



**CUMHURİYET ÜNİVERSİTESİ**

**Sosyal Bilimler Enstitüsü**

**Ekonometri Anabilim Dalı**

**KANTİL REGRESYON ANALİZİNDE BOOTSTRAP TAHMİNİ**

**Yüksek Lisans Tezi**

**Seçkin ÇAMURLU**

**Sivas**

**Nisan-2018**

**CUMHURİYET ÜNİVERSİTESİ**

**Sosyal Bilimler Enstitüsü**

**Ekonometri Anabilim Dalı**

**KANTİL REGRESYON ANALİZİNDE BOOTSTRAP TAHMİNİ**

**Yüksek Lisans Tezi**

**Seçkin ÇAMURLU**

**Tez Danışmanı**

**Doç. Dr. Necati Alp ERİLLİ**

**Sivas**

**Nisan-2018**

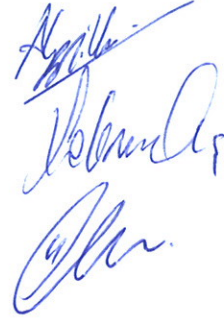
## KABUL VE ONAY

Üniversite: : Cumhuriyet Üniversitesi  
Enstitü : Sosyal Bilimler Enstitüsü  
Ana Bilim Dalı : Ekonometri  
Bilim Dalı : Ekonometri  
Tezin Başlığı : **Kantil Regresyon Analizinde Bootstrap Tahmin**  
Savunma Tarihi : 13/04/2018  
Danışmanı : Doç. Dr. Necati Alp ERİLLİ

Unvanı - Adı Soyadı

İmza

Jüri Başkanı : Doç. Dr. Necati Alp ERİLLİ  
Üye : Dr. Öğr. Ü. Şebnem ZORLUTUNA  
Üye : Dr. Öğr. Ü. Özge GÜNDOĞDU



Oy Birliği

Oy Çokluğu

Seçkin ÇAMURLU tarafından hazırlanan “**Kantil Regresyon Analizinde Bootstrap Tahmin**” başlıklı tez, kabul edilmiştir. ..../..../.....

Prof. Dr. Ahmet ŞENGÖNÜL  
Enstitü Müdürü

## ETİK İLKELERE UYGUNLUK BEYANI

Cumhuriyet Üniversitesi Sosyal Bilimler Enstitüsü bünyesinde hazırladığım bu Yüksek Lisans/Doktora/Sanatta Yeterlik tezinin bizzat tarafımdan ve kendi sözcüklerimle yazılmış orijinal bir çalışma olduğunu ve bu tezde;

- 1- Çeşitli yazarların çalışmalarından faydalandığımda bu çalışmaların ilgili bölümlerini doğru ve net biçimde göstererek yazarlara açık biçimde atıfta bulunduğumu;
- 2- Yazdığım metinlerin tamamı ya da sadece bir kısmı, daha önce herhangi bir yerde yayımlanmışsa bunu da açıkça ifade ederek gösterdiğimi;
- 3- Başkalarına ait alıntılanan tüm verileri (tablo, grafik, şekil vb. de dâhil olmak üzere) atıflarla belirttiğimi;
- 4- Başka yazarların kendi kelimeleriyle alıntıladığım metinlerini, tırnak içerisinde veya farklı dizerek verdiğim yine başka yazarlara ait olup fakat kendi sözcüklerimle ifade ettiğim hususları da istisnasız olarak kaynak göstererek belirttiğimi,

beyan ve bu etik ilkeleri ihlal etmiş olmam halinde bütün sonuçlarına katlanacağımı kabul ederim.

İmza

Seçkin ÇAMURLU

## TEŐEKKÜR

Çalıőmamın, baőlangıcından bitimine kadar her safhada yakın ilgi ve yardımlarını esirgemeyen, yapıcı ve öđretici eleőtirileriyle bana yol gösteren danıőman hocam Sayın Doç. Dr. Necati Alp ERİLLİ' ye,

Yönlendirmeleri, eleőtirileri ve önerileri ile yardımını gördüğüm yakın arkadaşım Betül'e,

Bugünlere gelmemde en büyük emeđi olan, en zor anlarımda desteklerini, sevgi ve hoşgörülerini esirgemeyen sevgili aileme,

Sonsuz teőekkürler...



# İÇİNDEKİLER

KABUL VE ONAY .....	i
ETİK İLKELERE UYGUNLUK BEYANI .....	ii
TEŞEKKÜR.....	iii
İÇİNDEKİLER.....	iv
KISALTMALAR.....	vi
TABLolar LİSTESİ.....	vii
ŞEKİL LİSTESİ .....	ix
ÖZET .....	x
ABSTRACT .....	xi
GİRİŞ.....	1
<b>1. LİTERATÜR TARAMASI .....</b>	<b>2</b>
1.1. Bootstrap Literatür Taraması.....	2
1.2. Kantil Regresyon Literatür Taraması .....	5
<b>2. REGRESYON ANALİZİ VE TÜRLERİ .....</b>	<b>8</b>
2.1. Parametrik ve Parametrik Olmayan Regresyon Analizi.....	9
2.2. Parametrik Regresyon Analizi.....	10
2.2.1. En Küçük Kareler Yöntemi.....	11
2.2.2. En Çok Olabilirlik Yöntemi .....	16
2.3. Parametrik Olmayan Regresyon Analizi .....	16
2.3.1. Düzgünleştirme Kavramı .....	18
2.4. Parametrik Regresyon ile Parametrik Olmayan Regresyon Analiz Yöntemleri Karşılaştırılması.....	18
<b>3. KANTİL REGRESYON .....</b>	<b>20</b>
3.1. Kantil Kavramı .....	20
3.2. Anakitlenin Modellenmesi .....	21
3.2.1. Kümülatif Dağılım Fonksiyonu .....	21
3.2.2. Olasılık Yoğunluk Fonksiyonu .....	22
3.2.3. Kantil Fonksiyon .....	23
3.2.4. Kantil Yoğunluk Fonksiyonu .....	24
3.3. Kantil Regresyon Analizi .....	25
3.3.1. Kantil Regresyon Yönteminin Özellikleri.....	28
3.3.2. Kantil Regresyonun Doğrusal Programlama Gösterimi.....	28
3.3.3. Kantil Regresyon Parametre Tahminleri.....	30

3.3.4.	Belirlilik Katsayısı ve Düzeltilmiş Belirlilik Katsayısı.....	31
3.3.5.	Eşit Varyans Özellikleri .....	32
3.3.6.	Güven Aralıkları.....	32
3.3.7.	Kovaryans ve Parametre Tahminlerinin Korelasyonu .....	33
<b>4.</b>	<b>BOOTSTRAP YÖNTEMİ VE TÜRLERİ.....</b>	<b>34</b>
4.1.	Bootstrap Yöntemi.....	34
4.1.1.	Parametrik Bootstrap Yöntemi.....	37
4.1.2.	Parametrik Olmayan Bootstrap Yöntemi .....	39
4.2.	Regresyon Analizinde Kullanılan Bootstrap Yöntemleri.....	40
4.2.1.	Hata Terimlerine Dayanan Bootstrap Yöntemi.....	41
4.2.2.	Çift Bootstrap Yöntemi .....	41
<b>5.</b>	<b>UYGULAMA.....</b>	<b>43</b>
<b>6.</b>	<b>SONUÇ VE ÖNERİLER.....</b>	<b>59</b>
<b>KAYNAKÇA .....</b>	<b>64</b>	
<b>ÖZGEÇMİŞ.....</b>	<b>71</b>	

## KISALTMALAR

<b>EKK:</b>	En Küçük Kareler
<b>CDF:</b>	Kümülatif Dağılım Fonksiyonu
<b>PDF:</b>	Olasılık Yoğunluk Fonksiyonu
<b>QDF:</b>	Kantil Yoğunluk Fonksiyonu
<b>LAD:</b>	En Küçük Mutlak Sapmalar
<b>OMS:</b>	Ortalama Mutlak Sapma
<b>HKOK:</b>	Hata Kareler Ortalamasının Karekökü
<b>ÜFE:</b>	Üretici Fiyat Endeksi
<b>GMM:</b>	Genelleştirilmiş Momentler Yöntemi

# TABLolar LİSTESİ

<b>Tablo No</b>	<b>Tablo Adı</b>	<b>Sayfa</b>
<b>Tablo 2. 1.</b>	<b>Parametrik Olmayan Regresyonda Değişken Sayısı Artıkça Gerekli Örnek Büyüklüğü</b>	<b>17</b>
<b>Tablo 4. 1.</b>	<b>Bootstrap Örneklemi</b>	<b>37</b>
<b>Tablo 5. 1.</b>	<b>Doğrusal Regresyon Modeli Analiz Sonuçları</b>	<b>43</b>
<b>Tablo 5. 2.</b>	<b>Doğrusal Regresyon Modeli Analiz Sonuçları</b>	<b>44</b>
<b>Tablo 5. 3.</b>	<b>Kantil Regresyon(Q<sub>1</sub>) Modeli Analiz Sonuçları</b>	<b>44</b>
<b>Tablo 5. 4.</b>	<b>Kantil Regresyon(Q<sub>2</sub>) Modeli Analiz Sonuçları</b>	<b>45</b>
<b>Tablo 5. 5.</b>	<b>Kantil Regresyon(Q<sub>3</sub>) Modeli Analiz Sonuçları</b>	<b>45</b>
<b>Tablo 5. 6.</b>	<b>Doğrusal Regresyon ve Kantil Regresyon Modeli OMS ve HKOK Analiz Sonuçları</b>	<b>46</b>
<b>Tablo 5. 7.</b>	<b>Bootstrap(%25) - Doğrusal Regresyon Modeli Analiz Sonuçları</b>	<b>46</b>
<b>Tablo 5. 8.</b>	<b>Bootstrap(%25) - Kantil Regresyon(Q<sub>1</sub>) Modeli Analiz Sonuçları</b>	<b>47</b>
<b>Tablo 5. 9.</b>	<b>Bootstrap(%25) - Kantil Regresyon(Q<sub>2</sub>) Modeli Analiz Sonuçları</b>	<b>47</b>
<b>Tablo 5. 10.</b>	<b>Bootstrap(%25) - Kantil Regresyon(Q<sub>3</sub>) Modeli Analiz Sonuçları</b>	<b>48</b>
<b>Tablo 5. 11.</b>	<b>Bootstrap(%25) - Doğrusal Regresyon ve Kantil Regresyon Modeli OMS ve HKOK Analiz Sonuçları</b>	<b>48</b>
<b>Tablo 5. 12.</b>	<b>Bootstrap(%50) - Doğrusal Regresyon Modeli Analiz Sonuçları</b>	<b>49</b>
<b>Tablo 5. 13.</b>	<b>Bootstrap(%50) - Kantil Regresyon(Q<sub>1</sub>) Modeli Analiz Sonuçları</b>	<b>49</b>
<b>Tablo 5. 14.</b>	<b>Bootstrap(%50) - Kantil Regresyon(Q<sub>2</sub>) Modeli Analiz Sonuçları</b>	<b>50</b>
<b>Tablo 5. 15.</b>	<b>Bootstrap(%50) - Kantil Regresyon(Q<sub>3</sub>) Modeli Analiz Sonuçları</b>	<b>50</b>
<b>Tablo 5. 16.</b>	<b>Bootstrap(%50) - Doğrusal Regresyon ve Kantil Regresyon Modeli OMS ve HKOK Analiz Sonuçları</b>	<b>51</b>
<b>Tablo 5. 17.</b>	<b>Bootstrap(%75) - Doğrusal Regresyon Modeli Analiz Sonuçları</b>	<b>51</b>
<b>Tablo 5. 18.</b>	<b>Bootstrap(%75) - Kantil Regresyon(Q<sub>1</sub>) Modeli Analiz Sonuçları</b>	<b>52</b>
<b>Tablo 5. 19.</b>	<b>Bootstrap(%75) - Kantil Regresyon(Q<sub>2</sub>) Modeli Analiz Sonuçları</b>	<b>52</b>
<b>Tablo 5. 20.</b>	<b>Bootstrap(%75) - Kantil Regresyon(Q<sub>3</sub>) Modeli Analiz Sonuçları</b>	<b>53</b>

