ve sondaki g adet hata teriminden önceki hata terimi atılan hata terimlerinin yerine alınmaktadır. Uygulamalarda α deđerinin sıklıkla 0,05 veya 0,1 olarak seçildiđi görülmektedir. Daha sonra hesaplanan bu yeni y deđerlerini kullanarak yeni parametre deđerleri bulunur. İkinci iterasyon için yeniden hata terimi deđerleri hesaplanır. Ardından y'' deđerleri üçüncü iterasyon için hesaplanır. Bu iterasyonlar böyle devam eder. Nevitt ve Tam (1998)' in çalışmaları 5 iterasyondan sonra Winsorized regresyon parametre tahminlerinin birbirine çok benzediđini göstermiştir. Bu nedenle 5 iterasyon yapmak Winsorization parametreleri hakkında karar vermek için yeterlidir (Yale ve Forsythe, 1976).

Seviyelendirme tekniğinde iterasyonların belirli sayıları için Winsorization prosedürü tekrarlanır. Her aykırı değerde winsorize noktalarının seviyeleri belirlenir. Her tekrarlama seviyeler yükseltilir (Yale ve Forsythe, 1976).

Artan seviyelerde iterasyon tekniği, iterasyon tekniğiyle hemen hemen aynıdır. Bu teknikte winsorize noktaları sabit kalır (Yale ve Forsythe, 1976).

Bununla beraber, bu çalışmada yukarıda belirtilen ilk iterasyon algoritması kullanılarak parametre tahmin değerleri elde edilmiştir.

2.10. Budanmış En Küçük Kareler (Least Trimmed Square-LTS) Tekniği

LTS tekniği Rousseeuw (1984) tarafından tanımlanmıştır.

LTS tekniğinde amaç aşağıdaki fonksiyonu minimize etmektir:

$$\min \sum_{i=1}^h (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 \quad (2.10.1)$$

Bu teknikte, ilk olarak kareleri alınan hata terimleri, küçükten büyüğe doğru sıralanır. Daha sonra bu hata terimlerinin ilk h tanesinin toplamı alınır. (2.10.1) no' lu eşitlik minimize edilir. Burada n gözlem sayısını göstermek üzere, $h = \left\lfloor \frac{n}{2} \right\rfloor + 1$ olarak alınır. LTS tekniğinde parametre tahmini yapılırken ilk h adet hata terimi kullanılırken LMS tekniğinde tüm hata terimleri kullanılır (Li, 2005).

LTS tekniği için bir diğer parametre tahmin tekniği aşağıdaki gibi verilmiştir.

Bu teknikte hata terimlerinin kareleri ε_i^2 değerleri hesaplanır ve küçükten büyüğe sıralanır. Hata terimlerinin küçükten büyüğe sıralanmasından sonra $0 < \alpha < 1/2$ olmak koşulu ile hata terimlerinin değerlerine ait en büyük αn değeri silinerek geriye kalan gözlem değerleri için EKK hesaplanması mantığına dayanmaktadır. Uygulamalarda α değerinin sıklıkla 0,05 veya 0,1 olarak seçildiği görülmektedir.

Yani ilk aşamada EKK tekniği kullanılarak hesaplanan büyük değerli hata terimlerine sahip gözlemler veri setinden silinir. Daha sonra kalan verilerle $\hat{\beta}_0$ ve $\hat{\beta}_1$ tahmin değerleri EKK tekniği ile hesaplanır (Nevitt ve Tam, 1998).

Bu çalışmada, yukarıda belirtilen gözlem değerlerinin veri setinden silindiği algoritma kullanılarak parametre tahmin değerleri elde edilmiştir.

2.11. Theil Tekniği

Theil'in regresyonu parametrik olmayan bir tekniktir ve aykırı değerli veri seti olduğunda sıklıkla kullanılır. Hata terimleri bir dağılıma uymadığında parametrik olmayan yöntemler bizi robust regresyon doğrusuna götürür ki bunun için birçok hesaplama yapmak gerekir. Bu teknik Theil (1950) tarafından önerilmiştir. Theil aykırı değerlerden şüphelenildiğinde bu tekniğin çok kullanışlı olduğunu kanıtlamıştır.

1950 yılında Theil tarafından ileri sürülen teknik, araştırmacıların en çok başvurdukları eğim bulma tekniklerinden birisidir. Bir doğrunun eğimi tahmininde kullanılan Theil metodu, (x_i, y_i) , (x_j, y_j) gözlem çiftlerinden hesaplanan eğim değerlerinin medyanı hesabına dayandırılmaktadır.

Sahip olunan $(x_1, y_1), \dots, (x_n, y_n)$ gözlem çiftlerinin, (2.1.1) eşitliğindeki basit doğrusal regresyon modeline uyduğu varsayalım. Bu denklemde x_i ve y_i ' ler gözlemlenen değerleri, β_0 ve β_1 bilinmeyen parametreleri göstermektedir. Ayrıca x_i ' ler birbirinden bağımsızdır ve $(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i), \dots, (x_n, y_n)$ şeklinde sıralandırılmışlardır.

Basit doğrusal regresyon modelinin eğim katsayısını elde etmek için veri setindeki tüm $(x_i, y_i), (x_j, y_j), \dots$ $1 \leq i < j \leq n$ ve $x_j \neq x_i$ değerleri için mümkün örneklem eğimleri,

$$b_{ij} = \frac{y_j - y_i}{x_j - x_i} \quad (2.11.1)$$

şeklinde hesaplanmalıdır (Tasker, 2000). Bu şekilde b_{ij} 'lerin $N = \frac{n!}{(n-2)!2!}$ tane

değeri elde edilecektir. Burada $b_{ij} = b_{ji}$ olur. $\hat{\beta}_1$ 'nin parametre tahmin değerini ise, b_{ij} değerlerinin medyanı verecektir. Bu belirtilenlere göre Theil'in eğim katsayısının tahmin edicisi $x_1 \leq x_2 \leq \dots \leq x_n$ olmak üzere,

$$\hat{\beta}_1 = med\{b_{ij} | x_j \neq x_i\} \text{ dır.} \quad (2.11.2)$$

$\hat{\beta}_1$ eğim tahmin edicisi için her gözlem değerine ilişkin $y_i - \hat{\beta}_1 x_i$ değerleri hesaplanmalı ve bunların medyanı alınarak regresyon doğrusunun $\hat{\beta}_0$ değeri bulunmalıdır.

Bu çalışmada (2.11.2) eşitliğindeki formül kullanılarak eğim tahmin değeri elde edilmiştir. Aynı zamanda eğim parametresi elde edildikten sonra bu eğim için sabit terim yukarıdaki gibi hesaplanmıştır (Nevitt ve Tam, 1998).

2.12. R-Regresyon Tekniği

R-kestiricileri, Hodges ve Lehman (1963)' in çalışmaları sayesinde ortaya çıkmıştır.

R-Regresyon tekniğinde, parametre tahmini yapılırken hata terimlerinin sıralandırılması gerekmektedir.

Bu teknikte amaç,

$$\min_{\hat{\beta}} \sum_{i=1}^n a_n(R_i) \varepsilon_i \quad (2.12.1)$$

fonksiyonunu minimize etmektir.

Burada R_i , ε_i hata teriminin sırasını, $a_n(i)$ ise $\sum_{i=1}^n a_n(i) = 0$ eşitliğini sağlayan monoton bir skor fonksiyonunu göstermektedir.

Skor fonksiyonları çeşitli biçimlerde tanımlanırlar.

Tablo 3. Skor Fonksiyonları

Skor Fonksiyonu	$a_n(i)$
Wilcoxon	$i - \frac{n+1}{2}$
Van Der Waerden	$\Phi^{-1}\left(\frac{i}{n+1}\right)$ (Φ^{-1} birikimli normal dağılım fonksiyonunun tersidir.)
Median	$\text{sgn}\left(i - \frac{n+1}{2}\right)$

R-Regresyon tekniğinde parametre tahminlerinin yapılabilmesi için gereken adımlar aşağıda gösterilmiştir.

1. Bağımsız değişken değerleri sıralanır. Bu değerlere ait bağımlı değişken değerleri de hesaba katılarak hata terimi değerleri bulunur. Daha sonra bulunan bu ε_i değerleri sıralanır ve R_i değerleri elde edilir. Tablo 3 de verilen skor fonksiyonu yardımıyla $a_n(R_i)$ değerleri hesaplanır. $\hat{\beta}_1$ parametresinin çeşitli değerlerine ait $a_n(R_i)\varepsilon_i$ ($i = 1, 2, \dots, n$) terimleri toplamı bulunur. Bu toplam değerinin minimum olduğu yerdeki parametre değeri grafiğe bakılarak tahmin edilir. Bu değer $\hat{\beta}_1$ parametresinin başlangıç değeri olarak kabul edilir.

2. 1. adımda elde edilen $\hat{\beta}_1$ değerini kapsayan daha küçük bir aralıktan seçilen $\hat{\beta}_1$ değerleri için 1. aşama tekrar edilir ve bu işlem $\hat{\beta}_1$ değerleri birbirine yaklaşıncaya kadar sürdürülür.

3. Daha sonraki aşamada ise $\hat{\beta}_0 = \text{med}\{y_i - \hat{\beta}_1 x_i\}$ değeri bulunur.

Bu konuda yapılan çalışmalar, burada bulunan parametre değerlerinin LAD regresyon tekniği ile bulunan parametre değerlerine benzediğini göstermiştir (Heikkilä, 2006).

Uygulama bölümünde Wilcoxon skorları kullanılarak analiz yapılacaktır.

3. UYGULAMA

Bu bölümde daha önce teorik açıklamaları yapılan robust regresyon teknikleriyle ilgili uygulama çalışmalarına yer verilmiştir.

3.1. Göl Dibi Sıcaklığı ile Göl Geometri Oranı Arasındaki İlişkinin Regresyon Denklemi Yardımıyla Belirlenmesi

Bu kısımda Gao ve Stefan (1999)' in göller üzerine yapmış olduğu çalışmadan alınan verilere robust regresyon teknikleri uygulanmış ve analiz sonuçları bir tablo halinde verilerek sonuçlar yorumlanmıştır.

Bu makalede, göllerdeki buz ve sıcaklığın göldeki ortama ve gölün biçimine bağlı olduğundan söz edilmiş ve Minnesota'daki 13 göle ait buz ve sıcaklık karakteristiklerinin gölün konumu, ortam parametresi, alanı ve derinliğine bağlı olarak değiştiği belirtilmiştir.

Göllerin buz veya sıcaklık karakteristiği, maksimum ve minimum göl dibi sıcaklığına ve maksimum ve minimum göl yüzeyi sıcaklığına bağlıdır. Maksimum göl dibi sıcaklığı (Maximum Bottom Temperatures-MBT) olarak tanımlanmıştır. Başka bir ifadeyle, maksimum göl dibi sıcaklığı bir değişken olarak çalışmada kullanılmıştır. Ayrıca Gao ve Stefan (1999) göldeki buz veya sıcaklığın gölün fiziksel karakteristikleri (gölün yüzey alanı ve maksimum derinliği) yardımıyla da açıklanabileceğini söylemiş ve LGR (Lake Geometry Ratio) değişkeni olarak adlandırılan yeni bir değişken tanımlamışlardır.

Göl yüzey alanı A_s ve gölün maksimum derinliği H_{\max} olarak tanımlandığında;

$$LGR = \frac{A_s^{0,25}}{H_{\max}}$$

olarak hesaplanır.

Makalede bahsi geçen MBT değişkeni bağımlı değişken olarak ve LGR değişkeni ise bağımsız değişken olarak alınmıştır. Analizde kullanılan veriler EK-1 'de verildiği gibidir.

EKK Tekniđi

Bu kısımda Gao ve Stefan (1999)' in göllerle ilgili yapmış olduđu çalışmasından alınan veriler kullanılarak EKK tekniđiyle belirlenen regresyon denklemi,

$$\hat{y}_i = 18,018 + 0,395x_i, \quad i = 1,2,\dots,13$$

olarak bulunur.

$\hat{\beta}_0 = 18,018$; doğrunun y eksenini kestiđi nokta olarak adlandırılır.Yani regresyon denkleminin sabit terimidir.

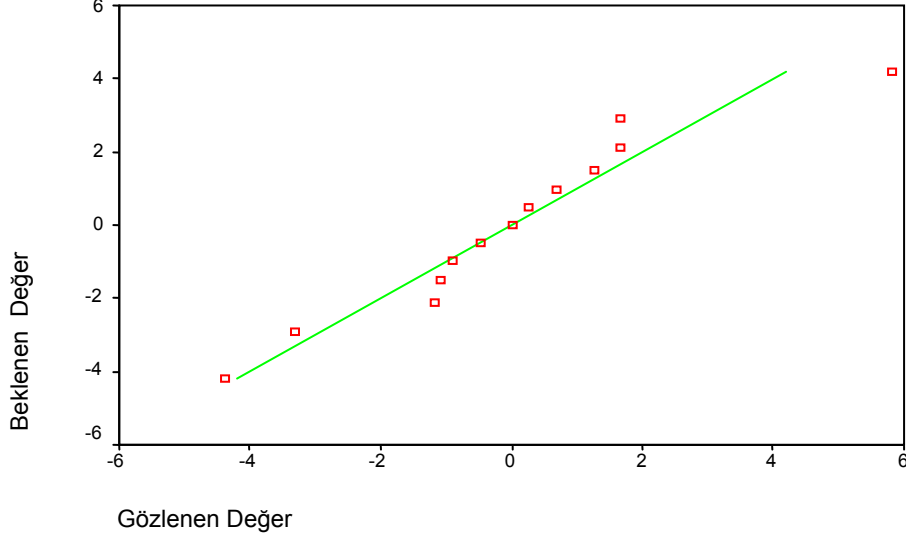
$\hat{\beta}_1 = 0,395$; doğrunun eğimi olarak adlandırılır. Yani x_i deđerindeki 1 br.lik artışın y_i deđerindeki göstereceđi artış miktarıdır.

