DOĞRUSAL TİP II REGRESYON TEKNİKLERİNİN MONTE-CARLO BENZETİM ÇALIŞMASI İLE KARŞILAŞTIRILMASI: SAĞLAM, BULANIK VE SAĞLAM BULANIK TEKNİKLER

Cengiz GAZELOĞLU

DANIŞMAN

Yrd. Doç. Dr. Sinan SARAÇLI

İSTATİSTİK

HAZİRAN, 2012

AFYON KOCATEPE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

DOĞRUSAL TİP II REGRESYON TEKNİKLERİNİN MONTE-CARLO BENZETİM ÇALIŞMASI İLE KARŞILAŞTIRILMASI: SAĞLAM, BULANIK VE SAĞLAM BULANIK TEKNİKLER

Cengiz GAZELOĞLU

DANIŞMAN
Yrd. Doç. Dr. Sinan SARAÇLI

İSTATİSTİK ANABİLİM DALI

HAZİRAN 2012

TEZ ONAY SAYFASI

Cengiz GAZELOĞLU tarafından hazırlanan "Doğrusal Tip II Regresyon Tekniklerinin Monte-Carlo Benzetim Çalışması ile Karşılaştırılması: Sağlam, Bulanık ve Sağlam Bulanık Teknikler" adlı tez çalışması lisansüstü eğitim ve öğretim yönetmeliğinin ilgili maddeleri uyarınca 26/06/2012 tarihinde aşağıdaki jüri tarafından oy birliği / oy çokluğu ile Afyon Kocatepe Üniversitesi Fen Bilimleri Enstitüsü **İstatistik Anabilim Dalı'nda YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Danışman **:** Yrd. Doç. Dr. Sinan SARAÇLI

Başkan **:** Doç. Dr. Süleyman DÜNDAR
Afyon Kocatepe Üniversitesi Fen Edebiyat Fakültesi

Üye **:** Prof. Dr. Birdal ŞENOĞLU
Ankara Üniversitesi Fen Fakültesi

Üye **:** Yrd. Doç. Dr. Sinan SARAÇLI
Afyon Kocatepe Üniversitesi Fen Edebiyat Fakültesi

Afyon Kocatepe Üniversitesi
Fen Bilimleri Enstitüsü Yönetim Kurulu'nun
........./......../........ tarih ve
………………. sayılı kararıyla onaylanmıştır.

……………………………….
Enstitü Müdürü
Prof. Dr. Mevlüt DOĞAN

BİLİMSEL ETİK BİLDİRİM SAYFASI

Afyon Kocatepe Üniversitesi

Fen Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmasında;

- Tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- Görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- Başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- Atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- Kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- Ve bu tezin herhangi bir bölümünü bu üniversite veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

26/06/2012

Cengiz GAZELOĞLU

ÖZET

Yüksek Lisans Tezi

DOĞRUSAL TİP II REGRESYON TEKNİKLERİNİN MONTE-CARLO BENZETİM ÇALIŞMASI İLE KARŞILAŞTIRILMASI: SAĞLAM, BULANIK VE SAĞLAM BULANIK TEKNİKLER

Cengiz GAZELOĞLU

Afyon Kocatepe Üniversitesi

Fen Bilimleri Enstitüsü

İstatistik Anabilim Dalı

Danışman: Yrd. Doç. Dr. Sinan SARAÇLI

Bu çalışmanın amacı Basit Doğrusal Regresyon Çözümlemesi gerçekleştirilirken, bağımlı ve bağımsız değişkenin her ikisinin de ölçüm hatası içerdiği durumlarda, En Küçük Kareler (EKK) Açıortay, Bulanık (Fuzzy) EKK Açıortay, Sağlam (Robust) EKK Açıortay ve Bulanık Sağlam EKK Açıortay tekniklerinin karşılaştırmalı olarak Monte-Carlo benzetim çalışması ile incelenmesidir.

İlgilenilen veri setinin aykırı değer içerdiği ve içermediği durumlarda, farklı örneklem büyüklüklerinde (n=10,50 ve 100) ve farklı teorik dağılış biçimlerinde ($t_{(4)}$, $t_{(10)}$ ve $t_{(30)}$) ilgili regresyon teknikleri Hata Kareler Ortalaması (HKO) kriterine göre karşılaştırılmıştır. Çalışmanın bulgularına göre tüm dağılış biçimi ve tüm örneklem hacmindeki, aykırı değer içeren ve içermeyen veri setleri için Bulanık Huber EKK Açıortay tekniği en düşük HKO değerine sahip teknik olarak belirlenmiştir.

2012, xiii + 46 sayfa

Anahtar Kelimeler: EKK Açıortay, Bulanık EKK Açıortay, Sağlam EKK Açıortay, Sağlam Bulanık EKK Açıortay, Monte-Carlo benzetimi

ABSTRACT
M.Sc Thesis

COMPARISON OF LINEAR TYPE II REGRESSION TECNIQUES VIA MONTE-CARLO SIMULATION STUDY:
ROBUST, FUZZY AND ROBUST FUZZY TECHNIQUES

Cengiz GAZELOĞLU
Afyon Kocatepe University

Graduate School of Natural and Applied Sciences

Department of Statistics

Supervisor: Assistant Professor Sinan SARAÇLI

The aim of this study is comparatively examine the OLS (Ordinary Least Squares) Bisector, Fuzzy OLS Bisector, Robust OLS and Fuzzy Robust OLS techniques via Monte-Carlo Simulation study, when both the dependent and independent variables includes measurement errors in a simple linear regression analysis.

In conditions, whether the interested data sets includes or not any outliers, the performance of the regression techniques are examined for different sample sizes (n=10, 50 and 100) and different distribution types ($t_{(4)}$, $t_{(10)}$ and $t_{(30)}$) according to Mean Square Error (MSE) criteria. According to findings of the study, the Fuzzy Huber OLS Bisector technique has the lowest MSE in all different distribution types and for all sample sizes for the data sets that either includes or not any outliers.

2012, xiii + 46 pages

Key Words: OLS Bisector, Fuzzy OLS Bisector, Robust OLS Bisector, Monte-Carlo simulation

TEŞEKKÜR

Bu çalışmanın konusu, deneysel çalışmaların yönlendirilmesi, sonuçların değerlendirilmesi ve yazımı aşamasında yapmış olduğu maddi ve manevi büyük katkılarından dolayı tez danışmanım Sayın Yrd. Doç. Dr. Sinan SARAÇLI, araştırma ve yazım süresince yardımlarını esirgemeyen Sayın Prof. Dr. Birdal ŞENOĞLU, Sayın Prof. Dr. Emine SOYTÜRK, yüksek lisans eğitim sürecindeki destek ve katkılarından dolayı İstatistik Anabilim Dalı Başkanı Sayın Doç. Dr. Süleyman DÜNDAR'a, her konuda öneri ve eleştirileriyle yardımlarını gördüğüm Sayın Prof. Dr. İsmet DOĞAN'a, Sayın Yrd. Doç Dr. İbrahim KILIÇ'a ve, arkadaşlarım Arş. Gör. İlkay DOĞAN, Huriye TELLİ ve Durmuş KARLIK'a teşekkür ederim.

Ayrıca her zaman desteğini arkamda gördüğüm, büyük bir özveri ve sabırla beni dinleyen, bunaldığım zamanlarda hep yanımda olan bir abi gibi gördüğüm Sayın Yrd. Doç. Dr. Sinan SARAÇLI'ya çok ama çok teşekkür ederim. İyi ki varsınız hocam.

Beni bugünlere getiren, maddi ve manevi desteklerini hiçbir zaman esirgemeyen, üzerimde çok büyük emekleri olan çok kıymetli aileme teşekkür ederim.

Cengiz GAZELOĞLU
AFYONKARAHİSAR, 2012

İÇİNDEKİLER DİZİNİ

SİMGELER ve KISALTMALAR DİZİNİ

Kısaltmalar	
EKK	En Küçük Kareler regresyon tekniği
EKK(Y/X)	Y'nin bağımlı X'in bağımsız olduğu En Küçük Kareler regresyon tekniği
EKK(X/Y)	X'in bağımlı Y'nin bağımsız olduğu En Küçük Kareler regresyon Tekniği
HKO	Hata Kareler Ortalaması
EKKAO	En Küçük Kareler Açıortay regresyon tekniği
Bulanık EKK(X/Y)	X'in bağımlı Y'nin bağımsız olduğu Bulanık En Küçük Kareler regresyon tekniği
Bulanık EKK(Y/X)	Y'in bağımlı X'nin bağımsız olduğu bulanık En Küçük Kareler regresyon tekniği
Bulanık EKK Açıortay	Bulanık En Küçük Kareler Açıortay regresyon tekniği
Huber(X/Y)	X'in bağımlı Y'nin bağımız olduğu Huber regresyon tekniği
Huber(Y/X)	Y'in bağımlı X'nin bağımız olduğu Huber regresyon tekniği
Huber Açıortay	Sağlam Açıortay regresyon tekniği
Bulanık Huber(X/Y)	X'in bağımlı Y'nin bağımsız olduğu Bulanık Huber regresyon tekniği
Bulanık Huber(Y/X)	Y'in bağımlı X'nin bağımsız olduğu Bulanık Huber regresyon tekniği
BulanıkHuber Açıortay	Bulanık Huber Açıortay regresyon tekniği
$t_{(4)}$	4 serbestlik derecesi ile Student dağılmış veri seti
$t_{(4)}$-A	4 serbestlik derecesi ile Student dağılmış ve aykırı değer içeren veri seti
$t_{(10)}$	10 serbestlik derecesi ile Student dağılmış veri seti

$t_{(10)}$-A	10 serbestlik derecesi ile Student dağılmış ve aykırı değer içeren veri seti
$t_{(30)}$	30 serbestlik derecesi ile Student dağılmış veri seti
$t_{(30)}$-A	30 serbestlik derecesi ile Student dağılmış ve aykırı değer içeren veri seti
R.E.	Göreli Etkinlik (Relative Efficiency)

Simgeler

β_0	Regresyon modelindeki sabit katsayı
β_1	Regresyon modelindeki eğim katsayısı
X_i	Değişkenin i. gözlem değeri
$\hat{\beta}_0$	Tahmin edilen regresyon modelindeki sabit katsayı
$\hat{\beta}_1$	Tahmin edilen regresyon modelindeki eğim katsayısı
e_i	i. hata terimi
$\hat{\beta}_{yx}$	Bağımlı değişkenin Y olduğu modeldeki eğim katsayısı
$\hat{\beta}_{xy}$	Bağımlı değişkenin X olduğu modeldeki eğim katsayısı
$\bar{X}$	X değişkeninin ortalaması
$\bar{Y}$	Y değişkeninin ortalaması
W	Sağlam regresyonda gözlem değerlerinin ağırlık matrisi
$\underline{\omega}_{i}$,	Bulanık değişkendeki sol yayılımı
$\overline{\omega}_i$	Bulanık değişkendeki sağ yayılımı
$\overline{m}$	Bulanık sayıdaki sağ yayılımı
$\underline{m}$	Bulanık sayıdaki sol yayılımı
$\Psi(z)$	Sağlam regresyonda gözlem değerlerinin ağırlık matrisi

ŞEKİLLER DİZİNİ

ÇİZELGELER DİZİNİ

Sayfa

1. GİRİŞ

Regresyon analizi değişkenler arasındaki ilişkinin modellenmesinde kullanılan istatistiksel yöntemlerden birisidir. Birçok istatistiksel analiz yönteminde olduğu gibi regresyon analizinin belirli bir veri setine uygulana bilmesi için belli başlı varsayımların sağlanması gerekir. Farklı varsayımlar için birçok farklı regresyon teknikleri mevcuttur. Uygulamada kullanılan regresyon teknikleri arasında en yaygın olanı En Küçük Kareler (EKK) tekniğidir.

EKK tekniğinin varsayımları, özellikle deneysel çalışmalar söz konusu olduğunda, sağlanamamakta ve amaca uygun olmamaktadır. Bu gibi durumlarda daha sağlıklı sonuçlar verebilecek alternatif regresyon tekniklerini kullanmak doğru olacaktır (Saraçlı 2008).

EKK tekniği bağımsız değişken(ler)in herhangi bir ölçüm hatası içermediğini, mevcut hatanın bağımlı değişkenden kaynaklandığını varsayımı altında gerçekleştirir. Ancak gerçek hayatta karşılaşılan birçok olayda bağımlı ve bağımsız değişkenin her ikisinin de ölçüm hatası içermesi söz konusu olabilir. Bu gibi durumlarda başvurulacak regresyon teknikleri Tip II regresyon teknikleri olarak adlandırılırlar. Tip I ve Tip II regresyon teknikleri arasındaki en önemli farklardan birisi; Tip I regresyon tekniklerinde hataların bağımlı değişkenden kaynaklandığı ve bağımsız değişkenlerin hata içermediği varsayılırken, Tip II regresyon tekniklerinde söz konusu bu hatalar hem bağımlı hem de bağımsız değişkenlerden kaynaklanabilir.

Bulanık regresyon analizi, klasik regresyon analizi için gerekli olan varsayımların sağlanamadığı durumlarda, alternatif bir yöntem olarak Hideo Tanaka tarafından ilk 1982 yılında geliştirilmiştir. Yöntemin en belirgin olarak sağladığı fayda; regresyon katsayılarını tek bir değer olarak değil, bir aralık olarak tahmin etmesidir. Tahmin edilen bulanık aralık, kullanım amacına göre daraltılabilir veya genişletilebilir. Araştırmacı çalışmanın amacına göre, sonsuz tane aralık tahmini yapabilme imkanına sahiptir (Yücel 2005).

Bulanık EKK Açıortay tekniğinde ise veri setindeki X ve Y değişkenleri bulanık mantık çerçevesinde sırasıyla ayrı ayrı bağımsız değişkenler gibi düşünülerek iki ayrı regresyon doğrusu tahmin edilir. Elde edilen bu iki regresyon doğrusunun açıortay doğrusunun hesaplanmasıyla değişkenlerin her ikisindeki hataları da hesaba katan Bulanık EKK Açıortay doğrusu elde edilir.