<b>Tablo 5. 21.</b> Bootstrap(%75) - Doğrusal Regresyon ve Kantil Regresyon Modeli OMS ve HKOK Analiz Sonuçları	53
<b>Tablo 5. 22.</b> Bootstrap(%100) - Doğrusal Regresyon Modeli Analiz Sonuçları	54
<b>Tablo 5. 23.</b> Bootstrap(%100) - Kantil Regresyon( $Q_1$ ) Modeli Analiz Sonuçları	54
<b>Tablo 5. 24.</b> Bootstrap(%100) - Kantil Regresyon( $Q_2$ ) Modeli Analiz Sonuçları	55
<b>Tablo 5. 25.</b> Bootstrap(%100) - Kantil Regresyon( $Q_3$ ) Modeli Analiz Sonuçları	55
<b>Tablo 5. 26.</b> Bootstrap(%100) - Doğrusal Regresyon ve Kantil Regresyon Modeli OMS ve HKOK Analiz Sonuçları	56
<b>Tablo 5. 27.</b> Bootstrap(%200) - Doğrusal Regresyon Modeli Analiz Sonuçları	56
<b>Tablo 5. 28.</b> Bootstrap(%200) - Kantil Regresyon( $Q_1$ ) Modeli Analiz Sonuçları	57
<b>Tablo 5. 29.</b> Bootstrap(%200) - Kantil Regresyon( $Q_2$ ) Modeli Analiz Sonuçları	57
<b>Tablo 5. 30.</b> Bootstrap(%200) - Kantil Regresyon( $Q_3$ ) Modeli Analiz Sonuçları	58
<b>Tablo 5. 31.</b> Bootstrap(%200) - Doğrusal Regresyon ve Kantil Regresyon Modeli OMS ve HKOK Analiz Sonuçları	58

## ŞEKİL LİSTESİ

Şekil No	Şekil Adı	Sayfa
Şekil 3. 1	Kümülatif Dağılım Fonksiyonu	22
Şekil 3. 2	Olasılık Yoğunluk Fonksiyonu	23
Şekil 3. 3	Kantil Fonksiyonu	23
Şekil 3. 4	Kantil Regresyon Amaç Fonksiyonu	30
Şekil 4. 1	Bilinmeyen Parametrenin Dağılımına İlişkin Tahminin Bootstrap Süreci	36



## ÖZET

Regresyon analizi, istatistiksel ve ekonometrik tahmin modellerinde sıklıkla kullanılan tahmin yöntemlerinden biridir. Regresyon analizinde yardımcı analiz yöntemlerden biri de kantil regresyon yöntemidir. Kantil regresyon yönteminde herhangi bir dağılım varsayımı gerekmemektedir ve çeşitli kantillere bağlı olarak parametre katsayılarını tahmin ettiği için aşırı değerlerin bulunduğu veri setlerinde geleneksel yöntemlere göre daha iyi tahminler vermektedir. Ayrıca kantil regresyon değişen varyansın belirlenmesine imkân sağlamaktadır. Doğrusal regresyon analizinde ise veri yapısının model için uygun olması gibi şartlar vardır. Bunlardan biri veri yapısındaki aşırı değerlere karşı iyi sonuçlar vermemesidir.

Bu çalışmada; Doğrusal regresyon ve Kantil regresyon yöntemleri tanıtılmış ve aralarındaki farklar belirtilmiştir. Bootstrap yöntemi hakkında bilgiler verilmiştir. Uygulama kısmında ise 2000-2017 yılları arası aylık Üretici Fiyat Endeksi(Üfe), Üfe(-2) Dönem Gecikmesi ve Beklenti Anketi verileri kullanılmıştır. Bu veriler Gretl programı yardımı ile Bootstrap yöntemi kullanılarak belirli düzeylerde veri sayıları artırılarak Doğrusal ve Kantil Regresyon yöntemlerinin sonuçları Ortalama Mutlak Sapma (OMS) ve Hata Kareler Ortalaması Karekökü (HKOK) değerleri karşılaştırılarak hangi yöntemin en uygun modeli tahmin ettiği belirlenmiştir. Sonuç olarak OMS ve HKOK değerleri karşılaştırıldığında en küçük değeri Doğrusal Regresyon yöntemi vererek en uygun model seçilmiştir. İkinci model olarak ise Kantil Regresyon( $Q_2$ ) yönteminin uygun model olduğu gösterilmiştir.

**Anahtar Kelimeler:** Doğrusal Regresyon, Kantil Regresyon Karşılaştırma, Bootstrap Tahmini, HKOK

## ABSTRACT

Regression analysis is one of the most commonly used estimation methods in statistical and econometric prediction models. One of the auxiliary analysis methods in regression analysis is the quantile regression method. There is no distributional assumption is required in the quantile regression model and it gives better estimates in the data sets of the structure where the outliers are estimated because it predicts the parameter coefficients depending on the various quantities. In addition, the quantile regression allows for the determination of the variance. In linear regression analysis, there are requirements such that data structure is suitable for the model. It does not provide good results against extreme values in data structure.

In this study; firstly linear regression and quantile regression methods are introduced, and than the differences between them are indicated. Information about the bootstrap method is also given. In the application part, monthly Producer Price Index (PPI), PPI (-2) period delay and Expectation Questionnaire data between 2000-2017 were used. This data was used to increase the number of data at certain levels by Bootstrap method and to compare the results of Linear and Quantile Regression methods with Mean Absolute Deviation (MAD) and Root Means Square of Error (RMSE) to determine which method predicts the most suitable model. Results for Linear and Quantile Regression ( $Q_2$ ) methods show that the two methods with the closest and smallest values of MAD and RMSE predict the most suitable models.

**Keywords:** Linear Regression, Quantile Regression Comparison, Bootstrap Estimation, RMSE

# GİRİŞ

Ekonometrik çalışmalarda en çok kullanılan araçlarından biri olan regresyon analizi, genellikle aklımıza ilk olarak parametrik regresyon modellerini getirir. Parametrik modellerde bağımlı ve bağımsız değişken arasındaki ilişkiyi açıklamak için, uygulanacak olan parametrik regresyon modelinin fonksiyonel şeklinin veri yapısına uygun olması, gerekli olan değişkenlerin modelde bulunuyor olması gibi bazı şartlar yer almaktadır. Bu modellerde aranan şartlar sağlanmış olsa bile değişkenler arasındaki ilişkiyi açıklayabilecek farklı modellerin olup olmayacağı her zaman bir araştırma konusu olmuştur. Hatta bazı durumlarda modellerin tahmin edeceği verinin yapısı, uç değerlerinin yer alması, serinin normal dağılmaması gibi sebeplerle parametrik modellerle iyi sonuçlar elde edilmemektedir. Bu yüzden alternatif modeller gerekmektedir.

Koenker ve Bassett (1978) tarafından geliştirilen kantil regresyon bu durumlarda tercih edilebilecek olan regresyon modellerinden birisi olarak yer almaktadır. Kantil regresyonda herhangi bir dağılım varsayımı gerektirmemektedir. Ayrıca kantillere bağlı olarak regresyon katsayılarını tahmin ettiği için aşırı değerlerin bulunduğu yapıdaki veri setlerinde daha iyi tahminler verdiği bilinmektedir (Koenker ve Bassett 1978; Leng ve Tong 2013). Ayrıca kantil regresyon değişen varyansın belirlenmesine yardımcı olmaktadır.

Bu çalışma ile veri yapısı değiştirilmeden bootstrap yöntemi kullanılıp yeni veriler türetilerek Doğrusal Regresyon ve Kantil Regresyon yöntemi analiz sonuçları karşılaştırılmış ve hangi yöntemin daha iyi sonuçlar verdiği incelenmiştir.

Çalışmamızın ilk bölümünde kantil regresyon ve bootstrap yöntemleri hakkında literatür taraması yapılmıştır. İkinci, üçüncü ve dördüncü bölümde ise çalışmamızda kullanacağımız yöntemler hakkında teorik bir çerçeve çizilmiştir. Beşinci bölümde üfe, üfe(-2) ve beklenti anketi verileri üzerine doğrusal regresyon ve kantil regresyon modeli tahmin sonuçları sunulmuştur. Altıncı ve son bölümde ise bulgularımız neticesinde elde ettiğimiz sonuçlar ve öneriler yer almaktadır.

# BÖLÜM 1

## 1. LİTERATÜR TARAMASI

### 1.1. Bootstrap Literatür Taraması

**Efron(1979)**, Standart hata tahmini için jackknife ve delta yöntemi gibi eskiden kabul edilen istatistiksel yöntemleri, parametrik olmayan bootstrap yöntemiyle birleştirdi.

**Bickel ile Freedman (1981) ve Singh (1981)**, Bootstrap yönteminin birtakım matematiksel koşul altında tutarlılığını gösteren ilk sonuçlarını ileri sürdüler.

**Efron(1982)**, Bootstrap tahmininin asimptotik tutarlılığı ve bazı karşı örnekler üzerine birçok teorik gelişmeler ortaya çıkartmıştır.