$R^2 = 0,541$ olarak hesaplanmıştır. Bađımlı deđişkendeki toplam deđişimin %54 'ünün bađımsız deđişken tarafından açıklandığını gösterir

EKK tekniđinin uygulanabilmesi için Bölüm 2.1.'de bahsedilen hata terimleri ε_i ' lerin bađımsız ve özdeş dağılıma sahip ortalaması sıfır, varyansı σ^2 olan normal dağılım şartını sağlaması gerekmektedir. Normal dağılım varsayımının sağlanıp sağlanmadığını görsel olarak kontrol edebilmek için hata terimlerinin Q-Q grafiđine bakılır.

EK-1' deki veriler ve EKK tekniđi ile elde edilen parametre tahminleri yardımıyla hesaplanan hata terimlerinin Q-Q grafiđi ařađıdaki gibidir.

Şekil 1. Uygulama 1' deki veriler için elde edilen hata terimlerinin Q-Q grafiği



Şekil 1.' den görüldüğü üzere, hata terimleri normal dağılıma sahip değildir. Çünkü kutucuklar doğru üzerinde değildir. Dolayısıyla robust regresyon tekniklerinin kullanılması daha güvenilir sonuçlar elde edilmesini sağlar.

Üç Grup Dayanıklı Doğru Tekniği

EK-1' deki veriler kullanılarak Üç Grup Dayanıklı Doğru Tekniği' ne ait regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 16,985 + 0,489x_i, \quad i = 1,2,\dots,13.$$

$R^2 = 0,502$ olarak hesaplanmıştır. Bağımlı değişkendeki toplam değişimin %50 'si bağımsız değişken tarafından açıklanır.

Wald Tekniđi

EK-1' deki veriler kullanılarak Wald Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 15,473 + 0,680x_i , \quad i = 1,2,\dots,13.$$

$R^2 = 0,259$ olarak hesaplanmıřtır. Wald Tekniđi kullanıldıđında bađımlı deđiřkendeki toplam deđiřimin %26 'sının bađımsız deđiřken tarafından ađıklandığı grlmřtr.

Nair ve Shrivastava Tekniđi

EK-1' deki veriler kullanılarak Nair ve Shrivastava Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 15,399 + 0,515x_i , \quad i = 1,2,\dots,13.$$

$R^2 = 0,301$ olarak hesaplanmıřtır. Nair ve Shrivastava Tekniđi kullanıldıđında bađımlı deđiřkendeki toplam deđiřimin %30 'unun bađımsız deđiřken tarafından ađıklandığı grlmřtr.

Bartlett Tekniđi

EK-1' deki veriler kullanılarak Bartlett Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 16,948 + 0,515x_i , \quad i = 1,2,\dots,13.$$

$R^2 = 0,490$ olarak hesaplanmıřtır. Bartlett Tekniđi kullanıldıđında bađımlı deđiřkendeki toplam deđiřimin %49 'unun bađımsız deđiřken tarafından ađıklandığı grlmřtr.

LAD Tekniđi

EK-1' deki veriler kullanılarak LAD Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 20,018 + 0,265x_i, \quad i = 1,2,\dots,13.$$

$R^2 = 0,426$ olarak hesaplanmıřtır. LAD Tekniđi kullanıldıđında bađımlı deđiřkendeki toplam deđiřimin %43 'ünün bađımsız deđiřken tarafından ađıklanıđı gürülmüřtür.

M-Regresyon Teknikleri

EK-1' deki veriler kullanılarak Huber-M algoritma Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 18,070 + 0,383x_i, \quad i = 1,2,\dots,13.$$

$R^2 = 0,731$ olarak hesaplanmıřtır. Huber-M tekniđi kullanıldıđında bađımlı deđiřkendeki toplam deđiřimin %73 'ünün bađımsız deđiřken tarafından ađıklanıđı gürülmüřtür.

$k = 1,5$ için Huber'ın ařađıda verilen ađırlık fonksiyonu kullanılarak parametre tahmini yapılmıřtır,

$$w = \begin{cases} 1 & , |r| \leq 1,5 \\ \frac{1,5}{r} & , |r| > 1,5. \end{cases}$$

EK-1' deki veriler kullanılarak Huber-M Tekniği' ne ait regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 18,071 + 0,383x_i, \quad i = 1,2,\dots,13.$$

$R^2 = 0,630$ olarak hesaplanmıştır. Huber-M Tekniği kullanıldığında bağımlı değişkendeki toplam değişimin %63 'ünün bağımsız değişken tarafından açıklandığı görülmüştür.

$a = 1,7, b = 3,4, c = 8,5$ için Hampel'in aşağıda verilen ağırlık fonksiyonu kullanılarak parametre tahmini yapılmıştır,

$$w = \begin{cases} 1 & , \quad 0 < |r| \leq 1,7 \\ \frac{1,7}{r} \text{sgn}(r) & , \quad 1,7 < |r| \leq 3,4 \\ \frac{1,7}{r} \left[\frac{8,5 - |r|}{8,5 - b3,4} \right] \text{sgn}(r) & , \quad 3,4 < |r| \leq 8,5 \\ 0 & , \quad 8,5 < |r| \end{cases} .$$

EK-1' deki veriler kullanılarak Hampel Tekniği' ne ait regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 17,916 + 0,394x_i, \quad i = 1,2,\dots,13.$$

$R^2 = 0,620$ olarak hesaplanmıştır. Hampel Tekniği kullanıldığında bağımlı değişkendeki toplam değişimin %62 'sinin bağımsız değişken tarafından açıklandığı görülmüştür.

$k = 1,5$ için Andrews'ın aşağıda verilen ağırlık fonksiyonu kullanılarak parametre tahmini yapılmıştır,

$$w = \begin{cases} \frac{\sin\left(\frac{r}{1,5}\right)}{r} & , |r| \leq 1,5\pi \\ 0 & , |r| > 1,5\pi . \end{cases}$$

EK-1' deki veriler kullanılarak Andrews Tekniği' ne ait regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 17,884 + 0,394x_i , \quad i = 1,2,\dots,13.$$

$R^2 = 0,810$ olarak hesaplanmıştır. Andrews Tekniği kullanıldığında bağımlı değişkendeki toplam değişimin %81 'inin bağımsız değişken tarafından açıklandığı görülmüştür.

$k = 5$ için Tukey'in aşağıda verilen ağırlık fonksiyonu kullanılarak parametre tahmini yapılmıştır,

$$w = \begin{cases} \left(1 - \left(\frac{r}{5}\right)^2\right)^2 & , |r| \leq 5 \\ 0 & , |r| > 5 . \end{cases}$$

EK-1' deki veriler kullanılarak Tukey Tekniği' ne ait regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 17,869 + 0,394x_i , \quad i = 1,2,\dots,13.$$

$R^2 = 0,817$ olarak hesaplanmıştır. Tukey Tekniği kullanıldığında bağımlı değişkendeki toplam değişimin %82 'sinin bağımsız değişken tarafından açıklandığı görülmüştür.

LMS Tekniđi

EK-1' deki veriler kullanılarak LMS Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 20,434 + 0,2457x_i , \quad i = 1,2,\dots,13.$$

$R^2 = 0,975$ olarak hesaplanmıřtır. LMS Tekniđi kullanıldıđında bađımlı deđiřkendeki toplam deđiřimin %98 'inin bađımsız deđiřken tarafından ađıklandığı grlmřtr.

Yeniden ađırlıklandırılmıř EKK Tekniđi uygulandıđında regresyon denklemi ařađıdaki gibi bulunmuřtur.

$$\hat{y}_i = 20,415 + 0,234x_i , \quad i = 1,2,\dots,13$$

$R^2 = 0,955$ biiminde hesaplanmıřtır. Yeniden ađırlıklandırılmıř EKK Tekniđi kullanıldıđında bađımlı deđiřkendeki toplam deđiřimin %96' sının bađımsız deđiřken tarafından ađıklandığı grlmřtr.

Winsorized En Kk Kareler Tekniđi

EK-1' deki veriler kullanılarak WLS Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 17,672 + 0,407x_i , \quad i = 1,2,\dots,13.$$

$R^2 = 0,728$ olarak hesaplanmıřtır. WLS Tekniđi kullanıldıđında bađımlı deđiřkendeki toplam deđiřimin %73 'nn bađımsız deđiřken tarafından ađıklandığı grlmřtr.

LTS Tekniđi

EK-1' deki veriler kullanılarak LTS Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 17,176 + 0,434x_i , \quad i = 1,2,\dots,13.$$

$R^2 = 0,734$ olarak hesaplanmıřtır. LTS Tekniđi kullanıldıđında bađımlı deđiřkendeki toplam deđiřimin %73 'ünün bađımsız deđiřken tarafından ađıklandığı gürülmüřtür.

Theil Tekniđi

EK-1' deki veriler kullanılarak Theil Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 17,630 + 0,420x_i , \quad i = 1,2,\dots,13.$$

$R^2 = 0,536$ olarak hesaplanmıřtır. Theil Tekniđi kullanıldıđında bađımlı deđiřkendeki toplam deđiřimin %54 'ünün bađımsız deđiřken tarafından ađıklandığı gürülmüřtür.

R-Regresyon Tekniđi

EK-1' deki veriler ve Wilcoxon skor fonksiyonu kullanılarak elde edilen regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 19,172 + 0,350x_i , \quad i = 1,2,\dots,13.$$

$R^2 = 0,490$ olarak hesaplanmıřtır. Bu deđer de bađımlı deđiřkendeki toplam deđiřimin %49 'unun bađımsız deđiřken tarafından ađıklandığını gsterir.

Sonuç Tablosu

Tüm algoritmalar uygulandıktan sonra sonuçları aşağıdaki tabloda özetleyecek olursak hangi regresyon denkleminin bu uygulama için en iyi sonucu verdiğini daha iyi görebilme şansını yakalamış oluruz.

Tablo 4. Uygulama 1 için Sonuç Tablosu

Teknik	Denklem	R^2
EKK	$\hat{y}_i = 18,018 + 0,395x_i$, $i = 1,2,\dots,13$	0,541
Üç Grup D.D. Tekniği	$\hat{y}_i = 16,985 + 0,489x_i$, $i = 1,2,\dots,13$	0,502
Wald	$\hat{y}_i = 15,473 + 0,680x_i$, $i = 1,2,\dots,13$	0,259
Nair ve Shrivastava	$\hat{y}_i = 15,399 + 0,515x_i$, $i = 1,2,\dots,13$	0,301
Bartlett	$\hat{y}_i = 16,948 + 0,515x_i$, $i = 1,2,\dots,13$	0,490
LAD	$\hat{y}_i = 20,018 + 0,265x_i$, $i = 1,2,\dots,13$	0,426
Huber-M Algoritma	$\hat{y}_i = 18,070 + 0,383x_i$, $i = 1,2,\dots,13$	0,731
Huber-M	$\hat{y}_i = 18,071 + 0,383x_i$, $i = 1,2,\dots,13$	0,630
Hampel	$\hat{y}_i = 17,916 + 0,394x_i$, $i = 1,2,\dots,13$	0,620
Andrews	$\hat{y}_i = 17,884 + 0,394x_i$, $i = 1,2,\dots,13$	0,810
Tukey	$\hat{y}_i = 17,869 + 0,394x_i$, $i = 1,2,\dots,13$	0,817
LMS	$\hat{y}_i = 20,415 + 0,234x_i$, $i = 1,2,\dots,13$	0,955
WLS	$\hat{y}_i = 17,672 + 0,407x_i$, $i = 1,2,\dots,13$	0,728
LTS	$\hat{y}_i = 17,176 + 0,434x_i$, $i = 1,2,\dots,13$	0,734
Theil	$\hat{y}_i = 17,630 + 0,420x_i$, $i = 1,2,\dots,13$	0,536
R-Regresyonu	$\hat{y}_i = 19,172 + 0,350x_i$, $i = 1,2,\dots,13$	0,490

Tablo 4' ten de görüleceği gibi bağımsız değişkenin bağımlı değişkeni açıklama yüzdesi bakımından en yüksek değeri LMS tekniği vermiştir. Daha sonra sırasıyla Tukey, Andrews, LTS, WLS, Huber-M Algoritma, Huber-M, Hampel, EKK, Theil, Üç Grup D. D. Tekniği, R-Regresyon, Bartlett, LAD, Nair ve Shrivastava, Wald teknikleri gelmektedir.

3.2. Balık Yakalamak için Sarfedilen Çaba ile Klorofil α Arasındaki İlişkinin Regresyon Denklemi Yardımıyla Belirlenmesi

Yapılan ikinci uygulamada Egertson ve Downing (2004)' un makalesindeki veriler alınarak robust regresyon teknikleri uygulanmış ve aşağıdaki sonuçlara ulaşılmıştır.

Bu makalede, çeşitli balık türlerini birim zamanda yakalamak için sarf edilen çaba ile göl ekosistem karakteristikleri arasında bir ilişkiden söz edilmiştir. Iowa'daki 32 gölden bu amaç için çeşitli veriler toplanmıştır. Her gece her balık ağından gram bazında yeni veriler alınmış ve balık yakalamak için sarfedilen çaba CPUE_w (Catch per unit effort by weight) değişkeni olarak tanımlanmıştır. Bu değişken çalışmada bağımlı değişken olarak ele alınmıştır. CPUE_w değişkeni, balık türlerini yakalamak için sarf edilen çaba değerlerinin toplamıdır. Analizde ele alınan balık türleri; Black Bullhead, Common Carp, Black Crappie, White Crappie, Bluegill ve Channel Catfish' dir.