Veri seti bulanıklaştırıldığında değişkenlerdeki söz konusu ölçüm hatalarının bu bulanıklık içerisine girdiğine dikkat edilmeli, yani değişkenlerin aldığı değerlerin kesin küme teorisinde olduğu gibi tek bir değere sahip olmadığına değişken değerleri bir aralık içerisinde yer alacağından söz konusu ölçüm hatalarının etkisini azalttığı göz önünde bulundurulmalıdır.

EKK tekniğinin iyi sonuç verebilmesi için veri setinde aykırı değer bulunmaması gerekir. İlgili veri seti aykırı değer içeriyorsa EKK tekniği ile tahmin edilecek regresyon doğrusu bu aykırı değerlerden etkilenecek dolayısıyla yanlış tahminde bulunulmasına sebep olacaktır. Veri seti aykırı değer içerdiğinde sağlam regresyon tekniklerinden yararlanmak daha doğru sonuçlar verecektir.

Sağlam sözcüğü istatistiksel varsayımlardan küçük sapmalara duyarsız istatistik işlemleri için kullanılmaktadır. Son yıllarda, en bilinen istatistiksel yöntemlerin varsayımlardan küçük sapmalara aşırı derecede duyarlı olduğu görülmüştür ve alternatif sağlam yöntemler önerilmiştir (Kavruk 2005).

Regresyon analizinde hatalar normal dağılış göstermiyorsa sağlam regresyon teknikleri önerilmektedir. Genellikle sağlam regresyon teknikleri, EKK tekniğinden daha fazla hesaplama gerektirmektedir (Draper ve Smith 1998, akt. Yıldırım 2010).

Bağımlı ve bağımsız değişkenlerin ikisinin de ölçüm hatası içerdiği ve veri setinde aykırı değer ya da değerler söz konusu olduğunda sağlam açıortay regresyon tekniğinden yararlanmak daha etkili sonuçlar elde edilmesine olanak sağlayacaktır. Bu teknikte robust temelde hesaplanan EKK(Y/X) ve EKK(X/Y) regresyon doğrularının açıortayından elde edilen doğru sağlam Tip II regresyon doğrusu olarak belirlenir.

Gözlem değerlerinin kesin bir şekilde açıklanamadığı durumlarda bulanıklık söz konusu olmaktadır. Bir kişi uzun, kısa, yaşlı, genç gibi ifadelerle tanımlanıyorsa, bir belirsizlik söz konusudur. Örneğin 20 yaşındaki bir birey farklı durumlar için yaşlı ya da genç kategorisine girebilir. Bu gibi bulanıklığın yanı sıra aykırı değer içeren bir veri seti söz konusu olduğunda Bulanık Sağlam teknikler yardımı ile tahmin edilecek regresyon doğrusu daha sağlıklı sonuçlar verecektir.

Tahmin problemlerinde verideki aykırı değerin etkisini azaltma özelliğine sahip yöntemler çok önemli ve yaygındır. Bu yöntemlerden en çok kullanılanları sağlam yöntemler olarak bilinir. Sağlam yöntemlerde farklı özelliklere sahip tahmin ediciler üzerine oldukça çok çalışma vardır. Bu çalışmalar Huber tahmin edicisiyle yaygınlaşmaya başlamış ve daha sonra hızla artış göstermiştir. Son yıllarda yapılan çalışmalarda ise sağlam yöntemlerde bulanık teori uygulamaları artmıştır. Olabilirlik teorisindeki olabilirlik dağılımları ya da bulanık küme teorisindeki üyelik fonksiyonları kavramı sağlam istatistiklerdeki ağırlık fonksiyonu kavramıyla anlatılabilir ve pek çok ortak noktaya sahiplerdir. Sağlam istatistikler ve bulanık küme teorisi son 20 yılda bağımsız olarak yavaş yavaş gelişen iki disiplindir. Tanaka vd. (1982) regresyon çözümlemesinde bulanık teoriyi kullanmışlardır. Sakawa ve Yano (1992) etkileşimli, çok amaçlı programlama problemini formüle etmişlerdir. Watada ve Yabuuchi (1994) bulanık regresyon ve sağlamlığı birlikte ele alan regresyon modelini önermişlerdir (Apaydın ve Şanlı 2004).

Her iki değişkenin de hata içerdiği göz önünde bulundurulursa ayrı ayrı bağımlı değişken olarak ele alınacak bu değişkenler için hesaplanacak bulanık sağlam regresyon doğrularının açıortayının alınması ile yani sağlam bulanık açıortay regresyon doğrusu ile yapılacak tahminler daha sağlıklı sonuçlar elde etmeyi sağlayacaktır.

Bu çalışmada ilk olarak yukarıda bahsedilen Tip I regresyon tekniklerinden EKK(X/Y), EKK(Y/X), Bulanık, Sağlam, Bulanık Sağlam ve Tip II regresyon tekniklerinden ise EKK Açıortay, Bulanık EKK Açıortay, Sağlam Bulanık EKK Açıortay teknikleri hakkında genel bilgiler verilmiş, uygulama bölümde ise bu tekniklerin performansları Monte- Carlo benzetim çalışması ile karşılaştırılmıştır. Performans kriteri olarak Hata Kareler Ortalaması (HKO) ve Relative Efficiency (R.E) değerleri dikkate alınmıştır. Son bölümde ise benzetim çalışması sonucunda elde edilen bulgular değerlendirilerek kullanılan tekniklere ilişkin çıkarsamalarda bulunulmuştur.

2. LİTERATÜR BİLGİLERİ

Regresyon analizinin ilk kullanım alanı astronomi olmuştur. Lengendre ve Gauss gezegenlerin yörüngelerini belirlemek amacıyla EKK olarak bilinen tekniğini oluşturmuşlardır. Konu ile ilgili değişkenlerden yola çıkarak bu değişkenler için bir regresyon modeli geliştirmişlerdir (Ergül 2006).

Regresyon (bağlanım), sözlük anlamı ile bir şeyi başka bir şeye bağlama işi ve biçimidir. Bilimsel olarak regresyon terimi, bir değişken ile başka bir ya da birden çok değişken arasında ilişki kurma işini ve ilişkinin biçimini anlatır (Şıklar 2000).

İki değişken arasında doğrusal bir ilişki olduğunda, bu ilişki dağılım grafiğindeki noktalar arasından geçen uygun bir doğru ile tanımlanabilir. Bu doğruya regresyon doğrusu denir ve matematiksel olarak bir denklem ile gösterilebilir. Bu denkleme de regresyon denklemi denir (Sümbüloğlu 2002).

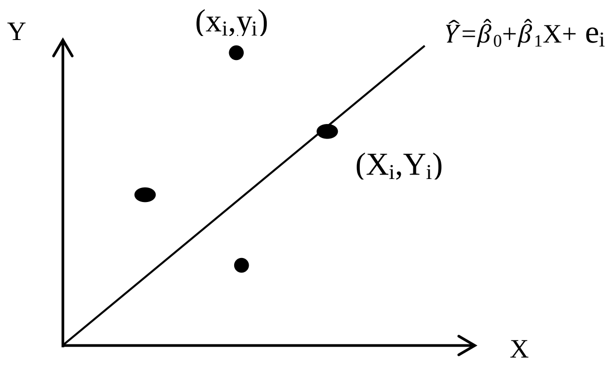
Regresyon analizi, bağımlı ve bağımsız değişkenlere ait veri kümesinden yararlanarak bağımlı değişkenin ortalama değerini tahmin etmek, modelde bulunan bağımsız değişkenlerin bağımlı değişken üzerinde beklenen etkiye sahip olup olmadığını test etmek için uygulanmaktadır. Bununla birlikte, bağımsız değişken(ler)in alabileceği değerlere karşılık bağımlı değişken değeri için öngörülerde bulunmak amacıyla da kullanılmaktadır (Çankal 2010).

İstatistiksel bir ilişkiden bahsedilebilmesi için Y bağımlı değişkeninin X bağımsız değişkeniyle sistematik olarak değişme eğiliminde olması yani X ve Y değişkenleri arasında ilişkinin doğrusal ya da eğrisel bir fonksiyon ile ifade edilebilmesi ve X bağımsız değişkeninin Y bağımlı değişkenini etkilenmesi gerekmektedir (Çankal 2010).

Basit doğrusal regresyon denklemi Eşitlik 2.1'de verildiği gibidir.

$$Y_i = \beta_0 + \beta_1 X_i + e_i \qquad (2.1)$$

Bu modelde yer alan β_0 sabit katsayısı, β_1 ise eğim katsayısını ifade etmektedir. X_i değerleri, herhangi bir ölçüm hatası içermeyen stokastik olmayan gözlem değerleridir. Modelde yer alan e_i hata terimleri ise stokastik, bağımsız değerlerdir (Tiku and Akkaya 2004).



Şekil 2.1. Regresyon denklemine ilişkin genel gösterim (Saraçlı 2008)

2.1 Tip I Regresyon Teknikleri

Sağlık, fen ve sosyal bilimlerdeki hemen hemen tüm disiplinlerde kullanıma sahip olan regresyon analizi, bağımlı değişken ile bağımsız değişken(ler) arasındaki ilişkiyi modellemede sıklıkla tercih edilen önemli bir istatistiksel yöntemdir. EKK tekniği, regresyon analizi içerisinde yaygın olarak kullanılan bir teknik olup, bu tekniğin kullanılabilmesi için bazı varsayımların gerçekleşmesi gerekmektedir. Bu varsayımlardan birisi, bağımsız değişkenlerin herhangi bir ölçüm hatası içermediğidir. Bu doğrultuda, bağımsız değişken(ler)in ölçüm hatası içermediği varsayılan tekniklere Tip I regresyon teknikleri adı verilmektedir. Söz konusu varsayımların sağlanamadığı durumlarda bağımlı ve bağımsız değişken(ler) arasındaki ilişkiyi modellemek için alternatif teknikler geliştirilmektedir (Gazeloğlu ve Saraçlı 2011).

2.1.1 En Küçük Kareler (Y|X) Regresyon Tekniği

En Küçük Kareler tekniği ilk olarak 1795 yılında Alman matematikçi Carl Friedrich Gauss tarafından ileri sürülmüştür. Gauss 1801 yılında En Küçük Kareler tekniğini kullanarak, keşfinden kısa süre sonra kaybedilen Ceres asteroidinin tekrar gözlemlenebileceği pozisyonu hesaplayabilmiştir. Bu başarısıyla büyük üne kavuşan Gauss bu tekniği ilk olarak 1809'da yayımlamıştır (Paris 2012).

$Y_i = \beta_0 + \beta_1 X_i + e_i$ doğrusal ilişkisi, X ve Y değişkenlerinin anakütleleri için geçerlidir. Ancak anakütleye ulaşmak her zaman mümkün olamayacağından anakütle içerisinden seçilen örneklem verilerinden β_0 ve β_1 parametreleri EKK tekniği ile tahmin edilebilir. Örneklem verileri yardımıyla tahmin edilen regresyon modeli;

$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + e_i$ şeklindedir.

En Küçük Kareler tekniğinde aşağıdaki varsayımlar sağlanmalıdır (Şıklar 2000).

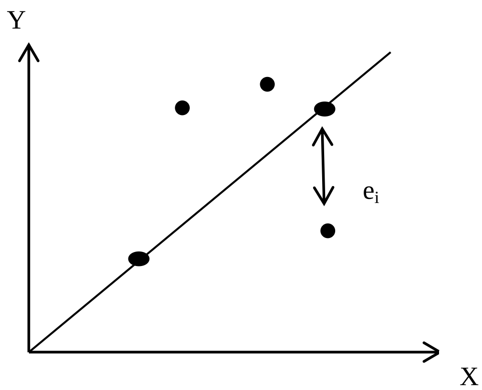
1) Hata terimi e_i ortalaması sıfıra eşit rassal bir değişkendir.
2) Hata terimi e_i' nin dağılımı normaldir.
3) Hata terimi e_i' nin değerleri arasında ilişki yoktur.
4) Hata terimi e_i' nin varyansı her X_i değeri için eşittir.
5) X değişkeni hata terimi e ile ilişkili olmayıp, stokastik değildir.
6) X değişkeni tekrar eden örnek değerlerine göre sabittir.
7) Model belirleme hatası taşımamaktadır.
8) Bağımsız değişkenler arasında tam veya kuvvetli bir ilişki yoktur. (Çoklu doğrusal regresyon modellerinde geçerlidir).

Regresyon analizinde bu varsayımların geçerli olduğu model "Tip I doğrusal regresyon modeli" olarak adlandırılır. Bu modellerde regresyon katsayıları EKK tekniği ile en iyi şekilde tahmin edilir (Maroco 2007).

Bağımlı ve bağımsız değişkenlerin hataya maruz kaldığı ve normal dağıldığı doğrusal regresyon modelleri "Tip II doğrusal regresyon" olarak ya da hata içeren değişkenler modelleri olarak adlandırılırlar (Maroco 2007).

EKK tahmini, diğer varsayımların dışında, tahmin hatalarının X'in tüm düzeyleri için sabit hata varsayımı ile normal dağıldığını varsayar. (($\varepsilon \sim N(0,\sigma^2)$) normalliği varsayımı, pratikte nadir olarak geçerlidir. Bu varsayım ihlali, genellikle gözlem değerleri arasında aykırı değerlerin olmasından kaynaklanır. Bu aykırı değer için veri seti, uzun kuyruğa sahip normal olmayan ya da varsayılan ε'den daha yüksek varyansa sahip bir normal dağılım sergileyebilir (Draper Smith 1981; Hamilton 1992 akt. Nevitt and Tam 1998).