**Staudte ve Sheather (1990)**, Tahminlerin standart hatalarını tahmin etmek için bootstrap yöntemini kullandılar. Hipotez testi ile ilgilenmelerine rağmen, hipotez testi problemleri için bootstrap tekniğini kullanmamışlardır.

**McLachlan ve Krishnan (1997)**, Bootstrap tekniğini kovaryans matrisinin dayanıklı tahmin aracı olarak ele aldılar. **Beran ve Srivastaya (1985)**, Kovaryans matrisi fonksiyonları için bootstrap testlerini buldu. **Lahiri(1992)**, M-estimator (Dayanıklı Konum Tahmincisi) için bootstrap tekniğini kullandı.

**Hahn ve Meeker (1991)**, Bootstrap güven aralıklarını ele aldılar. **Frangos ve Schucany(1990)**,  $BC_a$  güven aralığı yöntemi için Efron akselerasyon sabiti tahmininin teknik görüşünü ele aldılar. **Krieger ve Bickel (1989)**, Dağılım fonksiyonunda güven bantlarını elde etmek için bootstrap tekniğini kullandılar.

**Efron ve Gong(1983)**, Populasyon dağılımının normal ve negatif üstel olduğu durumlar için ortalamanın standart sapmanın bootstrap ve jackknife tahminleri elde edildiğini açıklamıştır.

**Wu(1986)**, Eşit varyanslılıkta  $\hat{\beta}'$  nın bootstrap varyans tahmini sapmasız olurken, bu yöntemin heterojen varyanslılıkta sapmalı ve tutarsız sonuç verdiği heterojen varyans durumunda tahmin yapmaya uygun olmadığını açıklamıştır.

**Stine(1985)**, Doğrusal regresyon modelinde, hata değerlerinin dağılımı için sapma  $\varepsilon = \ell_i - \bar{\ell}$  değerine  $1/n$  olasılığını vererek deneysel dağılımın tahmin edilebileceğini göstermiştir.

**Efron ve Tibshirani(1986)**, Standart sapmanın bootstrap tahminini ele almışlar; parametre tahmininin de hangi tahmincinin kullanılacağını; bu tahmincinin parametre için ne kadar doğru bilgi taşıdığı sorularından bootstrap tekniği ikinci soruya cevap bulmak için geliştirilmiş bir yöntem olduğunu açıklamıştır.

**Diciccioue ve Tibshirani (1987)**, Bootstrap güven aralıkları ile varyans ve korelasyon katsayısı için ilgili güven aralığı yöntemlerini kullanıp güven aralıkları elde etmişlerdir.

**Efron (1990)**, Bootstrap tekrarlama sayısını azaltma amacına yönelik daha etkili bootstrap hesaplamaları yapmış ve tekrarlama büyüklüğünün 50 ile 200 arasında olmasının yeterli olduğunu açıklamıştır.

**Money(1993)**, Bootstrap yönteminin örnekleme yöntemlerinden farklı bir yöntem olduğunu belirtmiştir. Bootstrap tekniğinin örnekleme dağılımını oluşturmak için güçlü varsayımlar ve analitik formüller yerine büyük sayıda yinelenmiş hesaplamalar içerdiğini ve bootstrap tekniğinin örnekleme yöntemlerinin yetersizliklerini ortadan kaldırmak, tahmin edicinin örnekleme dağılımının oluşturulmasını sağlamak hem de sapmanın azaltılması amacına yönelik parametrik olmayan bir yöntem olarak geliştirdiğini açıklamıştır.

**Shao(1996)**, Regresyon analizinde yöntemin kullanılması sırasında yeniden örnekleme sürecinin, analiz öncesinde gözlemlerin yeniden örneklenmesi veya analiz sonrasında hataların yeniden örneklenmesi biçiminde olacağını açıklamıştır.

**Jeremy(1996)**, Yeniden örnekleme yöntemlerinin tanımlı ve bağımsız veriler için kullanılabilirliği gibi zaman serilerinde de uygulanabileceğini bildirmiştir.

**Strawderman ve Wells(1997)**, Hazard (Birikimli Tehlike) ve yaşam fonksiyonları için bootstrap güven sınırlarının geçerliliği üzerine çalışmışlardır. Aynı zamanda rastgele olarak azalan veri için bootstrap yönteminin yararlı olacağını belirtmişlerdir.

**Zeng ve Davidan(1997)**, Bağışıklık sisteminde bootstrap güven aralıklarının kullanımı üzerinde durmuşlardır.

**Fox(1997)**, Bootstrap yönteminin hata terimlerinden elde edilen tekrarlı örnekleri seçip tahmin değerlerine ( $\hat{Y}_i$ ) eklenmesinden dolayı hata terimlerinin normal dağılıma sahip olduğunu varsaymıştır.

**Atakan(2003)**, Hata oranı tahmin edicilerinin jackknife ve bootstrap değerlendirmesi çalışmasında, hem gerçek hata oranı için ve hem de bu hata oranına ilişkin tahmin ediciler için tüm durumlarda genelde jackknife tekniğinden elde edilen sonuçlar, bootstrap tekniğini sonuçları bakımından gerçek değerlere daha yakın olduğu görülmektedir. Buradan jackknife yönteminin, bootstrap yöntemine tercih edilebilir olduğunu belirtmiştir.

**Topuz(2002)**, Klasik Örnekleme ve Yeniden Örnekleme(Bootstrap ve Jackknife) yöntemi sonuçlarını karşılaştırmıştır. Bootstrap yöntemi ile yapılan tahminlerin standart hataları klasik örnekleme yöntemi ile yapılan tahminlerin standart hatalarından daha küçük tahminler elde etmiştir. Bu şekilde jackknife ve bootstrap yöntemi sonuçları da karşılaştırılmış ve jackknife yönteminin, bootstrap yönteminden daha büyük standart hatalı tahminler vermiştir. Jackknife yöntemi de bootstrap yöntemi gibi parametrik varsayımlar kullanmak yerine verideki değişkenliği kullanmaya yönelik yöntemler olmasına rağmen bootstrap yönteminin jackknife yönteminden daha az hatalı parametre tahminleri yaptığı belirtilmiştir.

**Okutan(2009)**, Bootstrap tekniğini bazı özel dağılımlarının doğrusal regresyon modeli için uyarlanmış, En Çok Olabilirlik ve Klasik En Küçük Kareler tahmin edicisinin varyansını tahmin etmede kullanmıştır ve sonuçlara göre regresyonda bootstrap yönteminin kullanımının daha gerçek sonuçlar vereceğini belirtmiştir.

**Özel ve Sezgin(2011)**, Veri sayısının az oluşundan ve sağlam (robust) tahminler üretmek amacıyla çalışmalarında bootstrap tekniğini kullanmayı belirtmişlerdir.

## 1.2. Kantil Regresyon Literatür Taraması

**Koenker (2005)**, Kantil regresyonun ilk olarak regresyondaki klasik varsayımlardan bir olan hata terimlerinin normal dağılması varsayımını ihlal eden robust (sağlam) bir regresyon yöntemi olduğunu ve daha kapsamlı bir regresyon görüntüsü oluşturmak amacıyla tasarlanan bir yöntem olduğunu belirtmiştir.

**Chen ve Wei (2005)**, Kantil regresyonun özellikle koşullu kantillerin değişkenlik gösterdiği durumlarda kullanışlı olduğunu belirtmişlerdir.

**Koenker (2005)**, Kantil regresyonun özellikleri, asimptotik teori özellikleri, ağırlıklandırılmış kantil regresyon özellikleri, kantil regresyonun hesaplama yöntemleri ve belirsizlik durumundan son olarak da kantil regresyona ait sonuç çıkarımından bahsetmiştir. Kantil değerini değişkenin dağılımında bulunan ve bu dağılımı kendisinden büyük değerler ve küçük değerler olmak üzere ikiye bölen bir değer olduğunu belirtmiştir.

**Koenker ve Xiao (2002)**, Kantil regresyon, rassal değişkenler arasında stokastik ilişkilerin çok daha eksiksiz istatistiksel analizini yapabildiğini belirtmişlerdir.

**Baur, Saisana ve Niels (2004)**, En Küçük Kareler Yönteminde hata teriminin, değişkenlerin değerinden bağımsız olduğu (varyanslar homojen) varsayılmaktadır. Kantil regresyon modelindeyse, hata terimlerinin değişkenliğine izin verilmekte ve varyans yapısına ilişkin herhangi bir varsayımın bulunmadığı belirtmişlerdir.

**Hao ve Naiman (2007)**, Kantil regresyonda bağımsız değişkenlerin, bağımlı değişkenin dağılımını nasıl etkilediği yönünde bazı önemli bilgiler verdiği için sosyal bilimlerde oldukça geniş bir kullanım alanının olduğunu belirtmiştir.

**Gill (2002)**, Kantil regresyon tahminlerinin gözlem değerlerinin büyüklüğü yerine gözlem değerlerinin işaretlerine dayalı olması kantil regresyonunun robust bir yöntem olmasını sağladığını belirtmiştir.

**Akar (2013)**, Kantil regresyon yöntemini kullanarak geliştirmekte olan 8 ülkenin finansal piyasa istikrarları için oluşturulan modellerdeki eğim parametrelerinin çeşitli kantiller boyunca sabit kalıp kalmadığını ve bu eğim parametrelerinin ülkelere göre kantiller boyunca değişimini %95 güven aralığında incelemiştir.

**Ciner, Gurdgiev ve Lucey (2013)**, 1990-2010 dönemini içeren ABD ve İngiltere'ye ait günlük verileri kullanarak döviz kuru, petrol fiyatı, altın fiyatı, tahvil ve hisse senedi

arasındaki ilişkileri kantil regresyon yöntemi ile incelemişler. Belirlenen bu değişkenlerin birbirlerine karşı korunma oluşturduğunu ve altının döviz kuruna karşı güvenli bir liman olduğunu tespit etmişlerdir.

**Chen (2005)**, Özellikle uç noktaların önemli olduğu durumlarda kantil regresyon yönteminin kullanışlı olduğundan bahsetmiş ve halk sağlığı açısından çevresel çalışmalarda üst kantillerde hava kirliliğinin kritik seviyeye geldiği durumları örnek olarak göstermiştir.

**Kurtoğlu (2011)**, Kantil regresyon yönteminin özel bir hali olan En Küçük Mutlak Sapma (LAD) yöntemini ele almış ve bu yöntemlerle elde edilen sonuçları En Küçük Kareler yöntemi ile karşılaştırmıştır.

**Güriş, Çağlayan ve Saçıldı (2010)**, Kantil regresyonun hataların mutlak değerleri toplamını minimize edilerek tahmin edildiğini ve katsayılar vektörü için standart hatalar bootstrap yöntemi kullanılarak elde edilebileceğini belirtmişlerdir.