Göldeki balık yakalamak için sarf edilen çabanın göl ekosistem karakteristiklerine bağlı olduğu söylenerek chlorophyll α (klorofil α), total phosphorus (toplam fosfor) ve total nitrogen (toplam nitrojen) değişkenleri tanımlanmıştır. Ancak çalışmada sadece chlorophyll α değişkeni ile analiz yapılmıştır. Çünkü balık yakalamak için sarf edilen çaba ile chlorophyll α arasında pozitif ve anlamlı bir ilişki bulunmuştur. Bu değişken ise bağımsız değişken olarak çalışmada kullanılmıştır.

Ancak sabit varyanslılığı sağlamak için her iki değişkenin de logaritma değerleri alınarak analiz tamamlanmıştır. Analizde kullanılan veriler EK-2 'de verilmiştir.

EKK Tekniği

Bu kısımda Egertson ve Downing (2004)' in balık yakalamakla ilgili yapmış olduğu çalışmasından alınan veriler kullanılarak EKK tekniğiyle belirlenen regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 2,92 + 0,463x_i, \quad i = 1,2,\dots,32.$$

$\hat{\beta}_0 = 2,92$; regresyon denkleminin sabit terimidir.

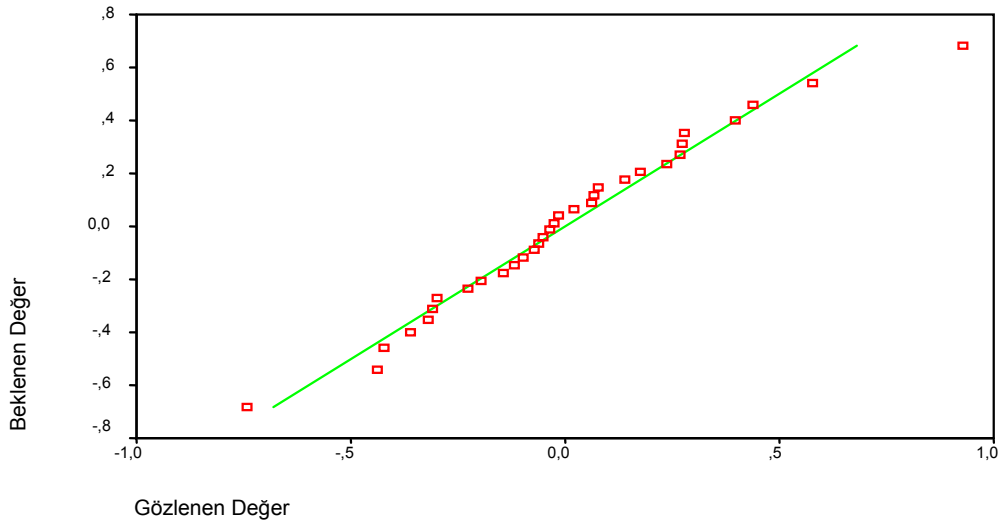
$\hat{\beta}_1 = 0,463$; regresyon denkleminin eğimi olarak adlandırılır.

$R^2 = 0,21$ olarak hesaplanmıştır. Bağımlı değişkendeki toplam değişimin %21 'inin bağımsız değişken tarafından açıklandığını gösterir.

EKK tekniğinin uygulanabilmesi için Bölüm 2.1.'de bahsedilen hata terimleri ε_i ' lerin bağımsız ve özdeş dağılıma sahip ortalaması sıfır, varyansı σ^2 olan normal dağılım şartını sağlaması gerekmektedir. Normal dağılım varsayımının sağlanıp sağlanmadığını görsel olarak kontrol edebilmek için hata terimlerinin Q-Q grafiğine bakılır.

EK-2' deki veriler ve EKK tekniği ile elde edilen parametre tahminleri yardımıyla hesaplanan hata terimlerinin Q-Q grafiği aşağıdaki gibidir.

Şekil 2. Uygulama 2' deki veriler için elde edilen hata terimlerinin Q-Q grafiği



Şekil 2.' den görüldüğü üzere, hata terimleri normal dağılıma sahip değildir. Çünkü kutucuklar doğru üzerinde değildir. Dolayısıyla robust regresyon tekniklerinin kullanılması daha güvenilir sonuçlar elde edilmesini sağlar.

Üç Grup Dayanıklı Doğrular Tekniği

EK-2' deki veriler kullanılarak Üç Grup Dayanıklı Doğru Tekniği' ne ait regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 2,775 + 0,556x_i , \quad i = 1,2,\dots,32.$$

$R^2 = 0,196$ olarak hesaplanmıştır. Bağımlı değişkendeki toplam değişimin %20 'sinin bağımsız değişken tarafından açıklandığını gösterir.

Wald Tekniği

EK-2' deki veriler kullanılarak Wald Tekniği' ne ait regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 3,104 + 0,325x_i , \quad i = 1,2,\dots,32.$$

$R^2 = 0,191$ olarak hesaplanmıştır. Bağımlı değişkendeki toplam değişimin %19 'unun bağımsız değişken tarafından açıklandığını gösterir.

Nair ve Shrivastava Tekniği

EK-2' deki veriler kullanılarak Nair ve Shrivastava Tekniği' ne ait regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 3,088 + 0,338x_i , \quad i = 1,2,\dots,32.$$

$R^2 = 0,195$ olarak hesaplanmıştır. Bağımlı değişkendeki toplam değişimin %20 'sinin bağımsız değişken tarafından açıklandığını gösterir.

Bartlett Tekniđi

EK-2' deki veriler kullanılarak Bartlett Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 3,086 + 0,338x_i , \quad i = 1,2,\dots,32.$$

$R^2 = 0,195$ olarak hesaplanmıřtır. Bađımlı deđiřkendeki toplam deđiřimin %20 'sinin bađımsız deđiřken tarafından ađıklanđını gsterir.

LAD Tekniđi

EK-2' deki veriler kullanılarak LAD Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 2,73 + 0,588x_i , \quad i = 1,2,\dots,32.$$

$R^2 = 0,193$ olarak hesaplanmıřtır. Bađımlı deđiřkendeki toplam deđiřimin %19 'unun bađımsız deđiřken tarafından ađıklanđını gsterir.

M –Regresyon Teknikleri

EK-2' deki veriler kullanılarak Huber-M algoritma Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 2,810 + 0,535x_i , \quad i = 1,2,\dots,32.$$

$R^2 = 0,356$ olarak hesaplanmıřtır. Bađımlı deđiřkendeki toplam deđiřimin %36 'sının bađımsız deđiřken tarafından ađıklanđını gsterir.

$k = 1,5$ için Huber'ın aşağıda verilen ağırlık fonksiyonu kullanılarak parametre tahmini yapılmıştır,

$$w = \begin{cases} 1 & , |r| \leq 1,5 \\ \frac{1,5}{r} & , |r| > 1,5 . \end{cases}$$

EK-2' deki veriler kullanılarak Huber-M Tekniği' ne ait regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 2,841 + 0,508x_i , \quad i = 1,2,\dots,32.$$

$R^2 = 0,268$ olarak hesaplanmıştır. Bağımlı değişkendeki toplam değişimin %27 'sinin bağımsız değişken tarafından açıklandığını gösterir.

$a = 1,7$, $b = 3,4$, $c = 8,5$ için Hampel'in aşağıda verilen ağırlık fonksiyonu kullanılarak parametre tahmini aşağıdaki gibi yapılır.

$$w = \begin{cases} 1 & , 0 < |r| \leq 1,7 \\ \frac{1,7}{r} \text{sgn}(r) & , 1,7 < |r| \leq 3,4 \\ \frac{1,7}{r} \left[\frac{8,5 - |r|}{8,5 - b3,4} \right] \text{sgn}(r) & , 3,4 < |r| \leq 8,5 \\ 0 & , 8,5 < |r| . \end{cases}$$

EK-2' deki veriler kullanılarak Hampel Tekniği' ne ait regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 2,827 + 0,524x_i , \quad i = 1,2,\dots,32.$$

$R^2 = 0,290$ olarak hesaplanmıştır. Bağımlı değişkendeki toplam değişimin %29 'unun bağımsız değişken tarafından açıklandığını gösterir.

$k = 1,5$ için Andrews'in aşağıda verilen ağırlık fonksiyonu kullanılarak parametre tahmini yapılmıştır,

$$w = \begin{cases} \frac{\sin\left(\frac{r}{1,5}\right)}{r} & , |r| \leq 1,5\pi \\ 0 & , |r| > 1,5\pi \end{cases}$$

EK-2' deki veriler kullanılarak Andrews Tekniği' ne ait regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 2,806 + 0,537x_i \quad , \quad i = 1,2,\dots,32.$$

$R^2 = 0,317$ olarak hesaplanmıştır. Bağımlı değişkendeki toplam değişimin %32 'sinin bağımsız değişken tarafından açıklandığını gösterir.

$k = 5$ için Tukey'in aşağıda verilen ağırlık fonksiyonu kullanılarak parametre tahmini yapılmıştır,

$$w = \begin{cases} \left(1 - \left(\frac{r}{5}\right)^2\right)^2 & , |r| \leq 5 \\ 0 & , |r| > 5 \end{cases}$$

EK-2' deki veriler kullanılarak Tukey Tekniği' ne ait regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 2,794 + 0,544x_i \quad , \quad i = 1,2,\dots,32.$$

$R^2 = 0,330$ olarak hesaplanmıştır. Bağımlı değişkendeki toplam değişimin %33 'ünün bağımsız değişken tarafından açıklandığını gösterir.

LMS Tekniđi

EK-2' deki veriler kullanılarak LMS Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 2,17 + 1,00x_i , \quad i = 1,2,\dots,32.$$

$R^2 = 0,48$ olarak hesaplanmıřtır. Bađımlı deđiřkendeki toplam deđiřimin %48 'inin bađımsız deđiřken tarafından ađıklandığını gsterir.

Yeniden ađırlıklandırılmıř EKK tekniđi uygulandıđında regresyon denklemi ařađıdaki gibi bulunur.

$$\hat{y}_i = 2,684 + 0,627x_i , \quad i = 1,2,\dots,32$$

$R^2 = 0,453$ olur. Bađımlı deđiřkendeki toplam deđiřimin %45 'inin bađımsız deđiřken tarafından ađıklandığını gsterir.

Winsorized En Kk Kareler Tekniđi

EK-2' deki veriler kullanılarak WLS Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 2,770 + 0,561x_i , \quad i = 1,2,\dots,32.$$

$R^2 = 0,445$ olarak hesaplanmıřtır. Bađımlı deđiřkendeki toplam deđiřimin %45 'inin bađımsız deđiřken tarafından ađıklandığını gsterir.

LTS Tekniđi

EK-2' deki veriler kullanılarak LTS Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 2,678 + 0,619x_i \quad , \quad i = 1,2,\dots,32.$$

$R^2 = 0,497$ olarak hesaplanmıřtır. Bađımlı deđiřkendeki toplam deđiřimin %50 'ünün bađımsız deđiřken tarafından ađıklandığını gsterir.

Theil Tekniđi

EK-2' deki veriler kullanılarak Theil Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 2,820 + 0,525x_i \quad , \quad i = 1,2,\dots,32.$$

$R^2 = 0,536$ olarak hesaplanmıřtır. Bađımlı deđiřkendeki toplam deđiřimin %54 'ünün bađımsız deđiřken tarafından ađıklandığını gsterir.

R-Regresyon Tekniđi

EK-2' deki veriler ve Wilcoxon skor fonksiyonu kullanılarak elde edilen regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 2,884 + 0,500x_i \quad , \quad i = 1,2,\dots,32.$$

$R^2 = 0,21$ olarak hesaplanmıřtır. Bađımlı deđiřkendeki toplam deđiřimin %21 'inin bađımsız deđiřken tarafından ađıklandığını gsterir.

Sonuç Tablosu

Tüm algoritmalar uygulandıktan sonra sonuçları aşağıdaki tabloda özetleyecek olursak hangi regresyon denkleminin bu uygulama için en iyi sonucu verdiğini daha iyi görebilme şansını yakalamış oluruz.

Tablo 5. Uygulama 2 için Sonuç Tablosu

Teknik	Denklem	R^2
EKK	$\hat{y}_i = 2,920 + 0,463x_i$, $i = 1,2,\dots,32$	0,210
Üç Grup D. D. Tekniği	$\hat{y}_i = 2,775 + 0,556x_i$, $i = 1,2,\dots,32$	0,196
Wald	$\hat{y}_i = 3,104 + 0,325x_i$, $i = 1,2,\dots,32$	0,191
Nair ve Shritastava	$\hat{y}_i = 3,088 + 0,338x_i$, $i = 1,2,\dots,32$	0,195
Bartlett	$\hat{y}_i = 3,086 + 0,338x_i$, $i = 1,2,\dots,32$	0,195
LAD	$\hat{y}_i = 2,730 + 0,588x_i$, $i = 1,2,\dots,32$	0,193
Huber-M Algoritma	$\hat{y}_i = 2,810 + 0,535x_i$, $i = 1,2,\dots,32$	0,356
Huber-M	$\hat{y}_i = 2,841 + 0,508x_i$, $i = 1,2,\dots,32$	0,268
Hampel	$\hat{y}_i = 2,827 + 0,524x_i$, $i = 1,2,\dots,32$	0,290
Andrews	$\hat{y}_i = 2,806 + 0,537x_i$, $i = 1,2,\dots,32$	0,317
Tukey	$\hat{y}_i = 2,794 + 0,544x_i$, $i = 1,2,\dots,32$	0,330
LMS	$\hat{y}_i = 2,684 + 0,627x_i$, $i = 1,2,\dots,32$	0,453
WLS	$\hat{y}_i = 2,770 + 0,561x_i$, $i = 1,2,\dots,32$	0,445
LTS	$\hat{y}_i = 2,678 + 0,619x_i$, $i = 1,2,\dots,32$	0,497
Theil	$\hat{y}_i = 2,820 + 0,525x_i$, $i = 1,2,\dots,32$	0,536
R-Regresyonu	$\hat{y}_i = 2,884 + 0,500x_i$, $i = 1,2,\dots,32$	0,210

Tablo 5' ten de görüleceği gibi bağımsız değişkenin bağımlı değişkeni açıklama yüzdesi bakımından en yüksek değeri Theil tekniği vermiştir. Daha sonra sırasıyla LTS, LMS, WLS, Huber-M Algoritma, Tukey, Andrews, Hampel, Huber-M, EKK, R-Regresyonu, Üç Grup D. D. Tekniği, Bartlett, Nair ve Shritastava, LAD, Wald teknikleridir.