Klasik regresyon tekniklerinden biri olan EKK tekniğinde amaç Şekil 2.2'de gösterildiği üzere gözlem değerlerinin tahmini regresyon doğrusuna olan dik uzaklıklarının karelerinin toplamını minimize etmektir.



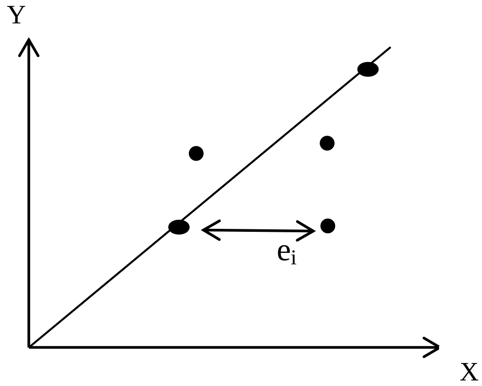
Şekil 2.2 EKK(Y/X) tekniğinde hatalara ilişkin grafik

En Küçük Kareler regresyon tekniğinde Şekil 2.2'de görülen hataların ($\mathbf{e_i}$) bağımlı değişken olan (Y)'den kaynaklandığını varsaymaktadır.

EKK tekniği kullanılarak bağımsız değişken (X) ile Bağımlı değişken (Y) ye ilişkin regresyon modeli literatür de en yaygın olarak kullanılan klasik bir tekniktir.

2.1.2 En Küçük Kareler (X|Y) Regresyon Tekniği

Klasik EKK tekniğinde kullanılan bağımlı ve bağımsız değişkenden oluşan bir model ve bu değişkenlerin birbirleriyle yer değiştirmesi sonucu ikinci bir modelin elde edilmesi yöntemine ters regresyon (EKK(X/Y)) tekniği denir (Schaefer and Visser 2003).



Şekil 2.3 Ters regresyon yönteminin grafiksel gösterimi

Ters regresyon tekniğinde amaç; gözlem değerlerinin tahmin edilen regresyon doğrusuna olan yatay uzaklıklarının karelerinin toplamını minimize etmektir.

Isobe et all., (1990) yaptıkları çalışmada EKK (Y|X) tekniği için elde edilen eğim katsayısı $\hat{\beta}_{1yx}$, Eşitlik 2.2'de görüldüğü gibi ve EKK (X|Y) tekniği için elde edilen eğim katsayısı $\hat{\beta}_{1xy}$, Eşitlik 2.3'de görüldüğü gibi hesaplanmış ve diğer teknikler için elde edilen eğim katsayıları bu iki eğim katsayısı cinsinden hesaplanmıştır. Eşitlik 2.4 ve Eşitlik 2.5'de görülen bu eğim katsayılarına ilişkin varyans hesaplamaları verilmiştir. (Billingsley 1986, akt. Saraçlı 2008).

$$\hat{\beta}_{1yx} = \frac{S_{xy}}{S_{xx}} \tag{2.2}$$

$$\hat{\beta}_{1xy} = \frac{S_{yy}}{S_{xy}} \tag{2.3}$$

$$\mathrm{Var}\hat{\beta}1\mathrm{yx}=\frac{1}{S_{xx}^2}[\sum_{i=1}^{n}(x_i-\overline{x})^2-\left(y_i-\hat{\beta}_{1yx}x_i-\overline{y}+\hat{\beta}_{1yx}\overline{x}\right)^2] \qquad (2.4)$$

$$\mathrm{Var}\hat{\beta}_{1xy}=\frac{1}{S_{xy}^2}[\sum_{i=1}^{n}(y_i-\overline{y})^2-\left(y_i-\hat{\beta}_{1xy}x_i-\overline{y}+\hat{\beta}_{1xy}\overline{x}\right)^2] \qquad (2.5)$$

Klasik regresyon denklemi için sabit katsayının ($\hat{\beta}_{0yx}$) hesabına ilişkin formül, Eşitlik 2.6'de verilmiştir.

$$\hat{\beta}_{0yx}=\overline{y}-\hat{\beta}_i\overline{x} \qquad (2.6)$$

2.1.3 Sağlam Regresyon Tekniği

Sağlam regresyon tekniklerinin geçmişi 19. yy' a kadar uzanmaktadır. Bu çalışmalardan ilki Simon Newcomb tarafından ele alınmıştır. Ancak en büyük atılım 1960 lı yıllarda ve 1970' li yılların başlarında John Tukey ve Peter Huber isimli bilim adamları tarafından yapılmıştır (Maronna at al. 2006).

Basit regresyonda olduğu gibi bilinen $\beta_0,\ldots,\beta_p$ parametreleri tahmin etmede, En Küçük Kareler tekniği varsayımlar açısından oldukça hassastır. Son yıllarda birçok istatistikçi varsayım bozulmasına karşı sağlam regresyonun önemine değinmiştir (Rousseeuw and Leroy 1986).

Bağımlı değişken Y normal dağılımdan geldiğinde EKK tekniği regresyon parametreleri için iyi tahminler vermektedir. Ancak gerçek hayatta normal dağılmış bir veri setini bulmak oldukça güçtür. Eğer gözlemler normal dağılmıyorsa ya da aykırı değer içeriyorsa EKK tekniği uygun olmamaktadır. Bu durumda sağlam regresyon tekniklerinin kullanılması gerekir (Mutan 2004).

Son yıllarda birçok istatistikçi, yaygın şekilde kullanılan bazı istatistiksel yöntemlerin varsayımlarından önemsiz küçük sapmalardan bile çok etkilendiğinin ve bu sapmalara karşı duyarlı olduğunun giderek artan bir oranda farkına varmıştır. Bu nedenle, varsayımlardan sapmalara daha az duyarlı olan “dayanıklı” alternatif yöntemler önerilmiştir. Bu sürecin sonucu olarak, birçok dayanıklı regresyon tekniği ortaya çıkmıştır. Bu teknikler, belli ortak özellikleri açısından sınıflanarak isimlendirilmektedir. Dayanıklı regresyon tahmin edicilerinin üç temel tipi L, M ve R tahmin edicileridir. Ayrıca, çeşitli dayanıklı tahmin ediciler de yüksek bozulma sınırına sahip olmaları nedeniyle ayrı bir sınıf olarak ele alınmaktadır (Türkay 2004).

Gözlem sonucu elde edilen veri setinde diğer gözlemlerden oldukça küçük ya da oldukça büyük olduğu görülen gözlem değerleri aykırı değerler olarak adlandırılırlar. Aykırı değerler, verilerin özelliklerini yansıtan gerçek değerler olabilecekleri gibi ölçme, kayıt altına alma, aktarma hatalarından dolayı gerçek değer yerine başka bir değerin alınmasından da kaynaklanabilirler (Vural 2007).

Veri setinde aykırı değer söz konusu olduğunda, sağlam regresyon tekniğini kullanarak tahmin edilen regresyon doğrusu gerçek değerlere daha yakın tahminler yapmamıza olanak sağlayacaktır. Eşitlik 2.7’da $\hat{\beta}$ vektörü sağlam bir şekilde tahmin edilmektedir.

$$\hat{\beta} = (X'WX)^{-1}X'WY \tag{2.7}$$

Sağlam regresyon tekniklerinden yaygın olarak kullanılan bir M regresyon tekniği olan Huber yönteminin hesaplanışı Eşitlikler 2.8 - 2.13’de verilmiştir.

t=2 için

Değişim aralığı $|z| \leq t$ ise

$$W(z)=1 \tag{2.8}$$

$$\Psi(z)=P'(z)=z \tag{2.9}$$

$$P(z)=\frac{1}{2}z^2 \tag{2.10}$$

$|z|>t$ ise

$$W(z)=\frac{t}{|z|} \tag{2.11}$$

$$\Psi(z)=P'(z)= t\mathrm{sig}(z) \tag{2.12}$$

$$P(z)= |z|t - \frac{1}{2}t^2 \tag{2.13}$$

Yukarıdaki eşitliklerden veri setindeki tüm gözlem değerleri için W(z) ağırlık değerleri hesaplanarak veri setindeki aykırı değerlerin diğer gözlem değerlerine göre düşük ya da yüksek bir ağırlığa sahip olacak şekilde etkisi azaltılır.

2.1.4 Bulanık Regresyon Tekniği

Bulanık mantık, proses kontrolü, uzman sistemler, iz tanıma, otomatik kontrol, karar verme, üretim planlama, görüntü tanıma, bilgi sistemleri gibi birçok alanda kullanılmaktadır. Uzakdoğu da özellikle Japonya'da oldukça fazla ilgi gören bulanık mantık yaklaşımının bugüne kadar en fazla uygulandığı alan ise denetim sistemleridir (Çevik ve Yıldırım 2010).

Bulanık mantık, belirsiz ve karmaşık sistemlerin çözümüne, basit çözümler getirmesi açısından avantajlıdır. Basit bir modelle karmaşık problemler kolaylıkla çözülerek daha az maliyetle daha hızlı sonuçlar elde edilmektedir. Ancak üyelik fonksiyonlarının belirlenmesi açısından daha fazla uzmanlık isteyen ve zamana bağlı bir model olması da dezavantajı olarak ifade edilebilir (Çevik ve Yıldırım 2010).

Bulanık regresyon analizi, bulanık bir çevrede bağımlı ve bağımsız değişkenler arasındaki fonksiyonun değerlendirilmesinde kullanılan klasik regresyon analizinin bulanık bir türüdür (Nasrabadi ve Nasrabadi 2004).

Verilerin her zaman kesin sayılarla ifade edilemediği durumlarda, özellikle veri için niteliksel tanımlamalar getirildiğinde bulanık sayı ifadesi ortaya çıkmakta ve bulanık küme teorisine ihtiyaç duyulmaktadır (Tekşen 2008).

Günlük hayatta kullandığımız birçok terim genellikle bulanık (net olmayan) bir anlam içerir. Bir durumu tanımlarken, bir olayı açıklarken, komut verirken ve daha birçok durumda kullandığımız sözel veya sayısal ifadeler bulanıklık içerir. Bu durumlara: yaşlı, genç, uzun, kısa, sıcak, soğuk, ılık, bulutlu, parçalı bulutlu, güneşli, hızlı, yavaş, çok, az, biraz, fazla, çok az, çok fazla, gibi daha pek çok sözel ifade örnek gösterilebilir.

İnsanlar bir olayı anlatıp, bir durum karşısında karar verirken bu tür kesinlik ifade etmeyen terimler kullanırlar. Kişinin yaş durumuna göre; yaşlı, orta yaşlı, genç, çok yaşlı ve çok genç denir. Yolun kayganlık ve rampa durumuna göre arabanın gaz veya fren pedalına biraz daha yavaş veya biraz daha hızlı basarız. Çalıştığımız odanın ışığı yetersiz ise onu biraz artırır, yeterinden fazla ise biraz azaltırız. Bütün bunlar insan beyninin belirsiz ve kesinlik içermeyen durumlarda nasıl davrandığına ve olayları nasıl değerlendirip, tanımlayıp, komut verdiğine dair birer örnektir. Bulanık mantığın ve bu mantık kurallarını kullanan bulanık küme teorisinin Lotfi A. Zadeh tarafından geliştirilip 1965 tarihli makalesinde yayınlanmasından sonra belirsizlik içeren sistemlerin incelenmesi yeni bir boyut kazanmıştır (Altaş 1999).

Açık veya net olmayan ya da bulanık olan verilerle pazarlama, ekonomi, kalite kontrol, görüntü tanımlama, yapay zeka gibi bir çok alanda karşılaşırız. Bu kesin olarak tanımlanmayan veya sınıflanmayan veriler için bulanık sayılar veya veriler tanımını kullanılmaktadır. Değişkenin aldığı sayısal değer sübjektif bir tanımlamaya yol açıyorsa bu veri bulanık veri olarak tanımlanır (Semiz ve Genç 2003).

Klasik matematiksel yöntemlerle karmaşık sistemleri modellemek ve kontrol etmek zordur, çünkü verilerin tam olması gerekir. Bulanık mantık kişiyi bu zorunluluktan kurtarır ve daha niteliksel bir tanımlama olanağı sağlar (Yanartaş 2009).

Klasik Regresyon, bağımsız nicel değişkenleri kullanarak, bağımlı nicel değişken değerlerini tahmin etme yöntemidir. Klasik regresyon mevcut ve/veya geçmiş dönemlere ait nicel verilere dayanarak tahmin yapma, karar destek sistemlerine katkılarda bulunma gibi önemli uygulama aracıdır. Klasik regresyondaki bazı önemli varsayımlara rağmen, gerçek hayatta var olan esnek düşünme yapısı metot içinde yer almamaktadır (Tanaka 1982).

Klasik Regresyon mevcut durumu veya geleceği tahmin etme amacı ile eldeki mevcut verilere dayanarak değişkenler arasındaki ilişkiyi modelleyen istatistiksel bir araç olarak kullanılmakta ve söz edilen değişkenler arasındaki ilişkiyi çok keskin tanımlamaktadır. Bilim adamları geleceği tahmin etmek amacıyla kullandıkları bulanık olmayan regresyonda bağımsız değişkenlerin etkilerini ve miktarını tam olarak yakalamaya çalışırlar. Hâlbuki doğadaki birçok belirsiz (bulanık) olayı yansıtmaya çalışan kesin matematiksel modeller her zaman çalışmaz. Klasik regresyon ancak doğada kesin bir şekilde veri elde edebildiğinde doğru sonuçlar vermektedir. Diğer bir deyişle, kendisine klasik mantığı temel almış regresyon, insan düşünüş tarzına yakın özellikler taşıyan bulanık sistemlerde yanlış kararlara sebep olmaktadır (Tanaka 1982).

L. Asker Zadeh ilk kez "bulanık küme" kavramından 1962'de yayınlandığı "Devre teorisinden sistem teorisine" adlı çalışmasında bahsetmiştir. Belirsizliğin matematiği olarak isimlendirilebilecek bu düşünceler ikinci çalışmada olgunlaşmıştır (Zadeh 1965).