**Çelik ve Selim (2013)**, Çalışmalarında Temel İnsan Sermayesi Modeli ve Genişletilmiş İnsan Sermayesi Modeli Sonuçlarını En Küçük Kareler yöntemine ek olarak Kantil regresyon yöntemini karşılaştırmalı olarak ele almışlardır. Çalışmadan elde edilen sonuçlara göre Türkiye’de yerleşim yerleri bakımından ücret farklılığının var olduğu söylemişlerdir.

**Saçaklı (2005)**, En Küçük Kareler Regresyonu, En Küçük Mutlak Sapma Regresyonu, En Küçük Medyan Kare Regresyonu, M regresyon ve Rank Regresyonundan bahsedip bu alternatif regresyon modellerini karşılaştırmıştır. Ayrıca, kantil regresyonun özelliklerini vermiş ve kantil regresyon analizine yönelik bir uygulama yapmıştır.

**Lingren (1997)**, Parametrik kantil fonksiyon tahminleri için yeni bir yöntem önermiştir. Kantil fonksiyonunun tahmini, simetrik olmayan  $L_1$  teknikleri ile birlikte sansür limitleri dağılımının ağırlıklı Kaplan-Meier tahmini olarak ele almıştır.

**Eide ve Showalter (1998)**, “Test skoru kazanımı” durumunun koşullu dağılımında farklı noktalardaki okul kalitesi ve performans arasındaki ilişkinin farklı olup olmadığını standartlaştırılmış testlerde kantil regresyon yöntemini kullanarak tahmin etmişlerdir.

**Kan ve Tsai (2004)**, Obeziteye neden olan değişkenlerin etkilerini incelemişlerdir. Taiwan’den alınan anket verileri temel alınmış ve Kantil Regresyon tekniği kullanılarak analiz edilmiştir. Sonuçlar ilişkinin var olduğunu, kadın ve erkekler arasında Beden Kitle Endeksi (BMI) açısından farklı ilişki olduğunu göstermişlerdir.

**Stifel ve Averett (2009)**, ABD’de aşırı kilolu (obez) çocukların yaygınlığının son yirmi yılda çarpıcı biçimde arttığı ve bunun halk sağlığı problemlerine neden olduğu, üstelik fazla kilolu olmayan çocukların da ağırlığının artması problemini ele almışlardır. Bunun için Beden Kitle Endeksi (BMI) ile ölçülen ağırlık durumu ve aşırı kilo arasındaki ilişkiyi bulmak için kantil regresyon yöntemini kullanmışlardır. Sonuç olarak kantil regresyon yönteminin daha detaylı bilgi verdiğini göstermişlerdir.

**Meligkotsidou ve ark. (2009)**, Risk faktörlerini kullanarak yatırım fonu getirilerinin koşullu kantillerle modellenmesi fikrini tanıtmışlardır. Kantil regresyon analizi, yatırım fonu getirisi ve risk faktörlerinin değişimi arasındaki ilişkinin koşullu getiri dağılımı boyunca nasıl değiştiğini anlamak için bir yol sunmuşlardır. En uygun risk faktörleri, farklı kantiller için tanımlanmış ve koşullu beklenen modelden elde edilenle karşılaştırılmıştır. Getirilerdeki faktör etkisindeki farklılığı kantiller ile elde etmişlerdir.

## BÖLÜM 2

### 2. REGRESYON ANALİZİ VE TÜRLERİ

İstatistiksel tahmin çalışmalarında en çok kullanılan yöntemlerden biri regresyon analizi yöntemidir. Regresyon analizi; bir bağımlı değişkenin bir veya birden fazla bağımsız değişkenle arasındaki ilişkinin matematiksel bir fonksiyon şeklinde yazılması olarak tanımlanabilir. Bu fonksiyona regresyon denklemi adı verilmektedir (Orhunbilge, 2000).

Regresyon analizi ile ilgili ilk çalışmalar; Francis Galton tarafından 19. Yüzyılın sonlarında yapılmıştır. Galton çalışmasında; anne-babaların boyu ile çocuklarının boyları arasındaki ilişkiyi incelemiş ve kısa boylu anne-babaların çocuklarının boylarının kısa, uzun boylu anne-babaların çocuklarının boylarının uzun olmasına rağmen, çocuklarının boylarının anakitle boy ortalamasına doğru yaklaşma eğiliminde olduğunu görmüştür. Bu eğilimi “ortaya doğru çekilme = regression to mediocrity” olarak adlandırmıştır. Galton’un çalışmaları bugün değişkenler arasındaki istatistik ilişkileri inceleyen Regresyon Analizi ’nin başlangıcı olmuştur (Galton, 1886).

$X$  bağımsız ve  $Y$  bağımlı değişkenini göstermek üzere; bu değişkenler arasında  $Y = f(X)$  biçiminde bir ilişki tanımlanıyorsa bu fonksiyonel bir ilişkidir. İstatistiki ilişki ise, fonksiyonel ilişkiden farklı olarak belli bir hata payı içermektedir. Diğer bir ifade ile gözlemler her zaman elde edilen doğru ya da eğrinin üzerine düşmez. Regresyon analizi, korelasyon analizi ile birlikte istatistiki ilişkinin araştırılmasında kullanılan en önemli analizdir. Regresyon analizinin amaçları(Maddala, 2001);

- Bağımsız değişken(ler)in bağımlı değişken üzerinde önemli bir etkiye sahip olup olmadığını incelemek
- Bağımlı değişkenin gelecekte alabileceği değerleri tahmin etmek
- Bağımsız değişkenlerdeki değişimin tek başına bağımlı değişken üzerinde yaptığı etkiyi analiz etmektir.

$$Y = \beta_0 + \beta_1 x + \varepsilon \quad (2.1)$$

eşitlik 2.1’de  $x$  bağımsız(açıklayıcı) değişkeni,  $y$  bağımlı değişkeni ve  $\varepsilon$  hata terimini (stokastik değişken) göstermektedir.  $\varepsilon$  stokastik değişkeninin hangi değeri alacağı önceden bilinmemektedir. Yani, tesadüfî bir şekilde ortaya çıkan bir değişkendir. Hata terimi  $\varepsilon$ , modele

dâhil edilemeyen değişkenleri içerir. Modelde yer alan parametreler  $\beta_0$ , doğrusal fonksiyon sabit parametresidir.  $X$  sıfır olduğunda regresyon doğrusunun dikey eksen olan  $y$  ile kesiştiği noktayı gösterir.  $\beta_1$  ise doğrusal fonksiyonun eğimidir ve  $\beta_1$  parametresi regresyon analizinde bağımsız değişken  $x$ 'deki bir birimlik değişimin bağımlı değişken  $y$ 'de ne kadar bir değişim yarattığını gösteren regresyon katsayısıdır.

En Küçük Kareler yönteminde anakitle regresyon denkleminin parametreleri tahmin edilirken belirli varsayımların sağlandığı kabul edilir. Bu duruma ek olarak; En Küçük Kareler tahminçileri, yansızlık özellikleri ile birlikte en küçük varyansa sahip olduklarından tüm doğrusal tahminçiler sınıfı içinde “en iyi, doğrusal yansız” tahminçilerdir(Gujarati,2004).

Hardle 1997’de yayımladığı kitabında bir regresyon eğrisine parametrik olmayan yaklaşımın dört amacından bahseder (Topal,1999).

- I. İki değişken arasındaki genel ilişkiyi inceleyen çok yönlü bir metot sağlar.
- II. Bir sabit parametrik modelin varlığı olmaksızın gözlemlere ilişkin tahminlerin yapılmasını sağlar.
- III. İzole edilmiş noktaların etkisini hesaba katarak sahte gözlemlerin bulunması için bir araç temin eder.
- IV. Ardışık  $x$  değerleri arasında interpolasyon yapmak ya da kayıp değerleri yedeklemek için esnek bir metot ortaya koyar.

## 2.1. Parametrik ve Parametrik Olmayan Regresyon Analizi

Parametrik modellerde genel olarak regresyon analizinde yer alacak değişkenler arasındaki ilişkilere ait spesifik form doğrusal olarak kabul edilir. Parametrik yaklaşımda fonksiyonel formun, belirli parametre setleri tarafından bütünüyle tanımlandığı varsayılır. Bilindiği gibi parametrik modeller arasında En Küçük Kareler (Ordinary Least Squares) ve En Çok Olabilirlik yöntemi (Maximum Likelihood) gibi yöntemler parametrik regresyon analizinde bilinmeyen parametrelerin tahmininde yaygın olarak kullanılmaktadır.

Parametrik olmayan modellerde ise değişkenlere ait matematiksel formun ne olduğu ile ilgilenilmez. Söz konusu regresyon eğrisi hiçbir spesifik form varsaymadan “düzgünleştirme” teknikleri ile doğrudan tahmin edilir. Parametrik olmayan regresyon teknikleri regresyon fonksiyonu hakkında bilgi edinmek için parametrik tekniklere göre veriye daha fazla güvenmektedir. Bu nedenle çıkarım problemleri için uygun olmaktadır. Bunun yanı sıra

parametrik olmayan modeller parametrik modelin geçerliliğini test etmek amacıyla da kullanılmaktadır. Mood-Brown, Theil ve K-NN(En Yakın Komşu) gibi yöntemleri parametrik olmayan regresyon analizinde yaygın olarak kullanılmaktadır.

Parametrik ve parametrik olmayan regresyon yöntemleri arasındaki en önemli fark regresyon fonksiyonu hakkında araştırmacıdan ve veriden elde edilen bilgilere duyulan güvene dayanmaktadır. Parametrik regresyon modelinde araştırmacı tüm eğriler içerisinde olası bir eğriler ailesi seçer ve regresyon fonksiyonunun biçimi hakkında çok özel niceliksel bilgilere ihtiyaç duyar. Parametrik olmayan regresyon modelinde ise araştırmacı regresyon fonksiyonunun ait olduğu uygun fonksiyon uzayını genellikle kendisi seçer ve regresyon fonksiyonunun biçimi hakkında niteliksel bilgilere ihtiyaç duyar (Eubank,1988).

## 2.2. Parametrik Regresyon Analizi

Parametrik regresyon analizi, açıklayıcı ve açıklanan değişkenler ile bu değişkenler arasındaki ortalama ilişkinin matematiksel bir fonksiyonla ifade edilmesidir. Bu fonksiyondaki parametre vektörlerinin açık bir şekilde gösterilmesiyle, parametrik regresyon modelinde modelin belirlenmesi ve model uygunluğunun araştırılması gerçekleştirilebilir.