3.3. Beton Kırma Deneyi Sırasında Açığa Çıkan Enerji Miktarı ile Darbe Değeri Arasındaki İlişkinin Regresyon Denklemi Yardımıyla Belirlenmesi

Altındağ (2003)' in yaptığı beton kırma denemesinde, açığa çıkan enerji miktarı ile darbe değeri arasında ilişki kurulmaya çalışılmıştır. Darbe değeri, gevreklik olarak adlandırılan be3 değişkeni değerleri ile ölçülür. Açığa çıkan enerji ise spesifik enerji olarak adlandırılır.

Çalışmada beton kırılırken açığa çıkan enerji miktarı (Spesifik Enerji) bağımlı değişken ve be3 kırılma ölçüsü (Gevreklik) değişkeni de bağımsız değişken olarak alınarak analiz yapılmıştır.

Altındağ (2003)' in verileri EK-3 de gösterildiği gibidir.

EKK Tekniği

Veriler kullanılarak hesaplanan regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 14,872 + 0,037x_i, \quad i = 1,2,\dots,25.$$

$$\hat{\beta}_0 = 14,872; \text{ regresyon denkleminin sabit terimidir.}$$

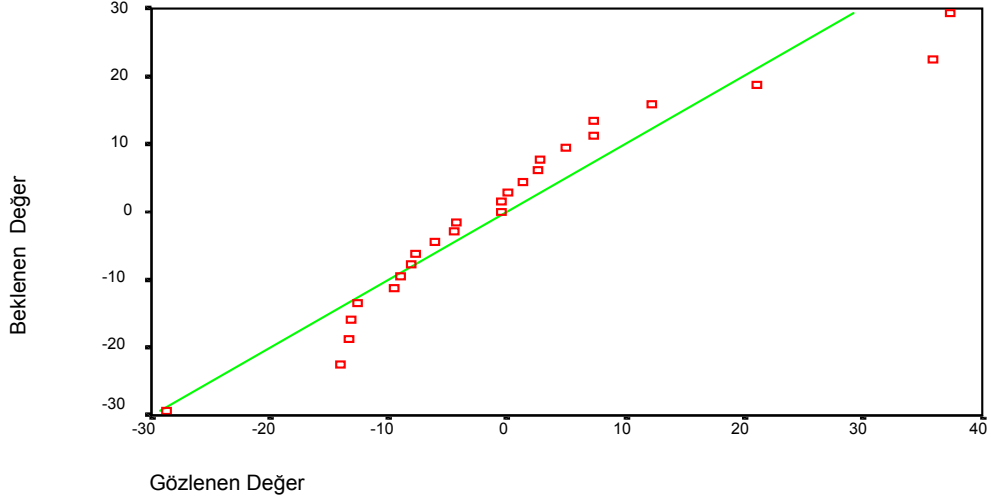
$$\hat{\beta}_1 = 0,0372; \text{ regresyon denkleminin eğimi olarak adlandırılır.}$$

$R^2 = 0,680$ olarak hesaplanmıştır. Bağımlı değişkendeki toplam değişimin %68 'inin bağımsız değişken tarafından açıklandığını gösterir.

EKK tekniğinin uygulanabilmesi için Bölüm 2.1.'de bahsedilen hata terimleri ε_i ' lerin bağımsız ve özdeş dağılıma sahip ortalaması sıfır, varyansı σ^2 olan normal dağılım şartını sağlaması gerekmektedir. Normal dağılım varsayımının sağlanıp sağlanmadığını görsel olarak kontrol edebilmek için hata terimlerinin Q-Q grafiğine bakılır.

EK-3' deki veriler ve EKK tekniği ile elde edilen parametre tahminleri yardımıyla hesaplanan hata terimlerinin Q-Q grafiği aşağıdaki gibidir.

Şekil 3. Uygulama 3’ deki veriler için elde edilen hata terimlerinin Q-Q grafiği



Şekil 3.’ den görüldüğü üzere, hata terimleri normal dağılıma sahip değildir. Çünkü kutucuklar doğru üzerinde değildir. Dolayısıyla robust regresyon tekniklerinin kullanılması daha güvenilir sonuçlar elde edilmesini sağlar.

Üç Grup Dayanıklı Doğrular Tekniği

EK-3’ deki veriler kullanılarak Üç Grup Dayanıklı Doğru Tekniği’ ne ait regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 3,775 + 0,054x_i , \quad i = 1,2,\dots,25.$$

$R^2 = 0,500$ olarak hesaplanmıştır. Bağımlı değişkendeki toplam değişimin %50 ’sinin bağımsız değişken tarafından açıklandığını gösterir.

Wald Tekniđi

EK-3' deki veriler kullanılarak Wald Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 8,124 + 0,058x_i \quad , \quad i = 1,2,\dots,25.$$

$R^2 = 0,471$ olarak hesaplanmıřtır. Bađımlı deđiřkendeki toplam deđiřimin %47 'sinin bađımsız deđiřken tarafından ađıklandığını gsterir.

Nair ve Shrivastava Tekniđi

EK-3' deki veriler kullanılarak Nair ve Shrivastava Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 6,189 + 0,057x_i \quad , \quad i = 1,2,\dots,25.$$

$R^2 = 0,483$ olarak hesaplanmıřtır. Bađımlı deđiřkendeki toplam deđiřimin %48 'inin bađımsız deđiřken tarafından ađıklandığını gsterir.

Bartlett Tekniđi

EK-3' deki veriler kullanılarak Bartlett Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 8,613 + 0,057x_i \quad , \quad i = 1,2,\dots,25.$$

$R^2 = 0,49$ olarak hesaplanmıřtır. Bađımlı deđiřkendeki toplam deđiřimin %49 'unun bađımsız deđiřken tarafından ađıklandığını gsterir.

LAD Tekniđi

EK-3' deki veriler kullanılarak LAD Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 14,485 + 0,036x_i , \quad i = 1,2,\dots,25.$$

$R^2 = 0,681$ olarak hesaplanmıřtır. Bađımlı deđiřkendeki toplam deđiřimin %68 'inin bađımsız deđiřken tarafından ađıklandıđını gsterir.

M-Regresyon Teknikleri

EK-3' deki veriler kullanılarak Huber-M algoritma Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 13,305 + 0,038x_i , \quad i = 1,2,\dots,25.$$

$R^2 = 0,832$ olarak hesaplanmıřtır. Bađımlı deđiřkendeki toplam deđiřimin %83 'ünün bađımsız deđiřken tarafından ađıklandıđını gsterir.

EK-3' deki veriler kullanılarak Huber-M Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 13,306 + 0,037x_i , \quad i = 1,2,\dots,25.$$

$R^2 = 0,758$ olarak hesaplanmıřtır. Bađımlı deđiřkendeki toplam deđiřimin %76 'sının bađımsız deđiřken tarafından ađıklandıđını gsterir.

EK-3' deki veriler kullanılarak Hampel Tekniği' ne ait regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 13,579 + 0,037x_i , \quad i = 1,2,\dots,25.$$

$R^2 = 0,743$ olarak hesaplanmıştır. Bağımlı değişkendeki toplam değişimin %74 'ünün bağımsız değişken tarafından açıklandığını gösterir.

EK-3' deki veriler kullanılarak Andrews Tekniği' ne ait regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 13,137 + 0,037x_i , \quad i = 1,2,\dots,25.$$

$R^2 = 0,772$ olarak hesaplanmıştır. Bağımlı değişkendeki toplam değişimin %77 'sinin bağımsız değişken tarafından açıklandığını gösterir.

EK-3' deki veriler kullanılarak Tukey Tekniği' ne ait regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 12,962 + 0,037x_i , \quad i = 1,2,\dots,25.$$

$R^2 = 0,781$ olarak hesaplanmıştır. Bağımlı değişkendeki toplam değişimin %78 'inin bağımsız değişken tarafından açıklandığını gösterir.

LMS Tekniği

EK-3' deki veriler kullanılarak LMS Tekniği' ne ait regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = -0,38081 + 0,32872x_i , \quad i = 1,2,\dots,25.$$

$R^2 = 0,638$ olarak hesaplanmıştır. Bağımlı değişkendeki toplam değişimin %64 'ünün bağımsız değişken tarafından açıklandığını gösterir.

Yeniden ağırlıklandırılmış EKK teniği uygulandığında LMS denklemi,

$$\hat{y}_i = 1,265 + 0,234x_i \quad , \quad i = 1,2,\dots,25.$$

$R^2 = 0,776$ olarak hesaplanır. Bağımlı değişkendeki toplam değişimin %78 'inin bağımsız değişken tarafından açıklandığını gösterir.

LTS Regresyonu

EK-3' deki veriler kullanılarak LTS Tekniği' ne ait regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 12,522 + 0,035x_i \quad , \quad i = 1,2,\dots,25.$$

$R^2 = 0,798$ olarak hesaplanmıştır. Bağımlı değişkendeki toplam değişimin %75 'inin bağımsız değişken tarafından açıklandığını gösterir.

Theil Tekniği

EK-3' deki veriler kullanılarak Theil Tekniği' ne ait regresyon denklemi aşağıdaki gibidir,

$$\hat{y}_i = 6,972 + 0,066x_i \quad , \quad i = 1,2,\dots,25.$$

$R^2 = 0,431$ olarak hesaplanmıştır. Bağımlı değişkendeki toplam değişimin %43 'ünün bağımsız değişken tarafından açıklandığını gösterir.

Winsorized En Küçük Kareler Tekniđi

EK-3' deki veriler kullanılarak WLS Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 12,869 + 0,038x_i \quad , \quad i = 1,2,..25.$$

$R^2 = 0,872$ olarak hesaplanmıřtır. Bađımlı deđiřkendeki toplam deđiřimin %87 'sinin bađımsız deđiřken tarafından ađıklandıđını gsterir.

R-Regresyon Tekniđi

EK-3' deki veriler kullanılarak R-Regresyon Tekniđi' ne ait regresyon denklemi ařađıdaki gibidir,

$$\hat{y}_i = 13,954 + 0,045x_i \quad , \quad i = 1,2,..25.$$

$R^2 = 0,650$ olarak hesaplanmıřtır. Bađımlı deđiřkendeki toplam deđiřimin %65 'inin bađımsız deđiřken tarafından ađıklandıđını gsterir.

Sonuç Tablosu

Tüm algoritmalar uygulandıktan sonra sonuçları aşağıdaki tabloda özetleyecek olursak hangi regresyon denkleminin bu uygulama için en iyi sonucu verdiğini daha iyi görebilme şansını yakalamış oluruz.

Tablo 6. Uygulama 3 için Sonuç Tablosu

Teknik	Denklem	R^2
EKK	$\hat{y}_i = 14,872 + 0,037x_i$, $i = 1,2,\dots,25$	0,680
Üç Grup D. D. Tekniği	$\hat{y}_i = 3,775 + 0,054x_i$, $i = 1,2,\dots,25$	0,500
Wald	$\hat{y}_i = 8,124 + 0,058x_i$, $i = 1,2,\dots,25$	0,471
Nair ve Shrivastava	$\hat{y}_i = 6,189 + 0,057x_i$, $i = 1,2,\dots,25$	0,483
Bartlett	$\hat{y}_i = 8,613 + 0,057x_i$, $i = 1,2,\dots,25$	0,490
LAD	$\hat{y}_i = 14,485 + 0,036x_i$, $i = 1,2,\dots,25$	0,681
Huber-M Algoritma	$\hat{y}_i = 13,305 + 0,038x_i$, $i = 1,2,\dots,25$	0,832
Huber-M	$\hat{y}_i = 13,306 + 0,037x_i$, $i = 1,2,\dots,25$	0,758
Hampel	$\hat{y}_i = 13,579 + 0,037x_i$, $i = 1,2,\dots,25$	0,743
Andrews	$\hat{y}_i = 13,137 + 0,037x_i$, $i = 1,2,\dots,25$	0,772
Tukey	$\hat{y}_i = 12,962 + 0,037x_i$, $i = 1,2,\dots,25$	0,781
LMS	$\hat{y}_i = 1,265 + 0,234x_i$, $i = 1,2,\dots,25$	0,776
LTS	$\hat{y}_i = 12,522 + 0,035x_i$, $i = 1,2,\dots,25$	0,798
Theil	$\hat{y}_i = 6,972 + 0,066x_i$, $i = 1,2,\dots,25$	0,431
WLS	$\hat{y}_i = 12,869 + 0,038x_i$, $i = 1,2,\dots,25$	0,872
R-Regresyonu	$\hat{y}_i = 13,954 + 0,045x_i$, $i = 1,2,\dots,25$	0,650

Tablo 6' dan da görüleceği gibi bağımsız değişkenin bağımlı değişkeni açıklama yüzdesi bakımından en yüksek değeri WLS tekniği vermiştir. Daha sonra sırasıyla Huber-M Algoritma, LTS, Tukey, LMS, Andrews, Huber-M, Hampel, LAD, EKK, R-Regresyonu, Üç Grup D. D. Tekniği, Bartlett, Nair ve Shrivastava, Wald, Theil teknikleridir.

4.SONUÇ VE ÖNERİLER

Bu çalışmada, doğrusal regresyon çözümlemesine giriş yapılmış ve aynı zamanda parametre tahminleri ve bu tahminlerin yapılabilmesi için gerekli koşulların varlığından bahsedilmiştir. EKK tekniğinin gözlem değerleri içinde bulunan aykırı değerlerden etkilendiği ve hata terimlerinin dağılımının normal olmadığı durumlarda alternatif regresyon tekniklerinin kullanılması gerekliliği üzerinde durulmuştur.