Zadeh, Kaufman'ın kitabı için yazdığı sunuş yazısında çok iddialı olarak bulanık küme teorisinin psikoloji, sosyoloji, sosyal bilimler, felsefe, ekonomi, dil bilimi, yöneylem araştırması, yönetim bilimi ve diğer alanlarda yapay zekâ sistemlerinin dizayınında temel olacağını belirtmiştir (Kaufman 1975).

1960'lı yılarda Zadeh'in geliştirdiği Bulanık Küme Teorisi yaklaşımı, nitel değişkenlere sayısal olarak değerlendirme olanağı yaratmıştır. Daha sonra Tanaka ve ark. (1982) Bulanık Küme Teorisinden yararlanarak Bulanık Doğrusal Regresyon yöntemini geliştirmişlerdir. Bu yöntemle her bir nitel gözlem üyelik derecesine göre modele sayısal olarak katılmaktadır (Düzyurt 2008).

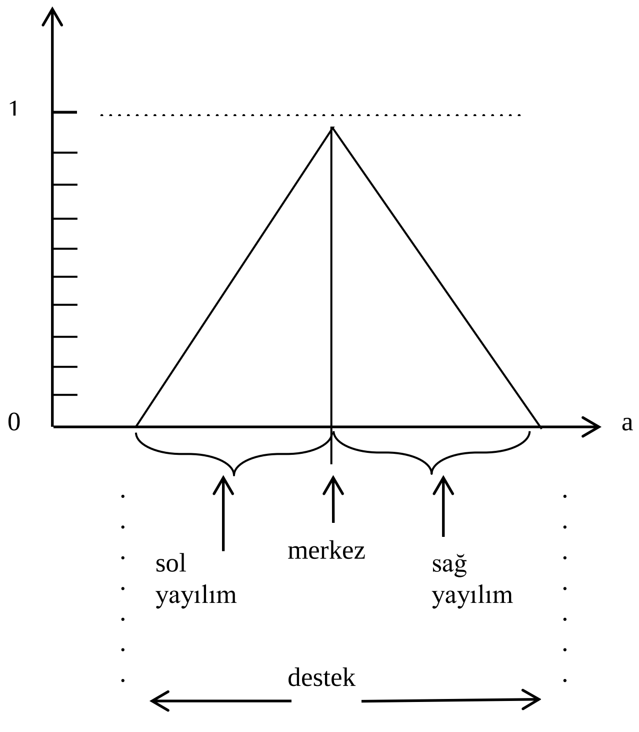
Diamond, 1988 yılında, bulanık parametreleri tahmin etmek için bulanık EKK yöntemini ilk kullanan kişi olmuştur (Shapiro 2005).

Bulanık regresyon, karar verme sürecinde hem nicel hem de nitel değişkenlerin dikkate alınmasına olanak sağlayan problemlerin çözümünde kullanılabilecek karar verme yöntemlerinden birisidir. Bulanık regresyon, sistem yapısındaki belirsizliğe bağlı olarak verilerin tamamının ya da bir kısmının bulanık olması veya sistem yapısının değişkenler arasında kesin ilişkiler tanımlanmasına imkân vermemesi gibi klasik regresyon uygulanmasının önerilmediği durumlarda kullanılan alternatif bir yöntemdir (Düzyurt 2008).

Osaka ve Ryukoku Üniversitelerinden Tanaka, Hayashi ve Watada'nın esnek düşünmek tabanına dayalı çalışmaları regresyon analizine bulanık boyut kazandırmıştır. Bu sayede doğada ve günlük yaşantımızda belirsizliğin var olduğu ve klasik metodun yetersiz kaldığı durumlarda kurulan doğrusal sistemlerin güvenilirliğini arttırdığı gibi, doğal düşünüşe uygun karar vermeyi sağlayabilmektedir (Güneş 2001).

Klasik regresyon modelinde hesaplanan değerler ile gözlenen değerler arasında ki farklılıkların ölçüm hataları olduğu kabul edilir. Bulanık regresyon modelinde ise bu farklılıkların sistem parametrelerinin kararlı olmayışına dayandığı varsayılır (Güneş 2001).

Bulanık bileşenleri üçgensel bulanık sayılar olarak varsayılır. Bu nedenle katsayılar aşağıdaki Şekil 2.4'de görüldüğü gibi bir üyelik fonksiyonu ile ifade edilebilir.



Şekil 2.4 Bulanık katsayı üyelik derecesi gösterimi (Shapiro 2005)

X=(m,$\underline{m}$,$\overline{m}$) şeklinde tanımlanan üçgensel bulanık sayıda m x'in merkezi, $\underline{m}$ sol yayılma ve $\overline{m}$ sağ yayılma olarak tanımlanır. Bulanık En Küçük Kareler yaklaşımı için X_i=(x_i, $\underline{\delta}_i$, $\bar{\delta}_i$), $Y_i = (y_i, \underline{\omega}_i, \overline{\omega}_i)$ üçgensel bulanık sayıları ele alındığında $Y = \alpha + \beta X$ modeli düşünülür. Burada α,β kesin sayılardır. Parametrelerin kesin olduğu model ele alındığında en küçük kareler optimizasyon problemi

Minimum $\quad r(\alpha,\beta) = \sum d(\alpha + \beta X_i, Y_i)^2$ (2.14)

olarak tanımlanır. Bu durumda

$$\boldsymbol{d(\alpha + \beta X_i, Y_i) = [\alpha + \beta x_i - y_i - (\beta \underline{\delta}_i - \underline{\omega})]^2 + [\alpha + \beta x_i - y_i + (\beta \bar{\delta}_i - \bar{\omega}_i)]^2 + (\alpha + \beta x_i - y_i)^2} \quad (2.15)$$

şeklinde tanımlanır (Diamond 1988 akt., Apaydın ve Şanlı 2004).

x'in kesin sayı ve Y=(y, $\underline{\omega}$, $\overline{\omega}$) üçgensel bulanık sayı olması durumda bulanık regresyon modeli

$$Y = A + BX \quad (2.16)$$

Olarak tanımlanır. Burada A=(α, $\underline{\alpha}$, $\bar{\alpha}$) ve B=(b,$\underline{\beta}$, $\bar{\beta}$) bulanık parametrelerdir. Eşitlik 2.16 'deki parametre tahminlerinin yapıla bilmesi için en küçük kareler optimizasyon problemi

Minumum $r(A,B)=\sum d(A + x_i B, Y_i)^2$ (2.17)

biçimindedir. Burada;

$$d(A+x_iB,Y_i)=(a+bx_i-y_i)^2+(a+bx_i-\underline{\alpha}-\underline{\beta}x_i-y_i+\underline{\omega}_i)^2+(a+bx_i+\bar{\alpha}+x_i\bar{\beta}-y_i-\bar{\omega}_i)^2 \quad (2.18)$$

olarak tanımlanır. $\frac{\partial d}{\partial a} = 0$ ve $\frac{\partial d}{\partial \beta} = 0$ çözümlemesi ile a ve b parametreleri ve $\frac{\partial d}{\partial \alpha} = 0$ ve $\frac{\partial d}{\partial \beta}$=0 çözümlemesi ile α veβ parametreleri elde edilir (Diamond 1988, Xu and Li 2001, akt. Apaydın ve Şanlı 2004).

Gözlem değerlerini bulanıklaştırmada önerilen yöntem için bağımsız değişken değerleri; merkez(x_i), sol yayılma $\underline{\delta}_i$=x_i/7, sağ yayılma $\bar{\delta}_i$=x_i/6 ve bağımlı değişken değerleri; merkez(y_i), sol yayılma $\underline{\omega}_i$=y_i/8, sağ yayılma $\bar{\omega}_i$=y_i/7 olarak çözümlenebilir (Apaydın ve Şanlı 2004).

Bulanık regresyon katsayıları;

$$\hat{\beta} = (X'X + A' + B'B)^{-1}(X'Y + A'C + B'D) \tag{2.19}$$

olarak bulunur. Burada;

$$X=\begin{bmatrix} 1 & x_{11} & . & \cdot & x_{p1} \\ 1 & x_{12} & . & \cdot & x_{p2} \\ . & . & . & \cdot & \cdot \\ . & . & . & \cdot & \cdot \\ 1 & x_{1n} & . & \cdot & x_{pn} \end{bmatrix} \tag{2.20}$$

$$Y=\begin{bmatrix} y_1 \\ \vdots \\ \vdots \\ y_n \end{bmatrix} \tag{2.21}$$

$$C=\begin{bmatrix} y_1 - \underline{\omega}1 \\ \vdots \\ y_n - \underline{\omega}\mathrm{n} \end{bmatrix} \tag{2.22}$$

$$D=\begin{bmatrix} y_1 + \bar{\omega}1 \\ \vdots \\ y_n + \bar{\omega}\mathrm{n} \end{bmatrix} \tag{2.23}$$

$$B=\begin{bmatrix} 1 & (x_{11}+\bar{\delta}_{11}) & . & . & (x_{p1}+\bar{\delta}_{p1}) \\ 1 & (x_{12}+\bar{\delta}_{12} & . & . & (x_{p2}+\bar{\delta}_{p2}) \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & (x_{1n}+\bar{\delta}_{1n} & . & . & (x_{pn}+\bar{\delta}_{pn}) \end{bmatrix} \tag{2.24}$$

$$A=\begin{bmatrix} 1 & (x_{11}-\underline{\delta}_{11}) & . & . & (x_{p1}-\underline{\delta}_{p1}) \\ 1 & (x_{12}-\underline{\delta}_{12} & . & . & (x_{p2}-\underline{\delta}_{p2}) \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & (x_{1n}-\underline{\delta}_{1n} & . & . & (x_{pn}-\underline{\delta}_{pn}) \end{bmatrix} \tag{2.25}$$

Gözlem sayısı i=1,2,…,n ve değişken sayısı j=1,2,…,p'dir.

2.2 Tip II Regresyon Teknikleri

Literatürde birçok farklı Tip II regresyon tekniği vardır. (bkz. Saraçlı ve diğerleri (2009)). Genel olarak gözlem değerlerinin elde edilen regresyon denklemine dik ya da hata miktarına bağlı olarak hesaplanan uzaklıklarının alınması sonucunda her iki değişkendeki hataları da dikkate alma mantığına dayanan bu teknikler, Ortogonal Regresyon, Deming Regresyon, York Regresyon teknikleri ve bunların çeşitli koşullar altında türetilmiş halleridir. Regresyon parametrelerini tahmin etmedeki hesaplanışları bakımından Ortogonal Regresyon Tekniği; Majör Eksen ve İndirgenmiş Majör Eksen olmak üzere ikiye, Deming Regresyon Tekniği; Deming, Optimal Deming ve Ağırlıklandırılmış Deming olmak üzere üçe, York Regresyon Tekniği ise York ve Optimal York Regresyon Tekniği olmak üzere iki gruba ayrılmaktadır. Passing-Bablok Regresyon Tekniği ise EKK Tekniğine alternatif olan ve parametrik olmayan diğer bir regresyon tekniğidir (Saraçlı 2010)

Saraçlı (2008) çalışmasında yukarıdaki Tip II regresyon tekniklerinin performanslarını karşılaştırmış ve en iyi performansı EKK Açıortay regresyon tekniğinin verdiğini tespit etmiştir. Bundan dolayı bu çalışmada EKK Açıortay tekniğinden yararlanılmıştır.

2.2.1 En Küçük Kareler Açıortay Tekniği

Tip II regresyon tekniklerinden biri olan EKK Açıortay tekniği, gözlem noktalarının tahmin edilen regresyon doğrusuna olan uzaklığını, EKK(Y/X) doğrusu ile EKK(X/Y) regresyon doğrusunun açıortayını dikkate alarak minimize etmeye çalışır. Literatür taramalarında EKK-Açıortay doğrusunun eksikliği hakkında hiçbir çalışmaya rastlanmamıştır (Isobe et all. 1990 akt. Saraçlı 2008). Ayrıca Saraçlı ve Çelik (2010) yatıkları benzetim çalışmasında, bağımlı ve bağımsız değişkenlerin ikisinin de ölçüm hatası içermesi durumunda bu tekniğin farklı dağılış biçimlerinde, farklı örneklem hacimlerinde ve veri setinin aykırı değer içerip içermediği durumlarda diğer Tip II basit doğrusal regresyon tekniklerinden üstünlüğünü vurgulamışlardır.

Isobe at. all. (1990) Bu teknik yardımıyla hesaplanacak olan eğim katsayısı ve varyans hesabı sırası ile Eşitlik 2.26, Eşitlik 2.30'da verilmiştir. Varyans hesaplamasında kullanılacak olan kovaryans terimlerinin hesabı ise Eşitlik 2.33'de açık olarak verilmiştir. X ve Y değişkenlerinin ortalamaları ise eşitlik 2.35 ve 2.36'da yer almaktadır.

$$\hat{\beta}_{1AO}=(\hat{\beta}_{1xy}+\hat{\beta}_{1yx})^{-1}\left[\hat{\beta}_{1xy}\hat{\beta}_{1yx}-1+\sqrt{(1+\hat{\beta}^2_{1xy})(1+\hat{\beta}^2_{1yx})}\right] \quad (2.26)$$

$$A=\frac{\hat{\beta}^2_3}{(\hat{\beta}^2_{1xy}+\hat{\beta}^2_{1yx})^2(1+\hat{\beta}^2_{1xy})(1+\hat{\beta}^2_{1yx})} \quad (2.27)$$

$$B=[(1+\hat{\beta}^2_{1yx})^2\widehat{var}(\hat{\beta}_{1xy})+2(1+\hat{\beta}^2_{1xy})(1+\hat{\beta}^2_{1yx})] \quad (2.28)$$

$$C=\widehat{cov}(\hat{\beta}_{1xy},\hat{\beta}_{1yx})+(1+\hat{\beta}^2_{1xy})^2\widehat{var}(\hat{\beta}_{1yx}) \quad (2.29)$$

$$Var(\hat{\beta}_{1AO})=A*B*C \quad (2.30)$$

$$E=(\hat{\beta}_{1xy}S^2_{xx})^{-1} \quad (2.31)$$

$$D= \left\{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})\left[Y_i - \bar{Y} - \hat{\beta}_{1xy}(X_i - \bar{X})\right][Y_i - \bar{Y} - \hat{\beta}_{1yx}(X_i - \bar{X}]\right\} \qquad (2.32)$$

$$\widehat{cov}\left(\hat{\beta}_{1xy}, \hat{\beta}_{1yx}\right) = E*D \qquad (2.33)$$

$$S_{xx} = \sum_{i=1}^{n}(X_i - \bar{X})^2 \qquad (2.34)$$

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad (2.35)$$

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \qquad (2.36)$$

2.2.2 Sağlam Regresyon Açıortay Tekniği

Sağlam Açıortay regresyon tekniğinde aykırı değer içeren X ve Y değişkenleri için ayrı ayrı üyelik fonksiyonları hesaplanarak, X ve Y değişkenleri sırası ile bağımlı değişken gibi düşünülüp tahmin edilen regresyon doğrularının açıortay doğrusu tahmin edilmektedir. Hesaplanan bu açıortay doğrusu bağımlı ve bağımsız değişkenin her ikisinin de ölçüm hatası içerdiği durumlardaki veri setinin matematiksel olarak modellendiği regresyon doğrusudur.