Parametrik bir regresyonun ekonometrik açıdan en önemli özelliği sınırlayıcı kısıtlamalarının varlığıdır. Parametrik bir modelin en önemli varsayımlarını şu şekilde sıralamak mümkündür (Güriş ve Çağlayan, 2000).

- Hata terimleri rassaldır ve beklenen değeri sıfırdır.  $E(u_i) = 0$
- Hata terimlerinin varyansları gözlemden gözleme değişmemektedir. Hata terimlerine ait varyanslar sabittir.  $E(u_i^2) = \sigma_u^2$
- Hata terimlerinin birbirini takip eden değerleri arasında ilişki yoktur.  $COV(u_i, u_j) = 0 \quad i \neq j$
- Hata terimleri ile bağımsız değişkenler arasında ilişki olmadığı varsayılmaktadır.  $COV(u_i, X_{ij}) = 0$

Burada doğrusallık kavramına ilişkin birtakım açıklamalarda bulunmakta fayda vardır. Parametrik yöntemlerle elde edilen regresyon denklemlerinde ilişkilerin doğrusal varsayılması, bağımsız değişkenlerin önünde bulunan katsayıların doğrusal olması anlamına gelmektedir. Parametrik modellerde yer alan bağımlı değişken ve bağımsız değişkenlerin, doğrusal olması

şartı aranmamaktadır. Parametrik regresyon denklemlerine, aşağıda yer alan denklemler örnek olarak gösterilebilir:

$$Y_i = \beta_0 + \beta_1 X_i$$

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2$$

$$Y_i = \beta_0 + \beta_1 \sqrt{X_i} \quad (2.2)$$

$$Y_i = \beta_0 + \beta_1 \left( \frac{1}{X_i} \right)$$

$$Y_i = \beta_0 \cdot X_i^{\beta_1} \cdot Z_i^{\beta_2}$$

Yukarıda eşitlik 2.2’de gösterilen tüm denklemler katsayıları itibariyle doğrusaldır. Bu örnek denklemlerden parametrik modellerde varsayılan doğrusallık varsayımının ne olduğu anlaşılmaktadır. Bağımsız değişkenlerin üstel, parabolik, ters ya da kareköklü olarak ifade edilmesi söz konusu parametrik modellerin doğrusallık özelliğini etkilememektedir. Söz konusu doğrusallık ilişkisi deneye dayalı bir yöntem olan serpilme(dağılma) diyagramları ile gözlemlenebilmektedir.

Parametrik modellere ait tahmin yapılırken sıklıkla kullanılan metot olan En Küçük Kareler Yöntemi, bu doğrusal varsayımı dikkate alarak parametre tahminlerini gerçekleştirir. En Küçük Kareler yöntemi parametrik modellere ait eğilimsiz, tutarlı, etkin ve doğrusal parametre tahmin ediciler elde edilmesinde uygun bir yöntem olarak uygulamalı çalışmalarda kullanılmaktadır.

### 2.2.1. En Küçük Kareler Yöntemi

Regresyon analizinde gözlem noktalarını temsil eden en iyi eğriyi bulma işleminde en çok kullanılan yöntem En Küçük Kareler (EKK) yöntemidir. En küçük kareler yöntemi, oluşacak hataların kareleri toplamı minimum olacak şekilde bir eğri denklemini bulma mantığına dayanır. Bu yaklaşımla bulunacak bir denklem en az hatalı ve en gerçekçi değeri verecek bir eğri denklemini olacaktır. EKK yönteminin başarılı bir şekilde uygulanabilmesi için birçok varsayım gerekmektedir. Bu varsayımların en önemlisi verilen noktalardaki ölçüm hatalarının normal dağılım göstermesidir.

Regresyon modeli; bağımsız değişkenlerin sayısına, fonksiyon özelliğine ve veri kaynağına göre gruplandırılır.

- Bağımsız değişkenlerin sayısına göre gruplandırıldığında basit regresyon analizi ve çoklu regresyon söz konusudur.
- Fonksiyon özelliğine göre gruplandırıldığında doğrusal regresyon analizi ve doğrusal olmayan regresyon analizini içerir.
- Veri kaynağına göre gruplandırıldığında ise ana kütle verileri ile regresyon analizi; örneklem verileri ile regresyon analizi ve zaman serilerinde regresyon analizi söz konusudur.

### 2.2.1.1. Basit Doğrusal Regresyon Modeli

Birbiriyle ilişkili olan iki değişkenin olduğu basit doğrusal regresyonda, değişkenlerden birisi bağımlı diğeri ise bağımsız değişken olarak kabul edilir(Baskan,1993).  $Y$  bağımlı değişkeni,  $x$  bağımsız değişkeni,  $\beta_0$  ile  $\beta_1$  bu değişkenin bilinmeyen parametrelerini ve  $\varepsilon$  hata terimlerini göstermek üzere anakitle için basit doğrusal regresyon denklemi aşağıdaki eşitlik 2.3'deki gibi gösterilir.

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, 2, 3, \dots, n \quad (2.3)$$

Doğrusal regresyon yönteminin en doğru sonucu vermesi için bazı varsayımları sağlaması gerekmektedir. Varsayımlardan herhangi birinin gerçekleşmemesi durumunda doğru sonuçlara ulaşılamaz. Bu varsayımları aşağıda verildiği gibi sıralayabiliriz (Berry, 1993):

- 1) Bağımsız değişkenlerin hepsi nicel veya nitel olarak ölçülmüş olması, bağımlı değişken  $Y$  'nin ise nicel ve sürekli olması gerekmektedir.  $X$  ve  $Y$  değişkenleri doğru olarak ölçülmelidir.
- 2) Tüm bağımsız değişkenlerin varyansının sıfırdan farklı olması gerekmektedir.
- 3) Bağımsız değişkenler arasında doğrusal bir ilişkinin olmaması gerekmektedir.
- 4) Hata terimleri ortalaması sıfırdır.  $E(\varepsilon_i) = 0$
- 5) Bağımsız değişkenler ve hata terimi arasında korelasyon olmamalıdır.
- 6) Hata terimlerinin varyansı sabit olmalıdır.  $E(\varepsilon_i^2) = \sigma^2$
- 7) Hata terimleri arasında korelasyon olmamalıdır.  $E(\varepsilon_j \varepsilon_i) = 0, (i \neq j)$
- 8) Hata terimleri  $\varepsilon_i$ , normal dağılmalıdır.

## Parametrelerin Tahmini

$\beta_0$  ve  $\beta_1$  'i tahmin etmek için birçok yöntem kullanılabilir. Bu yöntemlerden bir tanesi En Küçük Kareler yöntemidir. Burada farkların kareleri toplamı  $\sum_{i=1}^n (Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_i))^2$  minimize edilerek tahminler elde edilir. Regresyon modeli için bu fark  $Y_i$  ile  $\in (Y_i)$  arasındaki fark olacaktır. Yapılan işlemler sonucunda,

$$\sum Y_i = n\hat{\beta}_0 + \hat{\beta}_1 \sum X_i \quad (2.4)$$

$$\sum X_i Y_i = \hat{\beta}_0 \sum X_i + \hat{\beta}_1 \sum X_i^2 \quad (2.5)$$

eşitlik 2.4 ve eşitlik 2.5'deki gibi iki denklem elde edilir. Bu denklemlere 'Normal Denklemler' denir. Bunların çözümü ile parametreler tahmin edilir. Denklemlerden türetilen farklı formüller parametre tahmininde kullanılabilir.

En Küçük Kareler Regresyonu aşırı değerlerin etkisinde kalabilir. Hatalar normal dağılıyorsa değişkenler arasındaki ilişkiyi en iyi şekilde açıklar ancak hataların normal dağılmadığı durumlarda (aşırı değerlerin olması durumunda hatalar normal dağılmayabilir.) En Küçük Kareler tahminleri ve testleri iyi sonuç vermezler(Gujarati,2004).

## Belirlilik Katsayısı

Bağımlı değişkendeki değişimlerin bağımsız değişken veya değişkenler tarafından açıklanma oranını belirten katsayıya "Belirlilik Katsayısı( $R^2$ )" adı verilmektedir. Basit regresyonda tek bağımsız değişken olduğundan, bağımlı değişkendeki değişimler sadece bu değişken tarafından açıklanmaktadır.

$$R^2 = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (2.6)$$

bu katsayı eşitlik 2.6 gösterildiği gibi elde edilir.  $0 \leq R^2 \leq 1$  aralığında değer alır.  $R^2$ 'nin 1'e yaklaşması bağımlı değişkendeki değişimlerin bağımsız değişken tarafından iyi açıklandığını ortaya koyacaktır(Gujarati,2004).

### 2.2.1.2. Çoklu Doğrusal Regresyon Modeli

Çoklu doğrusal regresyon bir bağımlı değişken ile iki ve daha fazla bağımsız değişken arasındaki doğrusal bağıntıyı inceleyen bir yöntemdir. Bağımlı değişken ile  $n$  adet bağımsız değişken arasındaki ilişkiyi gösteren model genel olarak aşağıda verilen eşitlik 2.7’de olduğu gibi gösterilir.

$$Y = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \dots + \beta_k x_{ik} + \varepsilon_i \quad (2.7)$$

Çoklu doğrusal regresyon modelinde  $\beta_0$  parametresi regresyon sabitidir.  $\beta_1$ ’ den  $\beta_k$ ’ye kadar olan bağımsız değişkenlerin katsayıları ise kısmi regresyon katsayılarıdır.

Çoklu doğrusal regresyon analizinde elde edilen sonuçların geçerliliği temel varsayımların geçerli olmasına bağlıdır. Çoklu regresyon analizinde de basit regresyon varsayımları geçerlidir fakat iki varsayım daha vardır.

Çoklu doğrusal regresyonun temel varsayımları; verilerin normal dağılması, hataların sıfır ortalamaya sahip olması, değişen varyans ve otokorelasyon olmaması, bağımsız değişkenlerin tesadüfi değişken olmaması, bağımsız değişkenler arasında çoklu doğrusal bağlantının olmaması ve örnek hacminin bağımsız değişken sayısından fazla olmasıdır ( $n > k$ ).

Bağımsız değişkenlerden ikisi veya daha fazlası arasında kuvvetli doğrusal ilişki olması durumu “Çoklu Doğrusal Bağlantı” olarak adlandırılır. Tam çoklu doğrusal bağlantı olduğu zaman denklem sistemi çözülememekte, parametreler tahmin edilememekte veya bazı temel varsayımlar geçerli olmamaktadır.