Bilindiği üzere gerçek veriler ile çalışılırken en büyük sıkıntı verilerin dağılımının normal dağılıma uymamasıdır. İşte bu yüzden ki alternatif regresyon tekniklerine ihtiyaç duyulur. Çünkü EKK tekniği varsayımlarından birisi de hata terimlerinin dağılımının normal olmasıdır. Bahsedildiği üzere gerçek gözlem değerleri ile çalışıldığında böyle bir durumun olması genel olarak söz konusu değildir.

Çalışmanın uygulama bölümünde ise basit doğrusal robust regresyon teknikleri kullanılarak gerçek yaşam verileri üzerinde üç adet uygulama yapılmıştır. Sonuçlar elde edildikten sonra belirlilik katsayıları hesaplanmış, bağımlı değişkendeki toplam değişimin ne kadarının bağımsız değişken tarafından açıklandığı gösterilmiştir. Buna göre ilk uygulamada bağımsız değişkenin bağımlı değişkeni açıklama yüzdesi bakımından en yüksek değeri alan teknik, LMS tekniğidir. Daha sonra sırasıyla Tukey, Andrews, LTS, WLS, Huber-M Algoritma, Huber-M, Hampel, EKK, Theil, Üç Grup D. D. Tekniği, R-Regresyon, Bartlett, LAD, Nair ve Shritastava, Wald teknikleri gelmektedir. İkinci uygulamada bağımsız değişkenin bağımlı değişkeni açıklama yüzdesi bakımından en yüksek değeri alan teknik, Theil tekniğidir. Daha sonra sırasıyla LTS, LMS, WLS, Huber-M Algoritma, Tukey, Andrews, Hampel, Huber-M, EKK, R-Regresyonu, Üç Grup D. D. Tekniği, Bartlett, Nair ve Shritastava, LAD, Wald teknikleridir. Son uygulamada bağımsız değişkenin bağımlı değişkeni açıklama yüzdesi bakımından en yüksek değeri alan teknik, WLS tekniğidir. Daha sonra sırasıyla Huber-M Algoritma, LTS, Tukey, LMS, Andrews, Huber-M, Hampel, LAD, EKK, R-Regresyonu, Üç Grup D. D. Tekniği, Bartlett, Nair ve Shritastava, Wald, Theil teknikleridir.

Sonuçları genel olarak inceleyecek olursak, LMS, LTS, M-Regresyon ve WLS tekniklerinin belirlilik katsayılarının EKK tekniğine göre yüksek olduğu görülmüştür. Bu bilginin ışığında, bu uygulama çalışmaları için LMS, LTS, M-Regresyon ve WLS tekniklerinin EKK tekniğinden daha güvenilir sonuçlar verdiği söylenebilir.

Bu çalışmada, basit doğrusal regresyon için robust regresyon teknikleri ele alınmıştır. Çoklu doğrusal regresyon için robust regresyon teknikleri ve çok değişkenli robust regresyon teknikleri konuları üzerinde rahatlıkla çalışılabilir.

KAYNAKLAR DİZİNİ

- Abbott, A. (1997). "Of time and space: the contemporary relevance of the Chicago school," *Social Forces*, 75, 1149-1182.
- Altındağ, R. (2003). "Correlation of specific energy with rock brittleness concepts on cutting," *The Journal of the South African Institute of Mining and Metallurgy*, 15, 163-171.
- Angrist, J., Imbens, G., Rubin, D. (1996). "Identification of causal effects using instrumental variables," *J. Amer. Statist. Assoc.*, 91, 444-472.
- Andrews, D. (1974). "A robust method for linear regression," *Technometrics*, 16, 523-531.
- Bartlett, M. (1949). "Fitting a straight line when both variables are subject to error," *Biometrics*, 5, 207-212.
- Birkes, D. and Dodge, Y. (1993). *Alternative Methods of Regression*. NY: Wiley.
- David, A. (1981). *Introduction to Statistics*. St. Paul: West Publishing Company.
- Egertson, C. and Downing, J. (2004). "Relationship of fish catch and composition to water quality in a suite of agriculturally eutrophic lakes," *Can. J. Fish. Aquat. Sci.* 61, 1784-1796.
- Ertek, T. (1983). *Ekonometriye Giriş*. İstanbul: Ekin Yayınları.
- Freedman, D. (1999). "From association to causation: Some remarks on the history of Statistics," *Statistical Science*, 14(3), 243-258.
- Fox, J. (2002). *Robust Regression*. Appendix to An R and S-PLUS Companion to Applied Regression, S-PLUS Software.

- Gao, S. and Stefan, H. (1999). "Multiple linear regression for lake ice and lake temperature, characteristics," *Journal Of Cold Regions Engineering*, 86, 121-135.
- Gauss, C. (1809). *Theoria Motus Corporum Coelestium* (Reprinted 1963 by Dover). Hamburg: Perthes et Esser.
- Goldthorpe, J. (1998). *Causation, Statistics and Sociology*, Twenty-ninth Geary Lecture, Nufeld College, Oxford. (Published by Economic and Social Research Institute, Dublin, Ireland).
- Gujarati, N. (1999). *Basic Econometrics*. Çev. Ü. Şenesen, İstanbul: Literatür Yayıncılık.
- Hampel, F. (1971). "A general qualitative definition of robustness," *Ann. Math. Stat.*, 42, 1887-1896.
- Heikkila, J. (Erişim Tarihi: 02.04.2006), *Graudate course on Advanced statistical signal processing*. Web adresi: <http://www.ee.oulu.fi/~jth/robust.pdf>.
- Hoaglin, D., Mosteller F., Tukey J. W. (1983). *Understanding Robust and Exploratory Data Analysis*. New York: John Wiley & Sons, Inc.
- Hodges, J. and Lehmann E. (1963). "Estimates of location based on rank tests," *Ann. Math. Stat.*, 34, 598-611.
- Holland, P. (1988). *Casual Inference, Path Analysis and Recursive Structural Equations Models*, Sociological Methodology. Oxford: Blackwell.
- Huber, P. (1973). "Robust regression: Asymptoties, conjections and monte carlo," *Ann. Stat.*, 1, 799-821.
- Jabr, R. (2005). "Power system Huber M-estimation with equality and inequality constraints," *Electric Power System Research*, 74, 239-246.

- Legendre, A. (1805). *Nouvelles Methodes pour la Determination des Orbites des Cometes* (Reprinted 1959 by Dover). Paris: Courcier.
- Li, L. (2005). "An Algorithm for computing exact least-trimmed squares estimate of simple linear regression with constraints," *Computational Statistics and Data Analysis*, 18, 188-203.
- Liu, T. (1960). "Under-identification, structural estimation, and forecasting," *Econometrica*, 28, 855-865.
- Lucas, R. (1976). "Econometric policy evaluation: A critique (with discussion) In The Phillips Curve and Labor Markets (K. Brunner and A. Melzer, eds.) Carnegie Rochester Conferences on Public Policy (*Supplementary series to the Journal of Monetary Economics*)," Amsterdam, 19-64.
- Manski, C. (1995). *Identification Problems in the Social Sciences*. Massachusetts: Harvard Univ. Press.
- Mosteller, F. and Tukey, J. W. (1977). *Data Analysis and Regression*. Philippines: Addison-Wesley Publishing Company.
- Nair, K. and Shritastava, M. (1942). "On a simple method of curve fitting," *Sankhya*, 6, 121-132.
- Nevitt, J. and Tam, H (1998). "A comparison of robust and nonparametric estimators under the simple linear regression model," *Multiple Linear Regression Viewpoints*, 25, 54-69.
- Neyman, J. (1923). "Sur les applications de la theorie des probabilities aux experiences essai des in principes," *Roczniki Nauh Rolnichi*, 10, 1-51. (In Polish English translation by Dabrawaska and T. Speed. *Statist. Sci.* 5 (1990), 463-480).
- Öztürk, L. (2003). *Doğrusal Regresyonda Sağlam Kestirim Yöntemleri ve Karşılaştırılmaları*. Mimar Sinan Üniversitesi, Doktora Tezi, İstanbul.

- Pearl, J. (1995). "Casual diagrams for empirical research," *Biometika*, 82, 689-709.
- Rousseuw, P. J. (1984). "Least median of squares regression," *J. Am. Stat. Assoc.*, 79, 871-880.
- Rousseuw, P. J. and Leroy, A. (1987). *Robust Regression and Outlier Detection*. New York: John Wiley & Sons, Inc.
- Rubin, D. (1974). "Estimating casual effects of treatments in randomized and nonrandomized studies," *Journal of Educational Psychology*, 66, 688-701.
- Şanlı, K. (2005). *Bulanık Robust Regresyon Çözümlemesi*. Ankara Üniversitesi, Doktora Tezi, Ankara.
- Tasker, G. and Granato, G. (2000). "Statistical approaches to interpretation of local, regional and national highway -Runoff and Urban- Stormwater Data," *U.S. Geological Survey*, 1-47.
- Theil, H. (1950). "A rank-invariant method of linear and polynomial regression analysis," *Ned. Akad. Wetensch Proc.* 53, 386-392.
- Tukey, J. (1970). *Exploratory Data Analysis*. MA: Addison-Wesley.
- Wald, A. (1940). "The fitting of straight lines if both variables are subject to error," *Annals of Mathematical Statistics*, 11, 284-300.
- Yale, C. and Forsythe, A. (1976). "Winsorized regression," *Technometrics*, 18(3), 291-312.
- Yule, G. (1899). "An investigation into the causes of changes in pauperism in England, chiefly during the last two intercensal decades," *J. Roy. Stat. Soc.*, 62, 249-295.
- Yorulmaz, Ö. (2003). *Robust Regresyon ve Mathematica Uygulamaları*. Marmara Üniversitesi, Yüksek Lisans Tezi, Ankara.

EK-1**Minnesota Goller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları**

Gölün Adı	LGR(x)	MBT(y)
Mille Lacs Lake	13,5	23,6
Upper Red Lake	23,7	26,3
Lower Red Lake	15,4	24,1
Lake Of The Woods	17,9	23,9
Leech Lake	11,9	23,4
Big Bay	3,82	21,2
Lake Kabetogama	6,65	22,3
Lake Winnibigoshish	6,09	21,7
Lake Pepin	5,48	26,0
Rainy Lake	3,52	18,5
Cass Lake 1	2,55	15,7
Cass Lake 3	1,94	14,4
Cass Lake 4	3,45	18,9

EK-2

Iowa Goller Bölgesindeki Göllere ait CPUE_w ve Chlorophyll α Değerleri

CPUE _w	Chlorophyll α	logChlorophyll α (x)	logCPUE _w (y)
5150	47	1,672097858	3,711807229
2390	11,6	1,064457989	3,378397901
4834	32,4	1,51054501	3,684306646
1868	11,9	1,075546961	3,271376872
8035	98,5	1,99343623	3,904985881
2392	14,1	1,149219113	3,378761175
6033	50	1,698970004	3,780533325
5162	12,8	1,10720997	3,712818
22131	165	2,217483944	4,345001038
2373	17,3	1,238046103	3,375297738
20333	9,9	0,995635195	4,308201461
5153	26,1	1,416640507	3,712060142
8587	17,4	1,240549248	3,933841463
3755	11	1,041392685	3,574609941
4434	11,5	1,06069784	3,646795689
983	58	1,763427994	2,992553518
2149	20,9	1,320146286	3,332236415
2400	16,2	1,209515015	3,380211242
3434	23,2	1,365487985	3,535800291
1907	36,6	1,563481085	3,280350693
1831	17	1,230448921	3,262688344
6373	21,5	1,33243846	3,804343918
14311	26,4	1,421603927	4,155669982
3683	32	1,505149978	3,566201719
1881	25,7	1,409933123	3,274388796
746	7	0,84509804	2,872738827
1323	22,6	1,354108439	3,121559844
3416	98	1,991226076	3,533517862
1374	4	0,602059991	3,137986733
15173	135,4	2,131618664	4,181071458
3705	28,5	1,45484486	3,568788212
1208	11,1	1,045322979	3,082066934

EK-3**Beton Kırma Denemesine ait Enerji ve be3 değerleri**

be3(x)	Enerji(y)
20,63	2,3
6,8	2
14,51	1,3
80,4	8,3
24,83	6,77
114,08	26,3
35,14	16,2
63,25	16,8
46,2	19,2
42,6	16
77,87	29,9
167,3	22,4
60,95	19,8
36,63	8,16
6,58	7,4
1512,58	42,5
61,88	11
87,05	5,5
308,78	22
474,19	37,5
964,45	58
199,56	59,6
204,88	43,43
1006,5	88,24
2491,24	103,09

EK-4

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen EKK Sonuçları

Regression Analizi Sonuçları EKK Tekniği

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,736 ^a	,541	,499	2,6190

a. Predictors: (Constant), LGR

b. Dependent Variable: MBT

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	88,940	1	88,940	12,967	,004 ^a
	Residual	75,451	11	6,859		
	Total	164,391	12			

a. Predictors: (Constant), LGR

b. Dependent Variable: MBT

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	18,018	1,218		14,793	,000
	LGR	,395	,110	,736	3,601	,004

a. Dependent Variable: MBT

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen LAD Sonuçları

LAD Tekniği için 1. İterasyon Tablosu

y	x	m	Sıralı m	$x_i - x_0$	$ x_i - x_0 $	Kümülatif $ x_i - x_0 $
23,6	13,5	x	x	x	x	x
26,3	23,7	0,264706	-0,29925	10,2	8,02	8,02
24,1	15,4	0,263158	0,068182	1,9	4,4	12,42
23,9	17,9	0,068182	0,125	4,4	1,6	14,02
23,4	11,9	0,125	0,189781	-1,6	6,85	20,87
21,2	3,82	0,247934	0,247934	-9,68	9,68	30,55
22,3	6,65	0,189781	0,25641	-6,85	7,41	37,96
21,7	6,09	0,25641	0,263158	-7,41	1,9	39,86
26	5,48	-0,29925	0,264706	-8,02	10,2	50,06
18,5	3,52	0,511022	0,467662	-9,98	10,05	60,11
15,7	2,55	0,721461	0,511022	-10,95	9,98	70,09
14,4	1,94	0,795848	0,721461	-11,56	10,95	81,04
18,9	3,45	0,467662	0,795848	-10,05	11,56	92,6