Saraçlı ve diğerleri (2011) yaptıkları çalışmada sırası ile X ve Y değişkenlerini bağımlı olarak düşünerek sağlam tekniklerle elde ettikleri regresyon doğrularının açıortayını alarak, Sağlam Tip II regresyon tekniğinin performansını incelemiş ve veri setinin aykırı değer içerdiği ve her iki değişkeninde ölçüm hatasına maruz kaldığı durumlarda diğer tekniklere göre bu tekniğin üstünlüğünü vurgulamışlardır.

2.2.3 Bulanık En Küçük Kareler Açıortay Regresyon Tekniği

Bu teknikte, bulanık sayılardan oluşan X ve Y değişkenleri sırasıyla bağımlı değişken olarak düşünülerek hesaplanan regresyon doğrularının açıortayı olan yeni regresyon doğrusu elde edilir.

Bu teknik üzerine Kılıç ve diğerlerinin (2011) yaptığı çalışmada veri setinin bulanık olması durumunda farklı dağılış biçimleri ve farklı örneklem büyüklüklerinde veri setinin aykırı değer içerdiği ve içermediği durumlarda EKK Açıortay tekniğine göre daha iyi sonuçlar verdiği yapılan bir benzetim çalışması ile göstermişlerdir.

2.2.4 Sağlam Bulanık EKK Açıortay Regresyon Tekniği

Bu teknikte bulanık bir veri seti ve bu veri seti içerisinde aykırı değerler söz konusu olmakla beraber bağımlı ve bağımsız değişkenlerin her ikisinin de ölçüm hatası içermesi durumunda hesaplanan regresyon doğrularının açıortay doğrusu hesaplanarak çözümleme gerçekleştirilir.

Bilindiği üzere bulanıklık ya verilerin ya da modelin bulanıklaştırılması ile sağlanmaktadır. Bu teknik dikkate alınarak yapılan çözümlemelerde, veri seti sağdan ve soldan bulanıklaştırılarak içerisine aykırı değerler eklenir ve önce X'in daha sonra ise Y'nin bağımlı değişken olarak ele alınması ile hesaplanan iki regresyon doğrusunun açıortay doğrunun hesaplanması sonucunda sağlam bulanık EKK Açıortay doğrusu hesaplanır.

Literatürde sağlam EKK açıortay, bulanık EKK Açıortay, Sağlam Bulanık EKK Açıortay tekniklerine üzerine saraçlı vd den başkalarının bir çalışması bulunmamaktadır.

3. MATERYAL ve METOT

EKK Açıortay, Bulanık EKK Açıortay, Sağlam EKK Açıortay ve Bulanık Sağlam EKK Açıortay regresyon tekniklerinin karşılaştırmalı olarak incelenmesi ve matematiksel bir modelinin oluşturulmasının amaçlandığı çalışmanın bu bölümünde, MATLAB paket programı yardımıyla rassal olarak üretilen farklı örneklem büyüklüklerinde (n=10, 50 ve 100), farklı dağılış biçimlerinde ve aykırı değer içerip içermediği durumlar için daha önceki bölümde söz edilen regresyon tekniklerinin performansları HKO ve Göreli Etkinlik (R.E.) kriterine göre karşılaştırılmıştır.

Aykırı değerler için veri seti Dixon'un "Aykırı Değer Modeli" aracılığı ile elde edilmiştir (bkz. Dixon, 1950). Bu modelde, Eşitlik 4.1.'de görüldüğü üzere, simülasyon yardımı ile üretilecek olan N birimlik veri setinin N-r kadarı istenilen özelliklere uygun değerlerden, r kadarı ise aykırı değerlerden oluşmaktadır. Bu çalışmada aykırı değer içeren veri seti oluşturulurken, veri setinin %5'i aykırı verilerden (ortalaması daha yüksek olan), %95'i ise istenilen özellikteki (örneğin: 30 serbestlik derecesi ile Student dağılmış) verilerden oluşmaktadır (Saraçlı 2008).

$$R = [5+0.1*N] \tag{3.1}$$

Benzetim aracılığı ile yapılan tekrarlar $N = [100000/n]$ defa yapılmıştır. HKO değeri Eşitlik 3.2'de verildiği gibi hesaplanmaktadır.

$$\mathrm{HKO} = \frac{\sum(Y_i - \hat{Y}_i)^2}{n-k} \tag{3.2}$$

$$\mathrm{R.E.} = \frac{\text{Regresyon Tekniğinin HKO}}{\text{Açıortay Regresyon Tekniğinin HKO}} \tag{3.3}$$

MATLAB paket programı ile veri setleri Eşitlik 3.3 – 3.8'de verilmiştir.

$$X_i = T_{(4)} \ , X_i = T_{(10)} \ , X_i = T_{(30)} \tag{3.4}$$

$$x_i = X_i + \mathrm{e}_{\mathrm{ix}} \tag{3.5}$$

$$\mathrm{e}_{\mathrm{ix}} \sim \mathrm{N}(0,1) \tag{3.6}$$

$$e_{iy} \sim N(0,1) \tag{3.7}$$

$$Y_i = X_i + u_i \tag{3.8}$$

$$y_i = Y_i + \mathrm{e}_{\mathrm{iy}} \tag{3.9}$$

Aykırı değerler için

$$X_i = 5 + T_{(4)} \tag{3.10}$$

Veri seti ortalaması 0 varyansı 1 olan normal dağılımdan üretilmiştir. Aykırı değerler ise ortalaması 5 alınarak mevcut veri setinden farklı olarak üretilmiştir.

Bulanık veriler Eşitlik 3.10 – 3.13' de verildiği gibi türetilmiştir.

$$\underline{\delta}_{\mathrm{i}} = X_i/7 \tag{3.11}$$

$$\bar{\delta}_{\mathrm{i}} = X_i/6 \tag{3.12}$$

$$\underline{\omega}_{\mathrm{i}} = Y_i/8 \tag{3.13}$$

$$\overline{\omega}_{\mathrm{i}} = Y_i/7 \tag{3.14}$$

4. BULGULAR

Örneklem büyüklüğünün n=10 olması durumunda ve aykırı değer içermeyen 4, 10 ve 30 serbestlik dereceli olan t dağılımı için EKK regresyon tekniklerinin karşılaştırılmasına yönelik benzetim sonuçları aşağıda Çizelge 3.1' de verilmiştir.

Çizelge 3.1 $t_{(4)}$, $t_{(10)}$ ve $t_{(30)}$ dağılımlarında, n=10 için EKK regresyon teknikleri benzetim sonuçları

Regresyon Tekniği	n=10											
	$t_{(4)}$				$t_{(10)}$				$t_{(30)}$			
	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	*R.E*
EKK(X/Y)	3,946	0,605	1,600	104	4,497	0,552	1,546	103	4,816	0,517	1,582	103
EKK(Y/X)	5,593	0,443	1,015	164	6,194	0,378	0,895	179	6,495	0,350	0,857	191
EKK Açıortay	4,919	0,508	1,666	100	5,518	0,449	1,605	100	5,804	0,419	1,641	100

Serbestlik derecesinin 4, 10 ve 30 olduğu t dağılımında ve aykırı değer içermeyen bulanık veri setinde EKK(X/Y) regresyon tekniğinin HKO değerleri sırasıyla 1,600, 1,546 ve 1,582 olarak gözlemlenmiştir.

Y değişkeninin bağımlı, X değişkeninin bağımsız olduğu EKK(Y/X) regresyon tekniğinde HKO'sı sırasıyla 1,015, 0,895 ve 0,857 olarak hesaplanmış ve EKK(Y/X) regresyon tekniğinin EKK(X/Y) regresyon tekniğine göre daha az hatayla çözümleme yaptığı gözlemlenmiştir. Örneklem hacminin n=10 birim olduğu ve aykırı değer içermeyen veri setlerinde EKK açıortay regresyon tekniğinde HKO'sı 4, 10 ve 30 serbestlik dereceli dağılımlar için HKO değerleri 1,666, 1,605 ve 1,641 olarak hesaplanmıştır.

Bulanık veri setinde aykırı değer söz konusu olmadığı zaman, dağılımın serbestlik derecesi arttığında EKK(Y/X) regresyon tekniğinin daha etkili olduğu ve EKK(X/Y) regresyon tekniğinin serbestlik derecesinden etkilenmediği söylenebilir. Ayrıca 4, 10 ve 30 serbestlik derecelerinde EKK(Y/X) regresyon tekniği, EKK(X/Y) regresyon tekniğine göre daha etkili bir performansa sahip olduğu tespit edilmiştir.

n=10 birim örneklem büyüklüğünde, aykırı değer içermeyen ve farklı serbestlik dereceli t dağılımına sahip veri setleri için Bulanık EKK regresyon tekniklerinin karşılaştırılmasına yönelik benzetim sonuçları Çizelge 3.2'de verilmiştir.

Çizelge 3.2 $t_{(4)}$, $t_{(10)}$ ve $t_{(30)}$ dağılımlarında, n=10 için Bulanık EKK regresyon teknikleri benzetim sonuçları

Regresyon Tekniği	n=10											
	$t_{(4)}$				$t_{(10)}$				$t_{(30)}$			
	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	*R.E*
Bulanık EKK(X/Y)	-0,155	1,003	1,408	98,5	-0,159	1,008	1,346	98,6	-0,157	1,004	1,388	98,7
Bulanık EKK(Y/X)	0,382	0,945	1,297	107	0,381	0,941	1,245	106	0,381	0,944	1,281	106
Bulanık EKK Açıortay	0,262	0,972	1,388	100	0,294	0,972	1,328	100	0,268	0,972	1,370	100

Bulanık EKK regresyon tekniklerinin HKO'sı kriterine göre performanslarının karşılaştırıldığı ve örneklem hacminin n=10 olduğu farklı serbestlik derecelerinde Bulanık EKK(Y/X) regresyon tekniğinin HKO değerleri, Bulanık EKK(X/Y) regresyon tekniğinin HKO değerlerinden daha düşük çıkmıştır. Bulanık EKK açıortay regresyon tekniğinin 4, 10 ve 30 serbestlik dereceli t dağılımda ve aykırı değer içermeyen bulanık veri setin için hesaplanan HKO değerleri sırasıyla 1,388, 1,328 ve 1,370 olarak hesaplanmıştır. Bulanık bir veri setinde Bulanık EKK regresyon tekniklerinin klasik EKK regresyon tekniklerine göre daha iyi performans göstermesi yapılan analizlerin doğruluğunun bir göstergesidir.

Aykırı değer içermeyen n=10 örneklem hacmindeki bulanık veri setlerinde, serbestlik derecesinin artması Bulanık EKK(Y/X) ve Bulanık EKK(X/Y) regresyon tekniklerinin etkinliklerini etkilemediği söylenebilir. Ancak bu serbestlik derecelerinde Bulanık EKK(Y/X) regresyon tekniğinin Bulanık EKK(X/Y) regresyon tekniğine göre daha etkili olduğu söylenebilir.

Serbestlik derecesi 4, 10 ve 30 olan t dağılımı için örneklem büyüklüğünün n=50 olması durumunda aykırı değer içermeyen bulanık veri setinde, EKK regresyon tekniklerinin karşılaştırıldığı benzetim sonuçları Çizelge 3.3' da verilmiştir.

Çizelge 3.3 n=50 hacmindeki ve $t_{(4)}$, $t_{(10)}$ ve $t_{(30)}$ dağılımlarındaki veri setleri için EKK regresyon tekniklerine ait benzetim sonuçları.

Regresyon Tekniği	n=50											
	$t_{(4)}$				$t_{(10)}$				$t_{(30)}$			
	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	*R.E*
EKK(X/Y)	3,512	0,648	1,291	109	4,413	0,560	1,257	108	4,811	0,519	1,289	107
EKK(Y/X)	5,220	0,480	0,801	176	6,092	0,388	0,666	204	6,476	0,352	0,612	225
EKK Açıortay	4,404	0,559	1,415	100	5,318	0,469	1,363	100	5,684	0,431	1,383	100

Çizelge 3.3'da, örneklem büyüklüğünün n=50 olduğu ve 4 sd.li t dağılımı için EKK(Y/X) regresyon tekniğinde elde edilen HKO=0,801 değerinin, EKK(X/Y) regresyon tekniğinden elde edilen HKO=1,291 değerinden daha düşük olduğu görülmektedir. $t_{(10)}$ ve $t_{(30)}$ dağılımları için EKK(Y/X)'e ait HKO' değerinin EKK(X/Y) için bulunan HKO değerinden daha düşük olduğu görülmektedir. Aykırı değer içermeyen bulanık bir veri setinde EKK Açıortay regresyon tekniğine ait HKO değerleri 4, 10 ve 30 serbestlik derecesine sahip t dağılımları için sırasıyla 1,415, 1,363 ve 1,383 olarak hesaplanmıştır.