#### Parametrelerin Tahmini

Çoklu doğrusal regresyon modelinin parametreleri basit doğrusal regresyon modelinin parametreleri gibi aynı yöntemlerle tahmin edilebilir. Anakitle çoklu doğrusal regresyon modeli, eşitlik 2.8 gibi gösterilir.

$$Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \dots + \beta_k X_{ik} + \varepsilon \quad (2.8)$$

En Küçük Kareler Yönteminde  $Y_i$  ile  $E(Y_i)$  arasındaki farkların kareleri toplamı minimize edileceğinden yapılan işlemler sonucunda,

$$\begin{aligned}
\sum_{i=1}^n Y_i &= n\hat{\beta}_1 + \hat{\beta}_2 \sum_{i=1}^n X_{i2} + \dots + \hat{\beta}_k \sum_{i=1}^n X_{ik} \\
\sum_{i=1}^n X_{i2} Y_i &= \hat{\beta}_1 \sum_{i=1}^n X_{i2} + \hat{\beta}_2 \sum_{i=1}^n X_{i2}^2 + \dots + \hat{\beta}_k \sum_{i=1}^n X_{i2} X_{ik} \\
&\cdot \\
&\cdot \\
&\cdot \\
\sum_{i=1}^n X_{ik} Y_i &= \hat{\beta}_1 \sum_{i=1}^n X_{ik} + \hat{\beta}_2 \sum_{i=1}^n X_{ik} X_{i2} + \dots + \hat{\beta}_k \sum_{i=1}^n X_{ik}^2
\end{aligned} \tag{2.9}$$

eşitlik 2.9' da gösterilen denklemler elde edilir. Bu denklemler çoklu doğrusal regresyonun normal denklemleridir. Elde edilen normal denklemler gerçek değerlerle ifade edilmiştir. Bu denklemler yardımı ile,

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}_2 - \hat{\beta}_3 \bar{X}_3 - \dots - \hat{\beta}_k \bar{X}_k \tag{2.10}$$

eşitlik 2.10'da gösterildiği gibi tahmin edilecektir.

Çoklu doğrusal regresyonda tahmin edilecek parametre sayısı arttıkça tahmincilerin formüllerle ifadesi güçleşir. Bu nedenle parametrelerin tahmininde matrisler kullanılır.

### **Belirlilik Katsayısı ve Düzeltilmiş Belirlilik Katsayısı**

Çoklu doğrusal regresyonda birden fazla bağımsız değişken olduğundan, belirlilik katsayısı ( $R^2$ ) bağımlı değişkendeki değişimlerin bağımsız değişkenler tarafından açıklanma oranını verecektir. Basit regresyonda olduğu gibi teorik değerler çoklu regresyonda da hesaplanacak ancak işlem kolaylığı sağlanması açısından matrisler kullanılacaktır.

Aynı bağımlı değişkendeki değişimler farklı regresyon modelleri ile açıklanabilir. Bu modellerin matematiksel yapıları, gözlem sayıları ve değişken sayıları farklı olabilir. Bağımlı değişkeni aynı, bağımsız değişken sayıları farklı regresyon modellerinin karşılaştırılması için düzeltilmiş belirlilik katsayısı  $\bar{R}^2$  kullanılır.

$$\bar{R}^2 = 1 - \frac{n-1}{n-k} \cdot (1 - R^2) \tag{2.11}$$

Düzeltilmiş belirlilik katsayısı, eşitlik 2.11'de gösterildiği gibi hesaplanır.

## 2.2.2. En Çok Olabilirlik Yöntemi

Lojistik regresyonda bağımlı değişken  $Y$  ile bağımsız değişkenler arasındaki ilişki doğrusal olmadığından dolayı model parametreleri *EKK* yöntemi ile tahmin edilemez. Bunun yerine En Çok Olabilirlik yöntemi yaygın olarak tercih edilen tahmin yöntemlerinden biridir. Bu yöntemi uygulamak için öncelikle, olabilirlik fonksiyonunun oluşturulması gerekmektedir. Bu fonksiyon, gözlenen verilerin olasılıklarını bilinmeyen parametrelerin bir fonksiyonu olarak açıklar.

İkili lojistik regresyonda  $Y$ 'nin 0 ve 1 gibi iki değer aldığı bilinmektedir. Bağımlı değişken  $Y$ 'nin ilgilenen olay gerçekleşirse 1 değeri alma olasılığı  $P_i$  ve ilgilenilen olay gerçekleşmezse 0 değerini alma olasılığı  $1-P_i$  olarak ifade edilir.  $X$  değerleri verildiğinde,  $N$  tane örneklemlenmiş  $Y$  değerlerinden birisinin gözlemlenme olasılığı aşağıdaki eşitlik 2.12'de olduğu gibi ifade edilir.

$$P(Y/X) = P_i^{Y_i} (1 - P_i)^{1-Y_i} \quad (2.12)$$

Lojistik regresyonda amaç olabilirliği en büyük yapan  $\beta$  değerlerini bulmak olduğundan logaritması alınan fonksiyonun  $\beta$  'lara göre türevi alınarak sıfıra eşitlenir.

## 2.3. Parametrik Olmayan Regresyon Analizi

İktisadi değişkenler arasındaki ilişkilerin yapısı çoğu zaman başlangıçta belirlenemez. Verilerin fonksiyonel yapılarının belirlenmesi için önsel çalışmalar gerekebilir. Gerekli varsayımların sağlanamaması durumunda parametrik yöntemler devre dışı kalacaktır. Bu gibi durumlarda parametrik olmayan veya yarı parametrik regresyon analizi yöntemleri kullanılması tavsiye edilmektedir. Parametrik olmayan bir regresyon modeli aşağıdaki eşitlik 2.13'te gösterilmektedir:

$$Y_i = m(\chi_i) + u_i \quad (2.13)$$

Parametrik modellerin dezavantajları ışığında birçok parametrik olmayan tahminci,  $m(X_i)$ 'in "düzgün" tahminlerini üretmek amacıyla geliştirilmiştir (Pagan ve Ullah, 1999). Bu regresyon modelinde parametrik olmayan yaklaşım olarak tabir edilen yapı,  $m(X_i)$ ' in herhangi bir fonksiyonel form varsayılmadan tahmin edilmesidir.

Parametrik olmayan tahmin bu yapıyla ilgilenmektedir. Parametrik olmayan modellere özgü “düzgünleştirme(smoothing)” teknikleri ile  $m(X_i)$  anakitle regresyon fonksiyonuna ait en uygun düzgün tahmini üretmeye çalışmaktadır.

Parametrik olmayan yaklaşımla bir regresyon ilişkisini belirlemeye çalışmanın 4 avantajı bulunmaktadır. Bunlar:

- Parametrik olmayan tahmin yöntemi bağımlı ve bağımsız değişken arasındaki genel ilişkinin araştırılmasında çok yönlü ve esnek metot sağlamaktadır.
- Herhangi bir fonksiyonel forma ihtiyaç duyan sabit parametrik modeli temel almaksızın gözlemlere ait tahminler yapar.
- Uç değer olarak tabir edilen değerlerin etkisinin analizde yer alması sebebiyle “sahte” gözlemleri bulmak için bir araç niteliği taşımaktadır.
- Bağımsız değişkenler arasında enterpolasyon yapılabilmesi ve veri seti içerisinde kayıp gözlemlerin yerine konulabilmesi açısından esnek bir metot oluşturmaktadır.

**Tablo 2. 1.** Parametrik Olmayan Regresyonda Değişken Sayısı Arttıkça Gerekli Örnek Büyüklüğü

Boyut	Gerekli Örnek Büyüklüğü
1	4
2	19
3	67
4	223
5	768
6	2790
7	10700
8	43700
9	187000
10	842000

**Kaynak:** Silverman, Density Estimation for Statistics and Data Analysis, Chapman and Hall, 1986, s.94

### 2.3.1. Düzgünleştirme Kavramı

Parametrik olmayan regresyon analizinde düzgünleştirme kavramı önemli bir detay olarak karşımıza çıkmaktadır. Düzgünleştirmenin temelinde yerel olarak ortalama alma fikri yatmaktadır(Efron,1998). Düzgünleştirici; bir ya da birden fazla bağımsız değişkenin fonksiyonu olan bağımlı değişkenin sahip olduğu eğilimi ifade etmek için kullanılan bir araçtır. Burada amaç; bağımlı değişkenin kendisinden daha az değişken bir eğilim tahmini yapmasıdır. Bir düzgünleştiricinin en belirgin özelliği, değişkenler arasındaki ilişkinin biçimini kesin bir biçimde varsaymamasıdır ve bu özelliğinden dolayı parametrik olmayan regresyonda sık kullanılan bir araçtır.

Parametrik regresyonda elde edilen doğru, kesin parametrik bir biçime sahip olduğu için düzgünleştirici değildir. Bilinen en basit düzgünleştirici ise hareketli ortalamalar yöntemidir. Bir düzgünleştirici tarafından yapılan tahmin edilen eğriye “düzgün” adı verilmektedir. Tek bir bağımsız değişkenin olduğu durumlar ise “Serpilme Diyagramı Düzgünleştirme” olarak adlandırılmaktadır. Serpilme diyagramındaki her bir nokta, herhangi bir tesadüfi model olmaksızın düzlem üzerindeki noktaların birikimi olarak kabul edilmektedir(Ruppert vd., 2003).

Parametrik olmayan düzgünleştirme yöntemlerinin en önemli faydaları;

- Veri setinin özelliklerinin gösterilmesinde uygun, az ve öz bir araç olarak kullanılabilir ve böylelikle pratik bir parametrik modelin kurulmasına yardım eder.
- Tahmin edilmiş parametrik bir modelin kontrol edilmesinde kullanılabilir.
- Tamamen parametrik olmayan bir yapıda oldukça zayıf kısıtlamaların varlığı altında çıkarım yapılabilir.
- Parametrik olmayan tahmin ediciler, yarı parametrik modellerde Euclidian değerli büyüklüklerin tahminleyenlerinin oluşturulmasında kullanılabilir.

### 2.4. Parametrik Regresyon ile Parametrik Olmayan Regresyon Analiz Yöntemleri Karşılaştırılması

Parametrik regresyon analizinde, biçimi önceden belirlenmiş modelin parametreleri tahmin edilirken, parametrik olmayan regresyon analizinde ise amaç, regresyon fonksiyonu olan  $m(X_i)$ ’i doğrudan tahmin etmektir(Hardle, Müller, Sperlich ve Werwartz, 2004).