LAD Tekniği için 2. İterasyon Tablosu

k=2						
y	x	m	Sıralı m	$x_i - x_0$	$ x_i - x_0 $	Kümülatif $ x_i - x_0 $
23,6	13,5	0,264706	0,016465	-10,2	18,22	18,22
26,3	23,7	x	x	x	x	x
24,1	15,4	0,26506	0,234604	-8,3	17,05	35,27
23,9	17,9	0,413793	0,245763	-5,8	11,8	47,07
23,4	11,9	0,245763	0,256539	-11,8	19,88	66,95
21,2	3,82	0,256539	0,261215	-19,88	17,61	84,56
22,3	6,65	0,234604	0,264706	-17,05	10,2	94,76
21,7	6,09	0,261215	0,26506	-17,61	8,3	103,06
26	5,48	0,016465	0,365443	-18,22	20,25	123,31
18,5	3,52	0,386521	0,386521	-20,18	20,18	143,49
15,7	2,55	0,501182	0,413793	-21,15	5,8	149,29
14,4	1,94	0,546875	0,501182	-21,76	21,15	170,44
18,9	3,45	0,365432	0,546875	-20,25	21,76	192,2

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen LAD Sonuçları

LAD Tekniği için 3.İterasyon Tablosu

k=3						
y	x	m	Sıralı m	$x_i - x_0$	$ x_i - x_0 $	Kümülatif $ x_i - x_0 $
23,6	13,5	0,263158	-0,19153	-1,9	9,92	9,92
26,3	23,7	0,26506	-0,08	8,3	2,5	12,42
24,1	15,4	x	x	X	x	x
23,9	17,9	-0,08	0,2	2,5	3,5	15,92
23,4	11,9	0,2	0,205714	-3,5	8,75	24,67
21,2	3,82	0,250432	0,250432	-11,58	11,58	36,25
22,3	6,65	0,205714	0,257787	-8,75	9,31	45,56
21,7	6,09	0,257787	0,263158	-9,31	1,9	47,46
26	5,48	-0,19153	0,26506	-9,92	8,3	55,76
18,5	3,52	0,47138	0,435146	-11,88	11,95	67,71
15,7	2,55	0,653696	0,47138	-12,85	11,88	79,59
14,4	1,94	0,720654	0,653696	-13,46	12,85	92,44
18,9	3,45	0,435146	0,720654	-11,95	13,46	105,9

$$\text{LAD Tekniği için } \hat{\beta}_1 = \frac{y_k - y_0}{x_k - x_0} = 0,265 \quad , \quad \hat{\beta}_0 = y_0 - \hat{\beta}_1 x_0 = 20,018 \text{ olarak}$$

hesaplanmıştır.

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen Huber-M Algoritma Sonuçları

Huber-M Algoritma Tekniği için 1.İterasyon Tablosu

y	x	\hat{y}	ε_i	$\varepsilon_i - med\varepsilon_i$	$ \varepsilon_i - med\varepsilon_i $	ε^*	y^*
23,6	13,5	23,3505	0,2495	0,2505	0,2505	0,2495	23,6
26,3	23,7	27,3795	-1,0795	-1,0785	1,0785	-1,0795	26,3
24,1	15,4	24,101	-0,001	0	0	-0,001	24,1
23,9	17,9	25,0885	-1,1885	-1,1875	1,1875	-1,1885	23,9
23,4	11,9	22,7185	0,6815	0,6825	0,6825	0,6815	23,4
21,2	3,82	19,5269	1,6731	1,6741	1,6741	1,6731	21,2
22,3	6,65	20,64475	1,65525	1,65625	1,65625	1,65525	22,3
21,7	6,09	20,42355	1,27645	1,27745	1,27745	1,27645	21,7
26	5,48	20,1826	5,8174	5,8184	5,8184	2,6408	22,8234
18,5	3,52	19,4084	-0,9084	-0,9074	0,9074	-0,9084	18,5
15,7	2,55	19,02525	-3,32525	-3,32425	3,32425	-2,6408	16,38445
14,4	1,94	18,7843	-4,3843	-4,3833	4,3833	-2,6408	16,1435
18,9	3,45	19,38075	-0,48075	-0,47975	0,47975	-	18,9
		$med\varepsilon_i$	-0,001	med	1,1875		
				MSM	1,7605634		
				k	2,6408451		

Regression Analizi Sonuçları Huber-M Algoritma 1.İterasyon

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,851 ^a	,725	,700	1,7096

a. Predictors: (Constant), LGR

b. Dependent Variable: YSTAR

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	84,568	1	84,568	28,935	,000 ^a
	Residual	32,150	11	2,923		
	Total	116,718	12			

a. Predictors: (Constant), LGR

b. Dependent Variable: YSTAR

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen Huber-M Sonuçları

Regression Analizi Sonuçları Huber-M Algoritma 1.İterasyon

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	18,048	,795		22,699	,000
	LGR	,385	,072	,851	5,379	,000

a. Dependent Variable: YSTAR

Huber-M Algoritma Tekniği için Final İterasyon Tablosu

\hat{y}	ε_i	$\varepsilon_i - med\varepsilon_i$	$ \varepsilon_i - med\varepsilon_i $	ε^*	y^*
23,24651	0,35349	0,22854	0,22854	0,35349	23,6
27,15762	-0,85762	-0,98257	0,98257	-0,85762	26,3
23,97505	0,12495	0	0	0,12495	24,1
24,93366	-1,03366	-1,15861	1,15861	-1,03366	23,9
22,63301	0,76699	0,64204	0,64204	0,76699	23,4
19,5348	1,6652	1,54025	1,54025	1,6652	21,2
20,61994	1,68006	1,55511	1,55511	1,68006	22,3
20,40521	1,29479	1,16984	1,16984	1,29479	21,7
20,17131	2,57169	2,44674	2,44674	2,57169	22,743
19,41976	-0,91976	-1,04471	1,04471	-0,91976	18,5
19,04783	-2,5766	-2,70155	2,70155	-2,576597	16,47123
18,81393	-2,57662	-2,70157	2,70157	-2,576597	16,23733
19,39292	-0,49292	-0,61787	0,61787	-0,49292	18,9
$med\varepsilon_i$	0,12495	med	1,15861		
		MSM	1,717732		
		k	2,576597		

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen Huber-M Algoritma Sonuçları

Regression Analizi Sonuçları Huber-M Algoritma Final İterasyon

Model Summary ^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,855 ^a	,731	,707	1,6729

a. Predictors: (Constant), LGR

b. Dependent Variable: YSTAR7

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	18,070	,778		23,226	,000
	LGR	,383	,070	,855	5,474	,000

a. Dependent Variable: YSTAR7

İterasyon Sonuçları

İterasyon	$\hat{\beta}_0$	$\hat{\beta}_1$
1	18,04779593808	0,3850733632867
2	18,06026558419	0,384073489262
3	18,06578592047	0,3837170235886
4	18,06822224791	0,3835597133493
5	18,06934065830	0,3834875016576
6	18,07004937429	0,3834417440394
7	18,07005418076	0,3834414292509

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen Huber-M Ağırlıklı Regresyon Sonuçları

Huber-M Ağırlıklı Regresyon Tekniği İçin 1. İterasyon Tablosu

y	x	\hat{y}	ε_i	$\varepsilon_i - med\varepsilon_i$	$ \varepsilon_i - med\varepsilon_i $	r_i	$ r_i $	w_i
23,6	13,5	23,3505	0,2495	0,2505	0,2505	0,141716	0,141716	1
26,3	23,7	27,3795	-1,0795	-1,0785	1,0785	-0,61316	0,613156	1
24,1	15,4	24,101	-0,001	0	0	-0,00057	0,000568	1
23,9	17,9	25,0885	-1,1885	-1,1875	1,1875	-0,67507	0,675068	1
23,4	11,9	22,7185	0,6815	0,6825	0,6825	0,387092	0,387092	1
21,2	3,82	19,5269	1,6731	1,6741	1,6741	0,950321	0,950321	1
22,3	6,65	20,64475	1,65525	1,65625	1,65625	0,940182	0,940182	1
21,7	6,09	20,42355	1,27645	1,27745	1,27745	0,725024	0,725024	1
26	5,48	20,1826	5,8174	5,8184	5,8184	3,304283	3,304283	0,453956
18,5	3,52	19,4084	-0,9084	-0,9074	0,9074	-0,51597	0,515971	1
15,7	2,55	19,02525	-3,32525	-3,32425	3,32425	-1,88874	1,888742	0,794179
14,4	1,94	18,7843	-4,3843	-4,3833	4,3833	-2,49028	2,490282	0,602341
18,9	3,45	19,38075	-0,48075	-0,47975	0,47975	-0,27307	0,273066	1
		$med\varepsilon_i$	-0,001	med	1,1875			
				MSM	1,760563			

Huber-M Ağırlıklı Regresyonu için $w = \begin{cases} 1 & , |r| \leq 1,5 \\ \frac{1,5}{|r|} & , |r| > 1,5 \end{cases}$ ağırlık değerleri

kullanılmıştır.

Regression Analizi Sonuçları Huber-M Ağırlıklı Regresyon 1.İterasyon

Model Summary ^{b,c}

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,792 ^a	,627	,594	2,0660

a. Predictors: (Constant), LGR

b. Dependent Variable: MBT

c. Weighted Least Squares Regression - Weighted by W1

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen Huber-M Ağırlıklı Regresyon Sonuçları

Regression Analizi Sonuçları Huber-M Ağırlıklı Regresyon 1.İterasyon

Coefficients^{a,b}

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	18,047	1,035		17,444	,000
	LGR	,385	,089	,792	4,304	,001

a. Dependent Variable: MBT

b. Weighted Least Squares Regression - Weighted by W1

Huber-M Ağırlıklı Regresyon Tekniği için Final İterasyon Tablosu

\hat{y}	ε_i	$\varepsilon_i - med\varepsilon_i$	$ \varepsilon_i - med\varepsilon_i $	r_i	$ r_i $	w_i
23,24682	0,35318	0,22844	0,22844	0,205633	0,205633	1
27,15738	-0,85738	-0,98212	0,98212	-0,4992	0,499195	1
23,97526	0,12474	0	0	0,072628	0,072628	1
24,93373	-1,03373	-1,15847	1,15847	-0,60187	0,601872	1
22,6334	0,7666	0,64186	0,64186	0,44634	0,44634	1
19,53562	1,66438	1,53964	1,53964	0,969058	0,969058	1
20,62061	1,67939	1,55465	1,55465	0,977797	0,977797	1
20,40591	1,29409	1,16935	1,16935	0,753463	0,753463	1
20,17204	5,82796	5,70322	5,70322	3,393233	3,393233	0,442056
19,4206	-0,9206	-1,04534	1,04534	-0,536	0,536004	1
19,04872	-3,34872	-3,47346	3,47346	-1,94974	1,949737	0,769335
18,81485	-4,41485	-4,53959	4,53959	-2,57047	2,570473	0,58355
19,39377	-0,49377	-0,61851	0,61851	-0,28749	0,287489	1
$med\varepsilon_i$	0,12474	med	1,15847			
		MSM	1,717524			

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen Huber-M Ağırlıklı Regresyon Sonuçları

Regression Analizi Sonuçları Huber-M Ağırlıklı Regresyon Final İterasyon

Coefficients ^{a,b}

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	18,071	1,027		17,588	,000
	LGR	,383	,089	,793	4,324	,001

a. Dependent Variable: MBT

b. Weighted Least Squares Regression - Weighted by W7

Model Summary ^{b,c}

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,793 ^a	,630	,596	2,0428

a. Predictors: (Constant), LGR

b. Dependent Variable: MBT

c. Weighted Least Squares Regression - Weighted by W7

İterasyon Sonuçları

İterasyon	$\hat{\beta}_0$	$\hat{\beta}_1$
1	18,04681270224	0,3851028073005
2	18,06173347385	0,383989963941
3	18,06759816156	0,3836112093898
4	18,06986503050	0,3834659536072
5	18,07073882172	0,3834099835334
6	18,07107579286	0,3833883972465
7	18,07121089844	0,3833797018225

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen Hampel Ağırlıklı Regresyon Sonuçları

Hampel Ağırlıklı Regresyon Tekniği için 1. İterasyon Tablosu

y	x	\hat{y}	ε_i	$\varepsilon_i - med\varepsilon_i$	$ \varepsilon_i - med\varepsilon_i $	$ r_i $	w_i
23,6	13,5	23,3505	0,2495	0,2505	0,2505	0,141716	1
26,3	23,7	27,3795	-1,0795	-1,0785	1,0785	0,613156	1
24,1	15,4	24,101	-0,001	0	0	0,000568	1
23,9	17,9	25,0885	-1,1885	-1,1875	1,1875	0,675068	1
23,4	11,9	22,7185	0,6815	0,6825	0,6825	0,387092	1
21,2	3,82	19,5269	1,6731	1,6741	1,6741	0,950321	1
22,3	6,65	20,64475	1,65525	1,65625	1,65625	0,940182	1
21,7	6,09	20,42355	1,27645	1,27745	1,27745	0,725024	1
26	5,48	20,1826	5,8174	5,8184	5,8184	3,304283	0,514484
18,5	3,52	19,4084	-0,9084	-0,9074	0,9074	0,515971	1
15,7	2,55	19,02525	-3,32525	-3,32425	3,32425	1,888742	0,90007
14,4	1,94	18,7843	-4,3843	-4,3833	4,3833	2,490282	0,682654
18,9	3,45	19,38075	-0,48075	-0,47975	0,47975	0,273066	1
		$med\varepsilon_i$	-0,001	med	1,1875		
				MSM	1,760563		

$$w = \begin{cases} 1 & , 0 < |r| \leq 1,7 \\ \frac{1,7}{r} \operatorname{sgn}(r) & , 1,7 < |r| \leq 3,4 \\ \frac{1,7}{r} \left[\frac{8,5 - |r|}{8,5 - 3,4} \right] \operatorname{sgn}(r) & , 3,4 < |r| \leq 8,5 \\ 0 & , 8,5 < |r| \end{cases}$$

Hampel Ağırlıklı Regresyonu için

ağırlık değerleri kullanılmıştır.