Örneklem hacminin n=50 olduğu bulanık veri setlerinde, dağılımın serbestlik derecesinin değişmesi durumunda EKK(X/Y) regresyon tekniğinin etkinliğini değiştirmediği gözlemlenmiştir. EKK(Y/X) regresyon tekniğinin etkinliğinin ise dağılımın serbestlik derecesi ile birlikte arttığı görülmektedir.

4, 10 ve 30 serbestlik dereceli dağılımlarda EKK(Y/X) regresyon tekniğinin EKK(X/Y) regresyon tekniğine göre daha etkili olduğu dikkat çekmektedir.

Aykırı değer içermeyen bulanık veri setinde, serbestlik derecesinin 4, 10 ve 30 olduğu t dağılımları için örneklem büyüklüğünün n=50 olması durumunda, Bulanık EKK regresyon tekniklerinin karşılaştırılmasına yönelik benzetim sonuçları Çizelge 3.4'de verilmiştir.

Çizelge 3.4 $t_{(4)}$, $t_{(10)}$ ve $t_{(30)}$ dağılımlarında, n=50 için Bulanık EKK regresyon tekniklerine ait benzetim sonuçları

Regresyon Tekniği	n=50											
	$t_{(4)}$				$t_{(10)}$				$t_{(30)}$			
	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	*R.E*
Bulanık EKK(X/Y)	-0,148	1,000	1,089	98,3	-0,150	1,005	1,037	98,8	-0,150	1,002	1,072	98,4
Bulanık EKK(Y/X)	0,395	0,941	0,998	107	0,392	0,937	0,959	106	0,393	0,940	0,979	107
Bulanık EKK Açıortay	0,281	0,970	1,071	100	0,318	0,970	1,025	100	0,291	0,970	1,055	100

Çizelge 3.4 incelendiğinde, serbestlik derecesinin 4 olduğu, veri setinde Bulanık EKK(X/Y) regresyon tekniği için 1,089 olarak hesaplanan HKO değerinin, Bulanık EKK(Y/X) regresyon tekniği için 0,998 olarak hesaplanan HKO değerinden daha yüksek hesaplandığı görülmektedir.

Serbestlik derecelerinin 10 ve 30 olarak değiştiği t dağılımlarında Bulanık EKK(Y/X) regresyon tekniğinin EKK(X/Y) tekniğine göre daha az hatayla çözümleme yaptığı gözlemlenmiştir.

Dağılımın serbestlik derecesi 4, 10 ve 30 olarak değiştiği durumlarda, Bulanık EKK açıortay regresyon tekniğinin HKO'sı sırasıyla 1,188, 1,190 ve 1,146 olarak tespit edilmiştir.

n=50 örneklem hacmindeki aykırı değer içermeyen bulanık veri setlerinde, dağılımın serbestlik derecesinin değişmesi ile Bulanık EKK(X/Y) ve Bulanık EKK(Y/X) regresyon tekniklerinin etkinliklerinin değişmediği söylenebilir. Ayrıca 4, 10 ve 30 serbestlik derecelerinde EKK(X/Y) regresyon tekniğinin, Bulanık EKK(X/Y) regresyon tekniğine göre daha etkili olduğu söylenebilir.

n=100 birim örneklem hacmine sahip, aykırı değer içermeyen bulanık veri setinde, $t_{(4)}$, $t_{(10)}$ ve $t_{(30)}$ serbestlik dereceli t dağılımları için EKK regresyon tekniklerinin performanslarının karşılaştırılmasına yönelik benzetim sonuçları Çizelge 3.5 yer verilmiştir.

Çizelge 3.5 $t_{(4)}$, $t_{(10)}$ ve $t_{(30)}$ dağılımlarında, n=100 için EKK regresyon teknikleri benzetim sonuçları

Regresyon Tekniği	n=100											
	$t_{(4)}$				$t_{(10)}$				$t_{(30)}$			
	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	*R.E*
EKK(X/Y)	3,420	0,657	1,239	110	4,325	0,569	1,234	109	4,857	0,513	1,260	107
EKK(Y/X)	5,103	0,491	0,778	175	6,054	0,391	0,648	207	6,466	0,353	0,595	227
EKK Açıortay	4,292	0,570	1,365	100	5,251	0,476	1,347	100	5,688	0,430	1,352	100

Çizelge 3.5'de, örneklem hacminin n=100 olduğu ve farklı serbestlik derecelerindeki t dağılımları için EKK(Y/X) regresyon tekniğinin HKO değeri, EKK(X/Y) regresyon tekniğinden HKO değerinden daha düşük çıkmıştır. Ayrıca EKK açıortay regresyon tekniğinin 4, 10 ve 30 serbestlik dereceli t dağılımda ve aykırı değer içermeyen bulanık veri setinde HKO'sı sırasıyla 1,365, 1,347 ve 1,352 olarak hesaplanmıştır.

Farklı serbestlik derecelerinin EKK(X/Y) regresyon tekniğinin etkinliğini değiştirmediği ancak dağılımın serbestlik derecesi ile beraber EKK(Y/X) regresyon tekniğinin etkinliğinin arttığı gözlemlenmiştir. Ayrıca 4, 10 ve 30 serbestlik dereceli t dağılımlarında EKK(Y/X) regresyon tekniğinin EKK(X/Y) regresyon tekniğine göre daha etkili olduğu söylenebilir.

Serbestlik derecesi 4, 10 ve 30 olan t dağılımı için örneklem büyüklüğünün n=100 olması durumunda EKK regresyon tekniklerinin karşılaştırılmasına yönelik benzetim sonuçları Çizelge 3.6' de verilmiştir.

Çizelge 3.6 $t_{(4)}$, $t_{(10)}$ ve $t_{(30)}$ dağılımlarında, n=100 için Bulanık EKK regresyon tekniklerine ait benzetim sonuçları

Regresyon Tekniği	**n=100**											
	$t_{(4)}$				$t_{(10)}$				$t_{(30)}$			
	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	*R.E*
Bulanık EKK(X/Y)	-0,148	1,001	1,043	98,1	-0,150	1,005	1,011	99,0	-0,148	1,001	1,036	98,3
Bulanık EKK(Y/X)	0,395	0,940	0,955	107	0,395	0,936	0,935	107	0,393	0,940	0,949	107
Bulanık EKK Açıortay	0,283	0,970	1,024	100	0,327	0,970	1,001	100	0,289	0,970	1,019	100

Serbestlik derecesinin 4, 10 ve 30 olduğu t dağılımında aykırı değer içermeyen bulanık veri setinde, Bulanık EKK(X/Y) regresyon tekniğinin HKO değerlerinin sırasyla 1,043, 1,011 ve 1,036 olduğu görülmektedir. Y değişkeninin bağımlı, X değişkeninin bağımsız olarak ele alındığı EKK(Y/X) regresyon tekniğinde sırasıyla HKO değerleri 0,955, 0,935 ve 0,949 olarak hesaplanmıştır. Bulanık EKK(Y/X) regresyon tekniğinin Bulanık EKK(X/Y) regresyon tekniğine göre daha az hatayla çözümleme yaptığı tespit edilmiştir. Aykırı değer içermeyen bulanık bir veri setinde örneklem hacminin n=100 birim olduğu durumda EKK regresyon tekniklerini kullanmak yerine Bulanık EKK regresyon tekniklerini kullanmak, daha az hatayla çözümleme yapılmasına olanak sağlamaktadır.

Bulanık EKK(Y/X) ve Bulanık EKK(X/Y) regresyon tekniklerinin dağılımın serbestlik derecesinin 4, 10 ve 30 olarak artmasıyla, etkinliklerinin değişmediği söylenebilir. Ayrıca bu serbestlik derecelerinde Bulanık EKK(Y/X) regresyon tekniğinin, Bulanık EKK(X/Y) regresyon tekniğine göre daha etkili olduğu tespit edilmiştir.

Serbestlik derecesi 4, 10 ve 30 olan t dağılımı için örneklem büyüklüğünün n=10 olması durumunda aykırı değer içeren Sağlam regresyon tekniklerinin karşılaştırılmasına yönelik benzetim sonuçları Çizelge 3.7' de verilmiştir.

Çizelge 3.7 $t_{(4)}$, $t_{(10)}$ve $t_{(30)}$ dağılımlarında, n=10 ve aykırı değer içeren durumlar için Sağlam regresyon teknikleri benzetim sonuçları

Regresyon Tekniği	n=10											
	$t_{(4)}$-A				$t_{(10)}$-A				$t_{(30)}$-A			
	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	*R.E*
Huber(X/Y)	1,659	0,839	1,619	95,8	1,764	0,828	1,563	96,2	1,965	0,806	1,530	94,5
Huber(Y/X)	2,342	0,771	1,370	113	2,705	0,737	1,312	114	2,676	0,739	1,267	114
Huber Açıortay	2,103	0,796	1,552	100	2,341	0,773	1,504	100	2,393	0,764	1,447	100

Çizelge 3.7'de n=10 örneklem büyüklüğündeki 4 sd.li t dağılımına sahip ve aykırı değer içeren veri seti için X'in bağımsız Y'nin bağımlı değişken olarak ele alınması ile Huber regresyon tekniğine (Huber(Y/X)) göre yapılan çözümleme için elde edilen HKO=1.3709 değerinin, X'in bağımlı Y'nin bağımsız değişken olarak ele alınması ile (Huber(Y/X)) yapılan çözümlemenin HKO=1.6191 değerinden daha düşük olduğu tespit edilmiştir. Benzer olarak $t_{(10)}$ ve $t_{(30)}$ dağılımları için Huber(Y/X)'e ait HKO'nın Huber (X/Y) için bulunan HKO'dan düşük olduğu saptanmıştır.

n = 10 örneklem hacmindeki aykırı değer içeren veri setleri için sd.'si arttıkça Huber Açıortay regresyon tekniğine ait HKO değerlerinin azaldığı dolayısı ile sd arttıkça bu tekniğin performansının arttığı dikkat çekmektedir.

Çizelge 3.7'de verilen sağlam regresyon tekniklerinin göreli etkinlikleri incelendiğinde Huber(Y/X) regresyon tekniğinin Huber(X/Y) regresyon tekniğine göre daha etkin olduğu söylenebilir. Ayrıca dağılımın serbestlik derecesinin değişmesi tekniklerin etkinleri üzerinde etkili olmadığı görülmektedir.

4, 10, ve 30 serbestlik dereceli t dağılımı için örneklem hacminin 10 birim olduğu ve veri setinin aykırı değer içermesi durumunda Bulanık Sağlam regresyon tekniklerinin benzetim sonuçları Çizelge 3.8'de yer almaktadır.

Çizelge 3.8 $t_{(4)}, t_{(10)}$ve $t_{(30)}$ dağılımlarında, n=10 ve aykırı değer içeren durumlar için Bulanık Sağlam regresyon teknikleri benzetim sonuçları

Regresyon Tekniği	**n=10**											
	$t_{(4)}$**-A**				$t_{(10)}$**-A**				$t_{(30)}$**-A**			
	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	*R.E*
Bulanık Huber(X/Y)	1,626	0,840	1,655	96,1	1,723	0,830	1,601	96,0	1,906	0,809	1,563	95,9
Bulanık Huber(Y/X)	1,635	0,842	1,644	96,8	1,740	0,831	1,591	96,6	1,931	0,809	1,555	96,4
Bulanık Huber Açıortay	1,635	0,841	1,592	100	1,733	0,830	1,537	100	1,929	0,809	1,500	100

Çizelge 3.8'de, n=10 örneklem büyüklüğünde, 4 sd.li t dağılımına sahip ve aykırı değer içeren veri setinin bulanık olması durumunda X'in bağımsız Y'nin bağımlı değişken olarak ele alınması ile Bulanık Huber regresyon tekniği (Bulanık Huber(Y/X)) için elde edilen HKO=1,644, değerinin X'in bağımlı Y'nin bağımsız değişken olarak ele alınması (Bulanık Huber(Y/X)) ile elde edilen HKO=1.655 değerinden daha düşük olduğu tespit edilmiştir. Bu durum 10 ve 30 serbestlik dereceli t dağılımları içinde geçerlidir.

Veri setinin bulanık olması durumunda serbestlik derecesinin arttığı t dağılımları için Bulanık Huber Açıortay regresyon tekniğine ait HKO değerlerinin gittikçe azaldığı görülmekle beraber bu tekniğin diğer iki tekniğe göre daha iyi performans sergilediği de söylenebilir.

Bulanık bir veri setinde aykırı değer söz konusu olduğu zaman regresyon tekniklerinin birbirlerine göre çok fazla etkin bir çözümleme gerçekleştirdiği söylenemez. Ayrıca dağılımın serbestlik derecesinin değişmesi etkinlik üzerinde etkili olmadığı dikkat çekmektedir.

Serbestlik derecesi 4, 10 ve 30 olan t dağılımı için örneklem büyüklüğünün n=50 olması durumunda aykırı değer içeren veri seti için, sağlam regresyon tekniklerinin karşılaştırılmasına yönelik benzetim sonuçları Çizelge 3.9'de sunulmuştur.