Parametrik regresyon analizinde; regresyon fonksiyonunun biçimi ile birlikte  $X$  veri iken  $Y$  'nin koşullu varyansı ve hata terimi  $u$  ile  $X$  arasındaki otokorelasyon gibi diğer fonksiyonlar belirtilmiştir. Ayrıca  $X$  ve  $Y$  'nin parametrik bileşik yoğunluğunun genellikle normal olduğu varsayılır. Parametrik olmayan regresyon analizinde ise bu varsayımlar yapılmaz. Yapılan en önemli varsayım hata terimlerinin sıfır ortalama ve sonlu bir varyansa sahip bir dağılımdan geldiğidir. Parametrik olmayan regresyon için model belirlenirken bilinmeyen regresyon eğrisini içeren uygun fonksiyon uzayı seçilir. Bu seçim yapılırken regresyon fonksiyonunun süreklilik ve türevlenebilme gibi düzgünlük özelliklerine sahip olduğu noktasından hareket edilir(Eubank,1990).

Parametrik olmayan regresyon analizi; aykırı gözlemlerin (outliers) bulunduğu veri yapılarında başarılı bir şekilde kullanılan önemli bir analiz yöntemidir. Literatürde aykırı gözlemlerin etkilerini farklı biçimlerde ele alan güçlü (robust) parametrik yöntemler vardır. Yine de aykırı gözlemlerden dolayı parametreler bozulmalar yaşadığı için bu güçlü yöntemler bile uygun çözümler üretmeyebilirler ve verinin gerçek yapısı modele yansıtılamazlar. Bu durumda parametrik olmayan regresyon,  $X$  doğrultusunda belirli bir parametrik model olmaksızın ön bilgi sağlar(Hardle,1999).

Parametrik olmayan regresyonun sağladığı bu kolaylıkların yanında parametrik regresyona göre bazı dezavantajları da bulunmaktadır. Parametrik regresyon analizine göre artan işlem sayısı işlem karmaşıklığını artırmaktadır ve bazı durumlarda elde edilen sonuçların açıklanmasındaki zorluklar bunlardan bazılarıdır (Fox,2000).

En önemli noktalardan biri ise parametrik olmayan model gerçeğe daha yakın olsa bile parametrik model geçerli ise parametrik olmayan tahminleyenlerin etkinliği, parametrik tahminleyenlerden daha azdır (Moussa ve Cheema,1992).

## BÖLÜM 3

### 3. KANTİL REGRESYON

#### 3.1. Kantil Kavramı

Bu bölümde, kantil kavramı, kantil dağılım fonksiyonu, kantil yoğunluk fonksiyonu, kantil regresyon, kantil regresyonun doğrusal programlama gösterimi ele alınacaktır.

Serileri iki, dört, on ve yüz eşit parçaya ayıran değerler genel olarak kantil olarak adlandırılmaktadır. Seriyi iki eşit parçaya bölmek için hesaplanan değerlere medyan, dört eşit parçaya bölmek için hesaplanan değerlere çeyreklik (kantil), on eşit parçaya bölmek için hesaplanan değerlere ondalık (desil) ve yüz eşit parçaya bölmek için hesaplanan değerlere pörsentil adı verilmektedir (Serper, 2004:123).

##### a) Kantiller

Kantiller, büyüklük sırasına konulmuş bir seriyi dört eşit kısma bölen değerlerdir. Bir seride üç kantil mevcuttur ve ikinci kantil medyandır. Birinci kantil  $Q_1$ , ikinci kantil  $Q_2$  ve üçüncü kantil  $Q_3$  ile gösterilir. Basit ve sınıflandırılmış serilerde,

$$Q_1 = \frac{N+1}{4} \cdot \text{nci değer}$$

$$Q_2 = \frac{2(N+1)}{4} \cdot \text{nci değer ya da } \frac{N+1}{2} \cdot \text{nci değer}$$

$$Q_3 = \frac{3(N+1)}{4} \cdot \text{nci değerdir.}$$

##### b) Desiller

Desiller, büyüklük sırasına konulmuş bir seriyi 10 eşit kısma bölen değerlerdir. Dokuz adet desil vardır. Birinci desil  $D_1$ , ikinci desil  $D_2$ , ... dokuzuncu desil  $D_9$  ile gösterilir. Basit serilerde,

$$D_1 = \frac{N+1}{10}, \text{ uncu deęer,}$$

$$D_2 = \frac{2(N+1)}{10} = \frac{N+1}{5}, \text{ inci deęer,}$$

$$D_5 = \frac{5(N+1)}{10} = \frac{N+1}{2}, \text{ inci deęer} = Q_2 = \text{Medyan}$$

$$D_9 = \frac{9(N+1)}{10}, \text{ uncu deęerdir.}$$

### c) Pörsentil

Pörsentiller, büyüklük sırasına konulmuş bir seriyi 100 eşit kısma bölen deęerlerdir. Toplam 99 adet pörsentil vardır. En küçüğü birinci pörsentil ( $P_1$ ) en büyüğü doksan dokuzuncu ( $P_{99}$ ) pörsentillerdir. Basit serilerde,

$$P_1 = \frac{N+1}{100}, \text{ üncü deęer,}$$

$$P_2 = \frac{2(N+1)}{100} = \frac{N+1}{50}, \text{ nci deęer}$$

$$P_{50} = \frac{50(N+1)}{100} = \frac{N+1}{2}, \text{ nci deęer} = Q_2 = D_5 = \text{Medyan}$$

$$P_{99} = \frac{99(N+1)}{100}, \text{ üncü deęerdir.}$$

## 3.2. Anakitlenin Modellenmesi

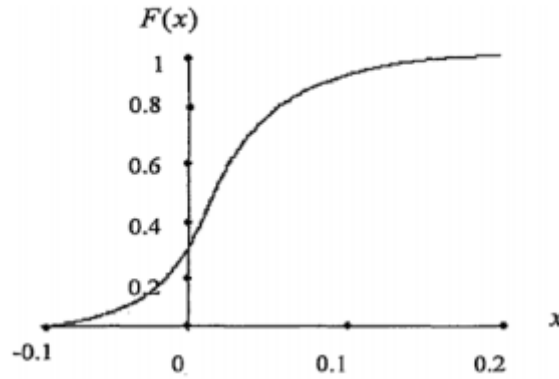
Herhangi bir dağılıma sahip örnek yapısı gösterilirken kümülatif dağılım fonksiyonu, olasılık yoğunluk fonksiyonu, kantil fonksiyonu ve kantil yoğunluk fonksiyonu olmak üzere dört farklı yol kullanılır.

### 3.2.1. Kümülatif Dağılım Fonksiyonu

Kümülatif dağılım fonksiyonu  $F(x)$ 'le gösterilir; verilen  $x$  deęerinde küçük ya da eşit olan  $X$  deęişkeninin olasılığı olarak tanımlanır ve

$$F(x) = P(X \leq x)$$

olarak ifade edilebilir. Azalan fonksiyon olan kümülatif dağılım fonksiyonu aşağıda şekil 3.1’de gösterilmiştir.



**Şekil 3. 1** Kümülatif Dağılım Fonksiyonu

Görüldüğü gibi fonksiyon sol taraftan '0' olasılık değerinden başlayarak, sağ tarafta bulunan '1' olasılığına kadar devam etmektedir. Herhangi bir olasılığın  $\theta$ , örnekten elde edilen değerinin  $x$  'e göre çizimini verir. Kümülatif dağılım fonksiyonu (*CDF*) dağılımının tek yönlü tanımlamasını yapar(Gilchrist, 2000: 10).

### 3.2.2. Olasılık Yoğunluk Fonksiyonu

Bir değişkenin alabileceği değerlerle bu değerleri alma olasılıkları arasındaki bağıntıyı gösteren fonksiyona ‘Olasılık Yoğunluk Fonksiyonu (*PDF*)’ denir (Saraçoğlu ve Çevik, 1995: 67).  $f(x)$  ile gösterilen olasılık yoğunluk fonksiyonu,

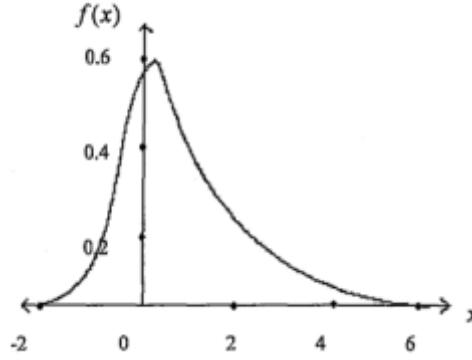
$$f(x)dx = P(x \leq X \leq x + dx) \quad (3.1)$$

eşitlik 3.1’de olduğu gibi tanımlanır. Burada  $dx$ ,  $x$  'in sonsuza doğru küçük aralığıdır.  $f(x)$  eğrisinin altındaki alan, herhangi gözlenen değer için toplam olasılığı 1 olmalıdır. Kümülatif dağılım fonksiyonu ve olasılık fonksiyonu arasındaki ilişki,

$$f(x)dx = F(x+dx) - F(x) = dF(x) \quad (3.2)$$

eşitlik 3.2’deki gibi olacaktır. Olasılık fonksiyonu, kümülatif dağılım fonksiyonunun türevine eşit olup eşitlik 3.3’de gösterildiği gibi elde edilir.

$$f(x) = dF/dx \quad (3.3)$$



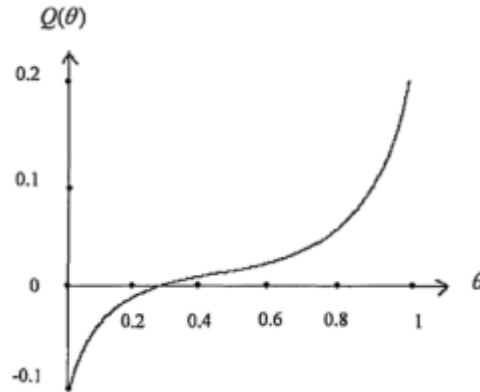
Şekil 3. 2 Olasılık Yoğunluk Fonksiyonu

Olasılık Yoğunluk Fonksiyonu 'da şekil 3.2' de olduğu gibidir.

### 3.2.3. Kantil Fonksiyon

Kantil fonksiyon  $QF$  ve  $Q(\theta)$  ile gösterilmektedir. Kantil değeri değişkenin dağılımında yer alan ve dağılımı, kendisinden büyük olanlar ve kendisinden küçük olanlar diye ikiye bölen herhangi değerdir. Yani değerlerin %  $\theta$ ' sı,  $\theta$ . kantilden daha küçüktür. ( $\theta$  olasılık değerini ifade etmektedir)

$x_0 = (X \leq x_0)$  olasılığı için  $x$ 'in değeridir.