Regression Analizi Sonuçları Hampel Ağırlıklı Regresyon 1.İterasyon

Model Summary^{b,c}

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,786 ^a	,617	,583	2,1681

a. Predictors: (Constant), LGR

b. Dependent Variable: MBT

c. Weighted Least Squares Regression - Weighted by W

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen Hampel Ağırlıklı Regresyon Sonuçları

Regression Analizi Sonuçları Hampel Ağırlıklı Regresyon 1. İterasyon

Coefficients^{a,b}

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	17,956	1,065		16,855	,000
	LGR	,392	,093	,786	4,212	,001

a. Dependent Variable: MBT

b. Weighted Least Squares Regression - Weighted by W

Hampel Ağırlıklı Regresyon Tekniği için Final İterasyon Tablosu

\hat{y}	ε_i	$\varepsilon_i - med\varepsilon_i$	$ \varepsilon_i - med\varepsilon_i $	$ r_i $	w_i
23,23714	0,36286	0,24891	0,24891	0,20647	1
27,2576	-0,9576	-1,07155	1,07155	0,54488	1
23,98605	0,11395	0	0	0,064838	1
24,97145	-1,07145	-1,1854	1,1854	0,609662	1
22,60647	0,79353	0,67958	0,67958	0,451524	1
19,42163	1,77837	1,66442	1,66442	1,011904	1
20,53712	1,76288	1,64893	1,64893	1,00309	1
20,31639	1,38361	1,26966	1,26966	0,787283	1
20,07595	5,92405	5,8101	5,8101	3,370821	0,504328
19,30339	-0,80339	-0,91734	0,91734	0,457134	1
18,92105	-3,22105	-3,335	3,335	1,832798	0,927544
18,68061	-4,28061	-4,39456	4,39456	2,435694	0,697953
19,27579	-0,37579	-0,48974	0,48974	0,213827	1
$med\varepsilon_i$	0,11395	med	1,1854		
		MSM	1,75745		

EK-4 Devam**Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen Hampel Ağırlıklı Regresyon Sonuçları**

Regression Analizi Sonuçları Hampel Ağırlıklı Regresyon Final İterasyon

Coefficients^{a,b}

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	
	B	Std. Error	Beta			
1	(Constant)	17,916	1,064		16,835	,000
	LGR	,394	,093	,787	4,237	,001

a. Dependent Variable: MBT

b. Weighted Least Squares Regression - Weighted by W7

İterasyon Sonuçları

İterasyon	$\hat{\beta}_0$	$\hat{\beta}_1$
1	17,95583080403	0,3917674854045
2	17,93280270908	0,3930842982297
3	17,92303608997	0,3937074666976
4	17,91884228015	0,3939765130449
5	17,91703966432	0,3940921632432
6	17,91626223698	0,3941420507908
7	17,91593088004	0,3941632899858
8	17,9157918585	0,3941721998536

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen Andrews Ağırlıklı Regresyon Sonuçları

Andrews Ağırlıklı Regresyon Tekniği için 1. İterasyon Tablosu

y	x	\hat{y}	ε_i	$\varepsilon_i - med\varepsilon_i$	$ \varepsilon_i - med\varepsilon_i $	r_i	$ r_i $	$\sin(r_i /1,5)$	w_i
23,6	13,5	23,3505	0,2495	0,2505	0,2505	0,14172	0,14172	0,0943	0,665
26,3	23,7	27,3795	-1,0795	-1,079	1,0785	-0,6132	0,61316	0,3975	0,6483
24,1	15,4	24,101	-0,001	0	0	-0,0006	0,00057	0,0004	0,6667
23,9	17,9	25,0885	-1,1885	-1,188	1,1875	-0,6751	0,67507	0,435	0,6444
23,4	11,9	22,7185	0,6815	0,6825	0,6825	0,38709	0,38709	0,2552	0,6593
21,2	3,82	19,5269	1,6731	1,6741	1,6741	0,95032	0,95032	0,592	0,623
22,3	6,65	20,6448	1,6553	1,6563	1,6563	0,94018	0,94018	0,5865	0,6239
21,7	6,09	20,4236	1,2765	1,2775	1,2775	0,72502	0,72502	0,4647	0,641
26	5,48	20,1826	5,8174	5,8184	5,8184	3,30428	3,30428	0,8068	0,2442
18,5	3,52	19,4084	-0,9084	-0,907	0,9074	-0,516	0,51597	0,3372	0,6536
15,7	2,55	19,0253	-3,3253	-3,324	3,3243	-1,8887	1,88874	0,9518	0,504
14,4	1,94	18,7843	-4,3843	-4,383	4,3833	-2,4903	2,49028	0,996	0,399
18,9	3,45	19,3808	-0,4808	-0,48	0,4798	-0,2731	0,27307	0,181	0,663
		$med\varepsilon_i$	-0,001	med	1,1875				
				MSM	1,7606				

Andrews Ağırlıklı Regresyonu için
$$w = \begin{cases} \frac{\sin(\frac{r}{1,5})}{r} & , |r| \leq 1,5\pi \\ 0 & , |r| > 1,5\pi \end{cases}$$

ağırlık değerleri kullanılmıştır.

Regression Analizi Sonuçları Andrews Ağırlıklı Regresyon 1. İterasyon

Model Summary ^{b,c}

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,805 ^a	,648	,616	1,6130

a. Predictors: (Constant), LGR

b. Dependent Variable: MBT

c. Weighted Least Squares Regression - Weighted by W

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen Andrews Ağırlıklı Regresyon Sonuçları

Regression Analizi Sonuçları Andrews Ağırlıklı Regresyon 1. İterasyon

Coefficients ^{a,b}

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	17,948	1,009		17,786	,000
	LGR	,390	,087	,805	4,496	,001

a. Dependent Variable: MBT

b. Weighted Least Squares Regression - Weighted by W

Andrews Ağırlıklı Regresyon Tekniği için Final İterasyon Tablosu

\hat{y}	ε_i	$\varepsilon_i - med\varepsilon_i$	$ \varepsilon_i - med\varepsilon_i $	$ r_i $	$ r_i /1,5$	$\sin(r_i /1,5)$	w_i
23,197	0,40314	0,24771	0,24771	0,22969	0,1531	0,15253	0,6641
27,211	-0,91089	-1,0663	1,06632	0,519	0,346	0,33913	0,6534
23,945	0,15543	0	0	0,08856	0,059	0,059	0,6663
24,928	-1,0284	-1,1838	1,18383	0,5859	0,3906	0,38077	0,6498
22,567	0,83279	0,67736	0,67736	0,47449	0,3163	0,31108	0,6556
19,387	1,81253	1,6571	1,6571	1,03271	0,6885	0,63536	0,6152
20,501	1,79884	1,64341	1,64341	1,02491	0,6833	0,63133	0,616
20,281	1,41922	1,26379	1,26379	0,80862	0,5391	0,51334	0,6348
20,041	5,95927	5,80384	5,80384	3,39536	2,2636	0,76948	0,2266
19,269	-0,76941	-0,9248	0,92484	0,4384	0,2923	0,28811	0,6572
18,888	-3,18768	-3,3431	3,34311	1,8162	1,2108	0,9359	0,5153
18,648	-4,24763	-4,4031	4,40306	2,4201	1,6134	0,99909	0,4128
19,242	-0,34186	-0,4973	0,49729	0,1948	0,1299	0,12949	0,6648
$med\varepsilon_i$	0,15543	med	1,18383				
		MSM	1,75512				

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen Andrews Ağırlıklı Regresyon Sonuçları

Regression Analizi Sonuçları Andrews Ağırlıklı Regresyon Final İterasyon

Coefficients^{a,b}

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	17,884	1,002		17,849	,000
	LGR	,394	,086	,810	4,573	,001

a. Dependent Variable: MBT

b. Weighted Least Squares Regression - Weighted by W6

İterasyon Sonuçları

İterasyon	$\hat{\beta}_0$	$\hat{\beta}_1$
1	17,94760583713	0,3902840116331
2	17,90743612079	0,3921047842147
3	17,89206093747	0,3930304623753
4	17,88666264399	0,3933726099856
5	17,88480762735	0,3934914593547
6	17,88417359200	0,3935321434652
7	17,88395904890	0,393546051285

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen Tukey Ağırlıklı Regresyon Sonuçları

Tukey Ağırlıklı Regresyon Tekniği için 1. İterasyon Tablosu

y	x	\hat{y}	ε_i	$\varepsilon_i - med\varepsilon_i$	$ \varepsilon_i - med\varepsilon_i $	$ r_i $	$1 - (r_i /5)^2$	w_i
23,6	13,5	23,351	0,2495	0,2505	0,2505	0,1417	0,999	0,998
26,3	23,7	27,38	-1,0795	-1,079	1,0785	0,6132	0,985	0,97
24,1	15,4	24,101	-0,001	0	0	0,0006	1	1
23,9	17,9	25,089	-1,1885	-1,188	1,1875	0,6751	0,982	0,964
23,4	11,9	22,719	0,6815	0,6825	0,6825	0,3871	0,994	0,988
21,2	3,82	19,527	1,6731	1,6741	1,6741	0,9503	0,964	0,929
22,3	6,65	20,645	1,6553	1,6563	1,65625	0,9402	0,965	0,931
21,7	6,09	20,424	1,2765	1,2775	1,27745	0,725	0,979	0,958
26	5,48	20,183	5,8174	5,8184	5,8184	3,3043	0,563	0,317
18,5	3,52	19,408	-0,9084	-0,907	0,9074	0,516	0,989	0,979
15,7	2,55	19,025	-3,3253	-3,324	3,32425	1,8887	0,857	0,735
14,4	1,94	18,784	-4,3843	-4,383	4,3833	2,4903	0,752	0,565
18,9	3,45	19,381	-0,4808	-0,48	0,47975	0,2731	0,997	0,994
		$med\varepsilon_i$	-0,001	med	1,1875			
				MSM	1,76056			

Tukey Ağırlıklı Regresyonu için $w = \begin{cases} \left(1 - \left(\frac{r}{5}\right)^2\right)^2 & , |r| \leq 5 \\ 0 & , |r| > 5 \end{cases}$ ağırlık değerleri

kullanılmıştır.

Regression Analizi Sonuçları Tukey Ağırlıklı Regresyon 1. İterasyon

Model Summary ^{b,c}				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,812 ^a	,659	,628	1,9143

a. Predictors: (Constant), LGR

b. Dependent Variable: MBT

c. Weighted Least Squares Regression - Weighted by W

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen Tukey Ağırlıklı Regresyon Sonuçları

Regression Analizi Sonuçları Tukey Ağırlıklı Regresyon 1.İterasyon

Coefficients^{a,b}

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	17,945	,986		18,193	,000
	LGR	,390	,085	,812	4,612	,001

a. Dependent Variable: MBT

b. Weighted Least Squares Regression - Weighted by W

Tukey Ağırlıklı Regresyon Tekniği için Final İterasyon Tablosu

\hat{y}	ε_i	$\varepsilon_i - med\varepsilon_i$	$ \varepsilon_i - med\varepsilon_i $	$ r_i $	$1 - (r_i /5)^2$	w_i
23,1833	0,4168	0,2479	0,24791	0,2374	0,9977	0,995
27,1983	-0,8983	-1,0672	1,06718	0,5117	0,9895	0,979
23,9312	0,1688	0	0	0,0962	0,9996	0,999
24,9153	-1,0153	-1,1841	1,18409	0,5783	0,9866	0,973
22,5534	0,8466	0,6777	0,67773	0,4822	0,9907	0,981
19,3729	1,8272	1,6583	1,65831	1,0408	0,9567	0,915
20,4868	1,8132	1,6443	1,64432	1,0328	0,9573	0,916
20,2664	1,4336	1,2648	1,26476	0,8166	0,9733	0,947
20,0263	5,9737	5,8049	5,80488	3,4028	0,5368	0,288
19,2548	-0,7548	-0,9236	0,92359	0,4299	0,9926	0,985
18,8729	-3,1729	-3,3418	3,34177	1,8074	0,8693	0,756
18,6328	-4,2328	-4,4017	4,40165	2,4112	0,7675	0,589
19,2272	-0,3272	-0,496	0,49604	0,1864	0,9986	0,997
$med\varepsilon_i$	0,1688	med	1,18409			
		MSM	1,75551			

EK-4 Devam**Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen Tukey Ağırlıklı Regresyon Sonuçları**

Regression Analizi Sonuçları Tukey Ağırlıklı Regresyon Final İterasyon

Coefficients ^{a,b}

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	17,869	,979		18,262	,000
	LGR	,394	,084	,817	4,704	,001

a. Dependent Variable: MBT

b. Weighted Least Squares Regression - Weighted by W7

İterasyon Sonuçları

İterasyon	$\hat{\beta}_0$	$\hat{\beta}_1$
1	17,94502975332	0,3897152733823
2	17,89979411685	0,3917548756164
3	17,88075805616	0,3929001586384
4	17,87338699667	0,3933659607803
5	17,87059765412	0,3935442179472
6	17,86954795692	0,3936113814125
7	17,86915355584	0,3936366729042
8	17,86900591736	0,3936460858312

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen LMS Sonuçları

Mümkün olan tüm ikili çiftleri için $med\varepsilon_i^2$ değerleri

$\hat{\beta}_0$	$\hat{\beta}_1$		$med\varepsilon_i^2$
0,264706	20,02647		0,26341
0,263158	20,04737		0,252636
0,068182	22,67955		3,027618
0,125	21,9125		1,4161
0,247934	20,25289		0,158682
0,189781	21,03796		0,316883
0,25641	20,13846		0,208313
-0,29925	27,6399		16,953
0,511022	16,7012		3,7966
0,721461	13,86027		6,506
0,795848	12,85606		10,25113
0,467662	17,28657		3,08958
0,26506	20,01807		0,26965
0,413793	16,4931		3,41611
0,245763	20,47542		0,074047
.	.		
.	.		
.	.		
3,555556	6,633333		43,38421
2,980132	8,618543		25,68003

Verilerden hareketle, iki noktası bilinen doğrunun eğimi bilgisini kullanarak, tüm mümkün durumlar için $\hat{\beta}_1 = \frac{y_i - y_0}{x_i - x_0}$ ve $\hat{\beta}_0 = y_0 - \hat{\beta}_1 x_0$ hesaplanır. Mümkün olan tüm veri çiftleri için $med\varepsilon_i^2$ değeri hesaplanır. Sonraki adımda min $med\varepsilon_i^2$ değerine sahip $\hat{\beta}_1 = \frac{y_i - y_0}{x_i - x_0}$ ve $\hat{\beta}_0 = y_0 - \hat{\beta}_1 x_0$ değerleri aranan parametrelerdir.