Çizelge 3.9 $t_{(4)}, t_{(10)}$ ve $t_{(30)}$ dağılımlarında, n=50 ve aykırı değer içeren durumlar için Sağlam regresyon teknikleri benzetim sonuçları

Regresyon Tekniği	n=50											
	$t_{(4)}$-A				$t_{(10)}$-A				$t_{(30)}$-A			
	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	*R.E*
Huber(X/Y)	0,586	0,948	1,147	103	0,466	0,959	1,113	106	0,600	0,945	1,099	104
Huber(Y/X)	1,102	0,900	1,096	108	1,331	0,878	1,111	107	1,191	0,893	1,065	107
Huber Açıortay	0,874	0,923	1,188	100	0,950	0,916	1,190	100	0,911	0,917	1,146	100

Serbestlik derecesinin 4 olduğu, aykırı değer içeren t dağılımı için Huber(X/Y) regresyon tekniğinin HKO'sı 1,147 iken, Huber(Y/X) regresyon tekniğinin HKO'sı 1,096 olarak (daha düşük) olarak hesaplanmıştır. Serbestlik derecelerinin 10 ve 30 olarak değiştiği dağılımlarda ise Huber(Y/X) regresyon tekniğinin daha düşük hatayla çözümleme yaptığı gözlemlenmiştir. Huber Açıortay regresyon tekniğinin, serbestlik derecesi 4, 10 ve 30 serbestlik dereceli t dağılımı için, HKO'sı sırasıyla 1,188, 1,190 ve 1,146 olarak hesaplandığı görülmektedir. Bu sonuca göre dağılım değiştikçe tekniğin performansı için bir şey söylemek mümkün değildir.

4 ve 30 serbestlik dereceli t dağılımlarında aykırı değer söz konusu olduğu zaman Huber(Y/X) regresyon tekniği, Huber(X/Y) regresyon tekniğine göre daha etkin bir performans gösterdiği R.E (göreli etkinlik) değerlerine bakılarak tespit edilmiştir. Ancak 10 serbestlik dereceli t dağılımında ise bu iki tekniğin birbirlerine göre daha etkin olmadığı söylenebilir.

Serbestlik derecesi 4, 10 ve 30 olan t dağılımı ($t_{(4)}$, $t_{(10)}$ve $t_{(30)}$) için örneklem büyüklüğünün n=50 olması ve verilerin aykırı değer içermesi durumunda sağlam regresyon tekniklerinin karşılaştırılmasına yönelik benzetim sonuçları Çizelge 3.10' de yer verilmiştir.

Çizelge 3.10 $t_{(4)}$, $t_{(10)}$ ve $t_{(30)}$ dağılımlarında, n=50 ve aykırı değer içeren durumlar için Bulanık Sağlam regresyon teknikleri benzetim sonuçları

Regresyon Tekniği	n=50											
	$t_{(4)}$-A				$t_{(10)}$-A				$t_{(30)}$-A			
	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	*R.E*
Bulanık Huber(X/Y)	0,563	0,949	1,152	98,9	0,435	0,961	1,117	98,9	0,550	0,948	1,100	98,8
Bulanık Huber(Y/X)	0,565	0,951	1,147	99,3	0,440	0,963	1,113	99,2	0,554	0,949	1,093	99,4
Bulanık Huber Açıortay	0,567	0,950	1,140	100	0,435	0,962	1,105	100	0,563	0,948	1,087	100

Çizelge 3.10'de verilen, örneklem büyüklüğü n=50 olan t dağılımı ($t_{(4)}$) için X'in bağımsız Y'nin bağımlı değişken olarak ele alındığı Bulanık Huber(Y/X) regresyon tekniğinin için HKO=1,147 iken, X'in bağımlı Y'nin bağımsız değişken olarak ele alındığı Bulanık Huber(Y/X) regresyon tekniği için HKO=1,152 olarak elde edilmiştir. t dağılımının serbestlik derecesi arttıkça ($t_{(10)}$ ve $t_{(30)}$ olduğunda) Bulanık Huber(Y/X) regresyon tekniğinin Bulanık Huber(X/Y) regresyon tekniğine göre daha düşük HKO'sına sahip olduğu gözlemlenmiştir. Bulanık Huber Açıortay regresyon tekniğinin $t_{(4)}$, $t_{(10)}$ ve $t_{(30)}$ dağılımlarında sırasıyla HKO'sının 1,140, 1,105 ve 1,087 olarak tespit edilmiştir dolayısıyla böyle bir veri seti için t dağılımının serbestlik derecesi arttıkça Bulanık Huber açıortay regresyon tekniğinin HKO'sı değerinin düştüğü tespit edilmiştir.

Göreli etkinlikleri bakımından incelendiğinde, Bulanık Huber(Y/X) regresyon tekniği, Bulanık Huber(X/Y) regresyon tekniğine göre sadece 4 serbestlik dereceli t dağılımı için etkinken, dağılım serbestlik derecesinin 10 ve 30 olduğu durumlarda böyle bir etkinlikten söz etmek mümkün olamamaktadır.

Örneklem hacminin 100 birim olduğu ve aykırı değer içeren veri seti için farklı serbestlik derecelerine sahip t dağılımı için sağlam regresyon tekniklerinin benzetim sonuçları Çizelge 3.11'de verilmiştir.

Çizelge 3.11 $t_{(4)}$, $t_{(10)}$ ve $t_{(30)}$ dağılımlarında, n=100 ve aykırı değer içeren durumlar için Sağlam regresyon teknikleri benzetim sonuçları

Regresyon Tekniği	n=100											
	$t_{(4)}$-A				$t_{(10)}$-A				$t_{(30)}$-A			
	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	*R.E*
Huber(X/Y)	0,552	0,952	1,110	103	0,433	0,963	1,065	106	0,588	0,948	1,066	103
Huber(Y/X)	0,986	0,912	1,056	108	1,195	0,892	1,041	109	1,043	0,907	1,022	108
Huber Açıortay	0,794	0,931	1,148	100	0,848	0,926	1,138	100	0,828	0,927	1,107	100

Çizelge 3.11'deki sonuçlara göre örneklem büyüklüğünün 100 olması durumda 4 serbestlik dereceli t dağılımı için X'in bağımsız Y'nin bağımlı değişken olarak ele alındığı Huber(Y/X) regresyon tekniğinde elde edilen HKO=1,056 değerinin, X'in bağımlı Y'nin bağımsız değişken olarak ele alındığı Huber(X/Y) regresyon tekniğinden elde edilen HKO=1,110 değerine göre daha düşük olduğu tespit edilmiştir. 10 ve 30 serbestlik derecelerine sahip dağılımlarda Huber(Y/X) regresyon tekniğinin Huber(X/Y) regresyon tekniğine göre daha düşük HKO'sına sahip olduğu gözlemlenmiştir. Bulanık Huber Açıortay regresyon tekniğinin $t_{(4)}$, $t_{(10)}$ ve $t_{(30)}$ dağılımlarında sırasıyla HKO değerleri sırasıyla 1,148, 1,138 ve 1,107 olarak hesaplanmıştır. Aykırı değer içeren ve 100 birim örneklem hacmine sahip bir veri setinde, dağılımın serbestlik derecesi arttıkça Huber açıortay regresyon tekniğinin HKO değerinin de azaldığı dikkatleri çekmektedir.

4, 10 ve 30 serbestlik dereceli t dağılımlarında, Huber(X/Y) regresyon tekniğinin en etkili çözümleme gerçekleştirdiği serbestlik derecesinin 10 olduğu söylenebilir. Ayrıca Huber (Y/X) regresyon tekniğinde örneklem hacminin 100 birim olduğu durumda dağılımın serbestlik derecesinin, tekniğin etkinliği üzerinde bir etkisinin bulunmadığı tespit edilmiştir.

Aykırı değer içeren ve serbestlik derecesi 4, 10 ve 30 olan t dağılımı için örneklem büyüklüğünün n=100 olması durumunda, Bulanık Sağlam regresyon tekniklerinin karşılaştırılmasına yönelik benzetim sonuçları Çizelge 3.12' da yer almaktadır.

Çizelge 3.12 $t_{(4)}$, $t_{(10)}$ ve $t_{(30)}$dağılımlarında, n=100 ve aykırı değer içeren durumlar için Bulanık Sağlam regresyon teknikleri benzetim sonuçları

Regresyon Tekniği	n=100											
	$t_{(4)}$-A				$t_{(10)}$-A				$t_{(30)}$-A			
	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	R.E	$\hat{\beta}_0$	$\hat{\beta}_1$	*HKO*	*R.E*
Bulanık Huber(X/Y)	0,529	0,953	1,111	99,3	0,398	0,965	1,066	99,2	0,538	0,950	1,066	99,2
Bulanık Huber(Y/X)	0,525	0,955	1,105	99,9	0,391	0,968	1,061	99,7	0,535	0,952	1,060	99,8
Bulanık Huber Açıortay	0,531	0,954	1,104	100	0,392	0,966	1,058	100	0,548	0,951	1,058	100

Çizelge 3.12'da verilen benzetim sonuçlarına göre 4, 10 ve 30 serbestlik derecesine sahip aykırı değer içeren veri setinde Bulanık Huber(X/Y) regresyon tekniği için HKO değerleri sırasıyla 1,111, 1,066 ve 1,066 olarak hesaplanmıtır. Bulanık Huber(Y/X) regresyon tekniğinin 4, 10 ve 30 serbestlik derecelerindeki t dağılımı için HKO değerleri ise sırasıyla 1,105, 1,061 ve 1,060 olduğu ve bu değerlerin Bulanık Huber(X/Y) regresyon tekniği için hesaplanan değerlerden daha düşük olduğu gözlemlenmiştir.

Bulanık Huber Açıortay regresyon tekniğinin serbestlik derecesi 4, 10 ve 30 olan t dağılımı için HKO değerleri sırasıyla 1,104, 1,058 ve 1,058 olarak hesaplanmış ve serbestlik derecesinin 10 ve daha fazla olması durumunda HKO değerinin değişmediği sonucuna gözlemlenmiştir.

Bulanık Huber (Y/X) regresyon tekniğinin örneklem hacminin n=100 birim olduğu durumlarda, veri setinde hem aykırı değer hem de bulanıklık söz konusu iken 4, 10 ve 30 serbestlik dereceli dağılımlarda etkinliğini değiştirmediği söylenebilir. Ancak bu serbestlik derecelerinde Bulanık Huber(X/Y) regresyon tekniği t dağılımın serbestlik derecesinin 10 olduğu durumda diğer tekniklere göre daha etkin olduğu gözlemlenmiştir. Ayrıca Bulanık Huber(Y/X) regresyon tekniğinin 10 serbestlik dereceli t dağılımında Bulanık Huber(X/Y) regresyon tekniğine göre daha etkili performans gösterdiği gözlemlenmiştir.

Farklı örneklem hacimlerindeki ve 4, 10 ve 30 serbestlik dereceli t dağılımına sahip regresyon tekniklerini R.E. kriterine göre karşılaştırılması Çizelge 3.13'de verilmiştir.

Çizelge 3.13 Regresyon tekniklerinin R.E. kriterine göre karşılaştırılması

Regresyon tekniği	$t_{(4)}$			$t_{(10)}$			$t_{(30)}$		
	n=10	n=50	n=100	n=10	n=50	n=100	n=10	n=50	n=100
EKK(X/Y)	104	109	110	103	108	109	103	107	107
EKK(Y/X)	164	176	175	179	204	207	191	225	227
EKK Açıortay	100	100	100	100	100	100	100	100	100
Bulanık EKK(X/Y)	98,5	98,3	98,1	98,6	98,8	99	98,7	98,4	98,3
Bulanık EKK(Y/X)	107	107	107	106	106	107	106	107	107
Bulanık EKK Açıortay	100	100	100	100	100	100	100	100	100
Huber(X/Y)	95,8	103	103	96,2	106	106	94,5	104	103
Huber(Y/X)	113	108	108	114	107	109	114	107	108
Huber Açıortay	100	100	100	100	100	100	100	100	100
Bulanık Huber(X/Y)	96,1	98,9	99,3	96	98,9	99,2	95,9	98,8	99,2
Bulanık Huber(Y/X)	96,8	99,3	99,9	96,6	99,2	99,7	96,4	99,4	99,8
Bulanık Huber Açıortay	100	100	100	100	100	100	100	100	100

4 serbestlik dereceli t dağılımında, örneklem hacminin 50 ve daha fazla olması EKK(Y/X) ve EKK(X/Y) regresyon tekniklerinin etkinliklerini değiştirmediği gözlemlenmiştir. Ancak t dağılımın serbestlik derecesinin 10 ve 30 olması durumunda durumun da, örneklem hacmi ile birlikte EKK(Y/X) regresyon tekniğinin etkinliğinin de arttığı dikkat çekmektedir.

Bulanık bir veri setine sahip farklı serbestlik dereceli t dağılımlarında örneklem hacmin değişmesi Bulanık EKK(X/Y) ve Bulanık EKK(Y/X) regresyon tekniklerinin etkinlikleri üzerinde bir değişikliğe sebep olmamaktadır.

Veri setinin aykırı değer içermesi durumunda, Huber(Y/X) ve Huber(X/Y) regresyon tekniklerinin göreli etkinlik değerlerinin dağılımın serbestlik derecesi ile bir bağlantısı olmadığı tespit edilmiştir. Ayrıca örneklem hacminin 50 ve daha fazla olması ile bu regresyon tekniklerinin etkinliklerinin değişmediği söylenebilir.

Bulanık bir veri setinde aykırı değer söz konusu olduğu zaman 4, 10 ve 30 serbestlik dereceli t dağılımlarında en etkili tekniğin Bulanık Huber açıortay regresyon tekniği olduğu gözlemlenmiştir. Örneklem hacminin 50 ve daha fazla olması durumunda ise Bulanık Huber(Y/X), Bulanık Huber(X/Y) ve Bulanık Huber açıortay regresyon tekniklerinin etkinliklerinin değişmediği dikkatleri çekmektedir.

5. TARTIŞMA ve SONUÇ

Sağlam ve Bulanık Doğrusal Tip II regresyon tekniklerinin performanslarının karşılaştırmalı olarak incelenmesi amaçlanmış olup, bu araştırmada ulusal çapta çok fazla olmayıp ancak araştırma ile ilgili bazı çalışmalar (Şanlı ve Apaydın 2004, Başaran ve Günay 2003, Güneş 2001, Yurtcu ve İçağa 2007, Çetin ve Orsoy 2001, Ergül 2006, Yorulmaz 2009) aşağıda sırasıyla kısaca özetlenmiştir.