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen LMS Sonuçları

LMS Tekniği için Ağırlıklanmış EKK Uygulamak İçin Oluşturulan Tablo

y	x	\hat{y}	ε_i	ε_i^2	$\frac{\varepsilon_i}{s_0}$	$\left \frac{\varepsilon_i}{s_0} \right $	w_i
23,6	13,5	23,79322	-0,19322	0,037334	-0,32927	0,329266	1
26,3	23,7	26,3	-3,1E-06	9,61E-12	-5,3E-06	5,28E-06	1
24,1	15,4	24,26017	-0,16017	0,025654	-0,27295	0,272945	1
23,9	17,9	24,87458	-0,97458	0,949802	-1,66077	1,660773	1
23,4	11,9	23,4	3E-07	9E-14	5,11E-07	5,11E-07	1
21,2	3,82	21,41423	-0,21423	0,045896	-0,36508	0,365076	1
22,3	6,65	22,10974	0,190256	0,036197	0,324214	0,324214	1
21,7	6,09	21,97212	-0,27212	0,074047	-0,46371	0,463713	1
26	5,48	21,8222	4,177799	17,454	7,119364	7,119364	0
18,5	3,52	21,34051	-2,84051	8,068473	-4,84049	4,84049	0
15,7	2,55	21,10212	-5,40212	29,18285	-9,20572	9,205716	0
14,4	1,94	20,9522	-6,5522	42,93133	-11,1656	11,16557	0
18,9	3,45	21,3233	-2,4233	5,872394	-4,12954	4,129536	0
			$med\varepsilon_i^2$	0,074047			
			$\sqrt{med\varepsilon_i^2}$	0,272117			
			$1 + \frac{5}{n-p}$	1,454545			
			s_0	0,586822			

Bu aşamada ise yeniden ağırlıklandırılmış EKK tekniği uygulanır.

Regression Analizi Sonuçları LMS İçin Yeniden Ağırlıklandırılmış EKK Tekniği

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,978 ^a	,956	,949	,3645

a. Predictors: (Constant), LGR

EK-4 Devam**Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen LMS Sonuçları****Regression Analizi Sonuçları LMS İçin Yeniden Ağırlıklandırılmış EKK Tekniği****Coefficients^{a,b}**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	20,415	,285		71,659	,000
	LGR	,234	,021	,978	11,402	,000

a. Dependent Variable: MBT

b. Weighted Least Squares Regression - Weighted by W

EK-4 Devam**Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen LMS PROGRESS Sonuçları**

LEAST SQUARES REGRESSION

VARIABLE	COEFFICIENT	STAND. ERROR	T - VALUE	P - VALUE
----------	-------------	--------------	-----------	-----------

lgr	.39490	.10967	3.60092	.00416
-----	--------	--------	---------	--------

CONSTANT	18.01776	1.21802	14.79266	.00000
----------	----------	---------	----------	--------

SUM OF SQUARES = 75.4508

DEGREES OF FREEDOM = 11

SCALE ESTIMATE = 2.61900

VARIANCE - COVARIANCE MATRIX =

.1203D-01

-.1072D+00 .1484D+01

COEFFICIENT OF DETERMINATION (R SQUARED) = .54103

THE F-VALUE = 12.967 (WITH 1 AND 11 DF) P - VALUE = .00416

EK-4 Devam**Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen LMS PROGRESS Sonuçları**

LEAST MEDIAN OF SQUARES REGRESSION

THE MINIMIZATION OF THE 7TH ORDERED SQUARED RESIDUAL IS PERFORMED.

VARIABLE COEFFICIENT

lgr .24576

CONSTANT 20.43449

FINAL SCALE ESTIMATE = .41821

COEFFICIENT OF DETERMINATION = .97544

REWEIGHTED LEAST SQUARES BASED ON THE LMS

VARIABLE COEFFICIENT STAND. ERROR T - VALUE P - VALUE

lgr .23420 .02054 11.40186 .00003

CONSTANT 20.41544 .28490 71.65862 .00000

WEIGHTED SUM OF SQUARES = .79714

DEGREES OF FREEDOM = 6

SCALE ESTIMATE = .36449

EK-4 Devam**Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen LMS PROGRESS Sonuçları**

VARIANCE - COVARIANCE MATRIX =

.4219D-03

-.5219D-02 .8117D-01

COEFFICIENT OF DETERMINATION (R SQUARED) = .9558

THE F-VALUE = 130.002 (WITH 1 AND 6 DF) P - VALUE = .00003

THERE ARE 8 POINTS WITH NON-ZERO WEIGHT

AVERAGE WEIGHT = .61538

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen WLS Sonuçları

WLS Tekniği için 1.İterasyon Tablosu

y	x	\hat{y}	ε_i	Sıralı ε_i	ε_i'	y'
23,6	13,5	23,34894	0,25106	-4,38387	-3,324	23,6
26,3	23,7	27,37694	-1,07694	-3,32476	-3,324	26,3
24,1	15,4	24,09925	0,00075	-1,18651	-1,18651	24,1
23,9	17,9	25,08651	-1,18651	-1,07694	-1,07694	23,9
23,4	11,9	22,71709	0,68291	-0,90781	-0,90781	23,4
21,2	3,82	19,52628	1,67372	-0,48017	-0,48017	21,19928
22,3	6,65	20,64386	1,65614	0,00075	0,00075	22,3
21,7	6,09	20,42271	1,27729	0,25106	0,25106	21,7
26	5,48	20,18182	5,81818	0,68291	0,68291	21,85482
18,5	3,52	19,40781	-0,90781	1,27729	1,27729	18,5
15,7	2,55	19,02476	-3,32476	1,65614	1,65614	15,70076
14,4	1,94	18,78387	-4,38387	1,67372	1,673	15,45987
18,9	3,45	19,38017	-0,48017	5,81818	1,673	18,9

Regression Analizi Sonuçları WLS Tekniği 1.İterasyon

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,853 ^a	,727	,703	1,7941

a. Predictors: (Constant), X

b. Dependent Variable: YBIR

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	17,673	,834		21,181	,000
	X	,407	,075	,853	5,416	,000

a. Dependent Variable: YBIR

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen WLS Sonuçları

WLS Tekniği için Final İterasyon Tablosu

\hat{y}	ε_i	Sıralı ε_i	ε_i'	y'
23,16614	0,43386	-3,00059	-3,00059	23,6
27,31737	-1,01737	-3,00059	-3,00059	26,3
23,93941	0,16059	-1,05687	-1,05687	24,1
24,95687	-1,05687	-1,01737	-1,01737	23,9
22,51496	0,88504	-0,60444	-0,60444	23,4
19,22653	1,95251	-0,17595	-0,17595	21,17903
20,3783	1,9217	0,16059	0,16059	22,3
20,15039	1,54961	0,43386	0,43386	21,7
19,90213	1,9525	0,88504	0,88504	21,85463
19,10444	-0,60444	1,54961	1,54961	18,5
18,70966	-3,00059	1,9217	1,9217	15,70907
18,4614	-3,00059	1,9525	1,9525	15,46081
19,07595	-0,17595	1,95251	1,9525	18,9

Regression Analizi Sonuçları WLS Tekniği Final İterasyon

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,853 ^a	,728	,703	1,7907

a. Predictors: (Constant), X

b. Dependent Variable: YBES

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	17,672	,833		21,220	,000
	X	,407	,075	,853	5,428	,000

a. Dependent Variable: YBES

EK-4 Devam**Minnesota Goller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen WLS Sonuçları**

İterasyon Sonuçları

İterasyon	$\hat{\beta}_0$	$\hat{\beta}_1$
1	17,67340847615	0,4069061243315
2	17,6719014937	0,406981454546
3	17,6718623479	0,4069835157664
4	17,67185549614	0,406983852892
5	17,67185393038	0,4069839422347

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen LTS Sonuçları

LTS Tekniği için Tablo

y	x	\hat{y}	ε_i	ε_i^2	Sıralı ε_i^2
23,6	13,5	23,3505	0,2495	0,06225	0,000001
26,3	23,7	27,3795	-1,0795	1,16532	0,06225
24,1	15,4	24,101	-0,001	0,000001	0,231121
23,9	17,9	25,0885	-1,1885	1,412532	0,464442
23,4	11,9	22,7185	0,6815	0,464442	0,825191
21,2	3,82	19,5269	1,6731	2,799264	1,16532
22,3	6,65	20,64475	1,65525	2,739853	1,412532
21,7	6,09	20,42355	1,27645	1,629325	1,629325
26	5,48	20,1826	5,8174	33,84214	2,739853
18,5	3,52	19,4084	-0,9084	0,825191	2,799264
15,7	2,55	19,02525	-3,32525	11,05729	11,05729
14,4	1,94	18,7843	-4,3843	19,22209	19,22209
18,9	3,45	19,38075	-0,48075	0,231121	33,84214

Regression Analizi Sonuçları LTS Tekniği

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,857 ^a	,734	,708	1,9478

a. Predictors: (Constant), LGR

b. Dependent Variable: MBT

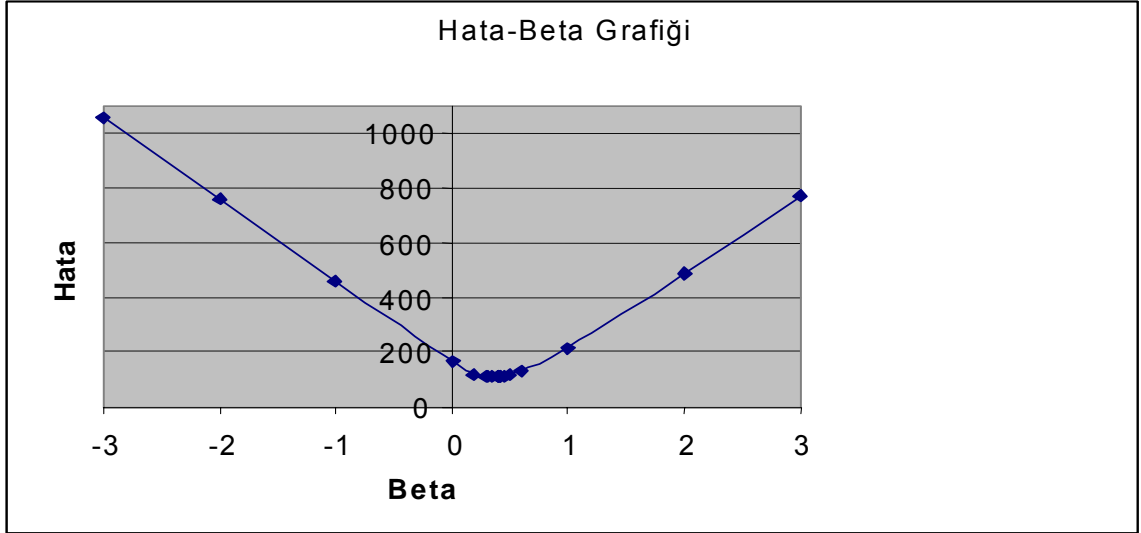
Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	17,176	,945		18,183	,000
	LGR	,434	,082	,857	5,258	,000

a. Dependent Variable: MBT

EK-4 Devam

Minnesota Göller Bölgesindeki Göllere ait Göl Geometri Oranları ve Maksimum Göl Dibi Sıcaklıkları Kullanılarak Elde Edilen R-Regresyonu Sonuçları



$\hat{\beta}_1 = 0,4$ için Wilcoxon skorları yardımıyla $\sum_{i=1}^n a_n(R_i)\varepsilon_i$ değerinin hesaplanması

y	sıra x	ε_i	R_i	$a_n(R_i)$	$a_n(R_i)\varepsilon_i$
14,4	1,94	13,624	1	-6	-81,744
15,7	2,55	14,68	2	-5	-73,4
18,5	3,52	17,092	5	-2	-34,184
18,9	3,45	17,52	6	-1	-17,52
21,2	3,82	19,672	12	5	98,36
26	5,48	23,808	13	6	142,848
21,7	6,09	19,264	10	3	57,792
22,3	6,65	19,64	11	4	78,56
23,4	11,9	18,64	9	2	37,28
23,6	13,5	18,2	8	1	18,2
23,9	17,9	16,74	3	-4	-66,96
24,1	15,4	17,94	7	0	0
26,3	23,7	16,82	4	-3	-50,46
				$\sum_{i=1}^n a_n(R_i)\varepsilon_i$	108,772