Tek bir gözlem bile regresyon çözümlemesinde büyük bir öneme sahiptir. Veri setinden aykırı değer olan bir veya birkaç gözlem değerlerinin çıkarılması regresyon denklemini tamamen değiştirebilir. Bulanık bir veri setinde aykırı gözlem değerleri söz konusu olduğu durumlarda üyelik fonksiyonları yardımıyla ağırlık matrislerinin oluşturulduğu bu çalışmada EKK, sağlam teknikler ve önerilen bulanık sağlam teknikler ile regresyon teknikleri tahmin edilmiş ve sonuçları karşılaştırılmıştır.

İlk kez Tanaka tarafından önerilen bulanık doğrusal regresyon analizi ve bulanık küme kuramı yöntemleri anlatıldığı bu çalışmada bu iki yöntemin bazı eksikliklerini gidermek için değişik yaklaşımlar önerilmiştir.

Bağımlı ve bağımsız değişkenler arasında var olan ve modellenebilen doğrusal ilişkilerin soyut modellerindeki katsayıların bulandırılmasının nasıl sağlandığı anlatılan bu çalışmada sayısal bir örnekle hataların minimize edilmesi amaçlanmıştır.

Belirsizliği modellemede başarılı bir yöntem olan bulanık mantığın ileri dönük tahmin konusundaki belirsizlikler için önerdiği bulanık regresyon yaklaşımı genel hatlarıyla incelendiği bu çalışma sayısal bir örnekle sunulmuştur.

Tıbbi bir veri seti üzerinde uygulanan bu çalışmada, regresyon analizinde çok sık kullanılan EKK yöntemi ile sağlam M yöntemi ve yüksek bozulma noktasına sahip En Küçük Medyan Kareler (EMK) ve MM yöntemlerinin çalışıldığı bu çalışmada EKK yöntemi ile karşılaştırılmıştır.

Sağlam regresyon teknikleri için parametre kestirim algoritmalarının incelendiği bu çalışmada basit doğrusal regresyon analizine kısa bir giriş yapılmış ve sağlam teknikler ayrıntılı olarak sunulmuştur. Ayrıca gerçek veriler üzerinde bir uygulama yapılmıştır.

Aykırı gözlem varlığında kullanılabilecek EKK tahmin tekniğine alternatif sağlam tahmin tekniklerinin bir kısmını tanıtarak, çeşitli veriler üzerinde aykırı gözlemlerin EKK ile sağlam tahmin ediciler üzerindeki etkisini incelemek ve aykırı gözlemleri teşhis etmeyi amaçlamıştır.

Matlab paket proğramı ile Monte-Carlo benzetim çalışmasının yapıldığı bu çalışmada benzetim sonuçları incelendiğinde;

EKK Açıortay regresyon tekniğinde, aynı serbestlik dercesine sahip farklı örneklem büyüklüklerin de bulanık veri setleri için HKO'sı kriterine bakarak performansları karşılaştırıldığında örneklem hacminin artmasıyla HKO'sı değeri de azaldığı dikkat çekmektedir. Ayrıca dağılımın serbestlik derecelerinin 4, 10 ve 30 olduğu durumlarda en düşük HKO'sı değeri 10 serbestlik dereceli t dağılımına ait olduğu gözlemlenmiştir.

Bulanık EKK Açıortay regresyon tekniğinde, farklı dağılış biçimlerinde ve farklı örneklem büyüklüklerinde, örneklem hacmin arttığı zaman HKO değerinin düştüğü ve 4, 10 ve 30 serbestlik dereceli t dağılımlarında en düşük HKO'sı değeri 10 serbestlik dereceli dağılıma ait olduğu tespit edilmiştir.

Huber Açıortay regresyon tekniğinde ise, farklı dağılış biçimlerinde, farklı örneklem büyüklüklerindeki veri setleri karşılaştırılmıştır ve sonuç olarak aynı serbestlik derecesine sahip farklı örneklem büyüklüklerinde üretilmiş veri setlerinde örneklem hacmi arttıkça HKO değeri düşmektedir. Dağılımın serbestlik derecesinin HKO üzerinde bir etkisinin olmadığı tespit edilmiştir.

Bulanık Huber Açıortay regresyon tekniğinde, farklı dağılış biçimlerinde, farklı örneklem büyüklüklerindeki veri setleri karşılaştırılmıştır ve sonuç olarak aynı serbestlik derecesine sahip farklı örneklem büyüklüklerinde üretilmiş veri setlerinde örneklem hacmi arttıkça HKO değeri düşmektedir. Ayrıca dağılımın serbestlik derecesi arttığı zaman HKO'nın yine düştüğü gözlemlenmiştir.

İlerleyen aşamalarda çalışmanın doğrusal olamayan modellere uyarlanması ve diğer Tip II tekniklerine kapsayacak şekilde genişletilmesi planlanmaktadır. Yapılan bu çalışmanın ölçüm hatalı regresyon teknikleri yardımıyla çözümleme gerçekleştirecek olan diğer araştırmacılara faydalı olması ve farklı bakış açıları sunması ümit edilmektedir.

6. KAYNAKLAR

Altaş, İ. H. (1999). Bulanık Mantık: Bulanıklık Kavramı Enerji, Elektrik, Elektromekanik-3e, Bilişim Yayıncılık A.Ş İstanbul, **62**: 80-85.

Apaydın, A. ve Şanlı, K. (2004). The Fuzzy Robust Regression Analysis, The Case of Fuzzy Data Set Has Outlier, *G.U Journal of Science* **17(3):** 71-84.

Başaran, A. ve Günay, S. (2003). Bulanık Doğrusal Regresyon, Ulusal Ekonometri ve İstatistik Sempozyumu, Gazi Üniversitesi, Ankara.

Çankal, E. (2010). İstatistik, Lisan Yayınları, Ankara, 290.

Çetin, M.C. ve Orsoy, A. (2001). Doğrusal Regresyonda Sağlam Tahmin Ediciler ve Bir Uygulama, *Anadolu Üniversitesi Bilim ve Teknoloji Dergisi*, **2:** 265-270

Çevik, O. ve Yıldırım, Y. (2010). Bulanık Doğrusal Programlama ile Süt Ürünleri İşletmesinde Bir Uygulama, *Karamanoğlu Mehmet Bey Üniversitesi Sosyal ve Ekonomik Araştırmalar Dergisi* **12 (18):** 15-26.

Dixon, W.J. (1950). Analysis of extreme value. Annals Math. Stat., **21:** 488-506

Düzyurt, S. (2008). Bulanık Regresyon ile Tahmin ve Bir Uygulama, Yüksek Lisans Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.

Ergül, B. (2006). Robust Regresyon ve Uygulamaları. Yüksek Lisans Tezi**,** Eskişehir Osmangazi Üniversitesi, Fen Bilimleri Enstitüsü, Eskişehir.

Gazeloğlu, C. ve Saraçlı, S. (2011). Çoklu Tip II Regresyon Analizi, 7. Uluslar Arası İstatistik Kongresi, Antalya.

Güneş, M. (2001). Bulanık Doğrusal Sistemler ve Regresyon Modellerine Uygulaması, *A Review of Social, Economic and Business Studies*, **1**: 176-192.

Isobe, T., Feigelson, E.D., Akritas, M.G. and Babu, G.J. (1990). Linear Regression in Astronomy I, *The Astrophysical Journal*, **364:** 104-113.

Kaufmann, A. , (1975). Introduction to the theory of Fuzzy Sub-sets, Academic Pres, 495-496

Kavruk, N. T. (2005). Doğrusal Regresyonda Sağlam Güven Aralıkları Robust Confidence Intervals In Linear Regression Analysis, Yüksek Lisans Tezi, Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, Ankara.

Kılıç İ., Doğan İ., Selvi P., Gazeloğlu C. ve Saraçlı S. (2011), Bulanık Basit Doğrusal Tip II Regresyon Analizi, 13. Biyoistatistik Kongresi, Ankara.

Maroco, J. (2007), Consistency and Efficiency of Ordinary Least Squares, Maximum Likelihood, and Three Type II Linear Regression Models, *Methodology*, **3(2):**81-88.

Mutan, O.C. (2004). Comparison of Regression Techniques Via Monte Carlo Simulation, Master of Science Thesis, Middle East Technical University, Institute of Science, Ankara.

Nevitt J. and Tam, H.P. (1998), A Comparison of robust and Nonparametric Estimators Under the Simple Linear Regression Model, M*ultiple Linear Regression Viewpoints*, **25**:54

Maronna, R.A., Martin, R.D., M. and Victor J. Y. (2006). Robust Statistics Theory and Methods, 14-17.

Nasrabadi, M.M. and Nasrabadi, E. (2004). A Mathematical-Programming Approach to Fuzzy Linear Regression Analysis, *Applied Mathematics and Computations* **155**: 873-881.

Paris, Q., (2012), The Dual of the Least-Squares Method, Department of Agricultural and Resource Economics University of California, Davis, Working Paper No. 12-001.

Rorusseeuw P.J and Leroy, A.M. (1986). Robust Regression and Outlier Detection, Fohn Wiley and Sons, Inc., Pulication, 75.

Saraçlı, S. (2008). Ölçüm Hatalı Modellerde Doğrusal Regresyon Tekniklerinin Karşılaştırılması –Monte-Carlo Simülasyon Çalışması-. Doktora Tezi, Eskişehir Osmangazi Üniversitesi, Fen Bilimleri Enstitüsü, Eskişehir.

Saraçlı S.,Gazeloğlu C. ve Altın Yavuz A. (2011), Robust Type II Sımple Regressıon, New Developments in Theory and Applications of Statistics An International Conference in Memory of Professor Moti Lal Tiku, Middle East Technical University, Ankara.

Saraçlı S., Yılmaz V. ve Doğan I., (2009), Simple Linear Regression Techniques in Measurement Error Models (Review), *Anadolu Üniversitesi Bilim ve Teknoloji Dergisi*, **10:** 335-342.

Saracli S., (2010) Tip II Regresyon Tekniklerinin Monte-Carlo Simülasyonu İle Karşılaştırılması", *E-Journal of New World Sciences Academy*, **6:** 26-35.

Saraçlı S. ve Çelik H.E., (2011) Performance of OLS-Bisector Regression in Method Comparison Studies, *World Applied Science Journal* **12(10):** 1860-1865.

Schaefer, K.C. and Visser, M.L. (2003). Reverse Regression and Orthogonal Regression in Employment Discrimination Analysis, *Journal of Forensic Economics*, **16(3),** 283–298.

Semiz, M. ve Genç A. (2003). Yığın Hacminin Tahmini İçin Bulanık Doğrusal Regresyon Modelinde Ters Tahmin Metodu, Selçuk Üniversitesi, Fen Edebiyat Fakültesi, *Fen Bilimleri Dergisi*, **22**: 65- 70, Konya.

Sümbüloğlu, K. ve Sümbüloğlu, V. (2002). Bioistatistik, Hatipoğlu Yayınları, Ankara.

Shapiro, F.A. (2005). Fuzzy regression models. Penn State University, **06**: 12

Şıklar, E. (2000). Regresyon Analizine Giriş, Anadolu Üniversitesi Fen-Edebiyat Fakültesi Yayınları, Eskişehir.

Tanaka, H. , Uejima, S. and Asai, K. , (1982). Linear regression analysis with fuzzy model, IEEE, Trans , *System Management, Sybernet*, **12(6):** 903 -907

Tekşen, Ü.M., (2008). Lineer Olmayan Bulanık Regresyonda Tahmin. Yüksek Lisans Tezi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya.

Türkay, H. (2004). Doğrusal Regresyon Modellerinin Robust (Dayanıklı) Yöntemlerle Tahmini ve Karşılaştırmalı Uygulamaları. Doktora Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.

Tiku, M.L. and Akkaya, A.D. (2004). Robust Estimation and Hypothesis Testing, New Age İnternational, New Delphi.

Vural, A., (2007). Aykırı Değerlerin Regresyon Modellerine Etkileri ve Sağlam Kestiriciler, Yüksek Lisan Tezi, Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.

Yanartaş, S. S. (2009). Bulanık Regresyonda Kullanılan Yöntemler. Yüksek Lisans Tezi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya.

Yorulmaz, Ö. (2009). Dayanıklı Regresyon Yöntemi ve Çeşitli Sosyal Veriler Üzerinde Aykırı Gözlemlerin Teşhisi, *Balıkesir Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, ,**21:**, 76-88

Yücel, L.İ. (2005). Bulanık Regresyon: Türkiye'de 1980-2004 Yılında Kayıt Dışı Ekonominin Bulanık Yöntemlerle Tahminine İlişkin Bir Uygulama. Yüksek Lisans Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.

Yıldırım, N. (2010). En Küçük Kareler, Ridge Regresyon ve Robust Regresyon Yöntemlerinde Analiz Sonuçlarına Aykırı Değerlerin Etkilerinin Belirlenmesi. Yüksek Lisans Tezi, Çukurova Üniversitesi, Fen Bilimleri Enstitüsü, Zootekni Anabilim Dalı, Adana.

Yurtcu, Ş. ve İçağa, Y. (2007) Bulanık Doğrusal Regresyona Genel Bir Bakış, *Yapı Teknolojileri Elektronik Dergisi*, **2:** 37-43.

Zadeh, L. A. , (1965). Fuzzy sets information and control, *Academic Press*, **8:** 338-353

ÖZGEÇMİŞ

Adı Soyadı	: Cengiz Gazeloğlu
Doğum Yeri ve Tarihi	: Delice/ 01.04.1988
Yabancı Dili	: İngilizce
İletişim (Telefon/e-posta)	: 0544 717 07 88/ cengiz_gazeloglu@hotmail.com

Eğitim Durumu (Kurum ve Yıl)

Lise	: Keçiören Kanuni Lisesi (2001-2004)
Lisans	: Afyon Kocatepe Üniversitesi (2006-2010)