Şekil 3. 3 Kantil Fonksiyonu

$x_\theta$ 'nın değeri, kitlenin  $\theta$ 'cı kantil olarak adlandırılır.  $x_\theta = Q(\theta)$  fonksiyonu,  $\theta$ . kantil,  $\theta$ 'nın bir fonksiyonu olarak ifade edilir ve kantil fonksiyon olarak adlandırılır. (0,1) aralığındaki her  $\theta$  ve rasgele  $Y$  değişkeni için,  $Y$ 'nin  $\theta$ . kantili her  $\varepsilon_\theta \in \mathbb{R}$

$$P(Y < \varepsilon_\theta) \leq \theta \leq P(Y < \varepsilon_\theta) \quad (3.4)$$

eşitlik 3.4'te gösterildiği gibi tanımlanır.

Kantil fonksiyon ve kümülatif dağılım fonksiyonu, herhangi  $(x, \theta)$  çifti için  $x_\theta = Q(\theta)$  ve  $\theta = F(x)$  şeklinde yazılabilir. Bu fonksiyonlar birbirlerinin tersine eşit ve sürekli artan fonksiyonlardır. Böylece;

$$Q(\theta) = F^{-1}(\theta) \text{ ve } F(x) = Q^{-1}(\theta) \quad (3.5)$$

eşitlik 3.5'te olduğu gibi gösterilir (Gilchrist, 2000: 13).

$Q(\theta)$  kantil fonksiyonu ise,  $\theta$ 'nin tüm olasılıkları için,  $0 \leq \theta \leq 1$  kantil değerlerini verir. Medyanda  $Q(0,5)$ 'tir ve benzer şekilde  $Q(1/4)$ ,  $Q(3/4)$  kantilleridir. Hesaplamalar için normal dağılım tablosundan faydalanılır. Kullanılan normal tablolar standart normal dağılım için kantil fonksiyon tablolarıdır.

Dağılımları modelleyebilmek için kantil fonksiyon kullanılabilir.  $x$  verilmişken  $y$ 'nin  $\theta$ . kantili;

$$Q_y(\theta/x) = \beta_x + \eta S(\theta) \quad (3.6)$$

eşitlik 3.6'da olduğu gibi gösterilir. Burada;

$\eta S(\theta)$  = artık terimi olarak ifade edilir.

$S(\theta)$  = simetrik olması gerekmeyen kantil fonksiyonudur.

$\eta$  = ölçek parametresidir.

' $y$ 'nin  $x$  üzerindeki kantil regresyon fonksiyonu ya da 'şartlı kantil fonksiyonu' olarak adlandırılır (Gilchrist, 2000: 13).

### 3.2.4. Kantil Yoğunluk Fonksiyonu

Dağılımları modelleyebilmek için, dağılım fonksiyonunun türevini alarak olasılık yoğunluk fonksiyonu elde edildiği gibi,  $QF$ 'in de türevi alınarak kantil yoğunluk fonksiyonu (QDF) belirlenebilir ve

$$q(\theta) = dQ(\theta) / d\theta \quad (3.7)$$

eşitlik 3.7' de olduğu gibi gösterilir.  $Q(\theta)$  azalmayan bir fonksiyon olduğu için eğimi  $q(\theta)$  negatif değildir, her zaman  $0 \leq \theta \leq 1$  birim aralığında yer almakta iken, olasılık yoğunluk fonksiyonu  $f(x)$  ise sonsuz tanım aralığında yer almaktadır.

Serinin mod değerinin olasılığı ( $p - \text{mod} \geq 0,5$ ) ise dağılım sola çarpıktır ve  $q(\theta)$  yoğunluk fonksiyonu  $q(\theta) \leq q(1 - \theta)$  durumunu sağlar,  $0 \leq \theta \leq 0,5$  'tir. Kantil fonksiyonu da  $Q(\theta) + Q(1 - \theta) \leq 2Q(0,5)$  durumunu sağlar ve ortalama  $\leq$  medyan  $\leq$  mod sıralaması sağlanır.

Benzer şekilde serinin mod değerinin olasılığı ( $p - \text{mod} \geq 0,5$ ) ise dağılım sağa çarpıktır ve  $q(\theta)$  yoğunluk fonksiyonu durumunu sağlar,  $0 \leq \theta \leq 0,5$  'tir.

Kantil fonksiyonu da  $Q(\theta) + Q(1 - \theta) \geq 2Q(0,5)$  durumunu sağlar ve ortalama  $\geq$  medyan  $\geq$  mod sıralaması sağlanır.

### 3.3. Kantil Regresyon Analizi

Kantil regresyon, ilk olarak regresyondaki klasik varsayımlardan hata terimlerinin normal dağılması varsayımını ihmal eden robust (sağlam) bir regresyon tekniği olarak ortaya çıkmıştır. Gelir dağılımındaki eşitsizlik ya da ücretlerdeki farklılık gibi dağılımın bozulduğu konularda kullanımı yaygın olan Kantil Regresyon, daha kapsamlı bir regresyon görüntüsü sunmak amacıyla tasarlanan bir yöntemdir (Koenker, 2005:112).

Uygulamalı istatistiğin önemli bir kısmı lineer regresyon modeli ve bu modelin tahmininde sıklıkla kullanılan En küçük Kareler tahmin metotlarının detaylı bir şekilde incelenmesi olarak görülebilir (Koenker, 2005:115).

Kantil regresyon modelleri şartlı ortalama fonksiyonları ve şartlı kantil fonksiyonları için tahmin yapılmasında kullanılır. Mostsellers ve Tukey; “regresyon eğrisinin yaptığı şey  $x$  'lerin kümesine karşılık gelen dağılımların ortalamasının özet bilgisini vermektedir. Daha fazla bilgi için dağılımların çeşitli yüzde puanlarına denk gelen farklı regresyon eğrileri hesaplanabilir, bu sayede de kümenin daha detaylı bilgisi elde edilebilir. EKK yönteminde bu yapılmaz ve bu nedenle genellikle regresyon eğrisi bize eksik bilgi verir. Sadece ortalama değerinin bir dağılım için eksik bilgi vermesi gibi, EKK regresyon eğrisi de dağılım kümeleri hakkında eksik bilgi verir.” şeklinde ifade etmişlerdir (Koenker, 2005).

Kantil Regresyon, özellikle koşullu kantillerin değişkenlik gösterdiği durumlarda kullanışlıdır. Kantillere bağlı olarak regresyon katsayılarını belirler (Chen, 2005).

Çoklu doğrusal regresyon modelinde hata teriminin değişkenlerin değerinden bağımsız olduğu (varyanslar homojen) varsayılır. Tam tersine kantil regresyon modelinde hata terimlerinin değişkenliğine izin verilir ve varyans yapısına ilişkin herhangi bir varsayımı bulunmamaktadır (Baur ve ark. 2004:4695).

Bağımsız değişken  $x$  'in değişen değerlerine karşı kantil regresyonlarla analiz yapmak, aşırı değerlerin varlığı durumunda daha etkin sonuçlar ortaya koymaktadır. Çoklu doğrusal regresyon doğrusu, aşırı değerleri yakalayamazken farklı kantillerdeki kantil regresyon doğruları aşırı değerleri rahat bir şekilde yakalayabilir (Wang, 2007:9-10).

Kantil regresyon modeli aslında bir yerleşim modelidir. Basit yerleştirme modeli;

$$Y_i = \beta + e_i \quad (3.8)$$

eşitlik 3.8'de olduğu gibi ifade edildiğinde, burada yer alan  $Y_i$ , simetrik  $F$  dağılım fonksiyonuna sahip, bağımsız, özdeş dağılımlı,  $\beta$  medyanlı tesadüfi değişkendir. Bu modelde  $\theta$ . kantili;

$$\min_{\beta} \frac{1}{n} \left\{ \sum_{i: y_i \geq \beta} \theta |y_i - \beta| + \sum_{i: y_i < \beta} (1 - \theta) |y_i - \beta| \right\} \quad (3.9)$$

eşitlik 3.9'daki ifadenin minimizasyonu ile ifade edilir (Judge, 1985:834). Bu doğrusal regresyon modeli eşitlik 3.10'daki gibi hesaplanır.

$$y_i = x_i \beta + e_i \quad (3.10)$$

$\theta$ . kantil regresyon gözlem değerlerinin işaretlerine bağlı olarak

$$\theta \min_{\beta} \frac{1}{n} \sum_{i=1}^n \left( \theta - \frac{1}{2} + \frac{1}{2} \operatorname{sgn}(y_i - x_i \beta) \right) (y_i - x_i \beta) \quad (3.11)$$

eşitlik 3.11'de olduğu gibi tahmin edilir. Burada,  $\operatorname{sgn}(\alpha)$   $\alpha$  'nın işaretidir ve  $\alpha$  pozitif ise 1, negatif veya 0 ise -1 değerini alır. Tahmincilerin gözlem değerlerinin büyüklüğü yerine işaretlerine dayalı olması kantil regresyonun robust bir yöntem olmasını sağlamaktadır. Minimizasyon için birinci mertebe koşulunun sağlanması gerektiğinden birinci mertebe koşulunun  $K \times 1$  vektörü,

$$\theta \min_{\beta} \frac{1}{n} \sum_{i=1}^n \left( \theta - \frac{1}{2} + \frac{1}{2} \operatorname{sgn}(y_i - x_i \beta) \right) (y_i - x_i \beta) = 0 \quad (3.12)$$

eşitlik 3.12'de olduğu gibi gösterilir. Bu ifade birinci mertebe koşulu genelleştirilmiş momentler yöntemi (GMM)'ne uyan bir moment fonksiyonudur. Moment fonksiyonu,

$$\psi(x_i, y_{i,\beta}) = \left( \theta - \frac{1}{2} + \frac{1}{2} \operatorname{sgn}(y_i - x_i \beta) \right) x_i \quad (3.13)$$

eşitlik 3.13'te olduğu gibi tanımlanabilir.  $\psi(\cdot)$  'nin fonksiyon olarak geçerli olabilmesi için belirli düzenleme şartları altında

$$\in [\psi(x_i, y_i, \beta_0)] = 0 \quad (3.14)$$

eşitlik 3.14'te olduğu gibi olması gerekir. Asimptotik dağılımın genel gösterimi;

$$\in [f_{u\theta}(0/x_i)x_i x_i'] = \partial \in [\psi(x_i, y_i, \beta_0)] / \partial \beta_\theta' \theta(1-\theta) \in [x_i x_i'] = \in [\psi(x_i, y_i, \beta_0) \psi(x_i, y_i, \beta_0)'] \quad (3.15)$$

eşitlik 3.15'deki şeklinde olup genelleştirilmiş momentler yöntemi kullanılarak elde edilen parametre tahminçileri tutarlı ve asimptotik olarak normal olacaktır. Belirli düzenleme şartları altında,

$$\sqrt{n}(\hat{\beta}_\theta - \beta_\theta) \rightarrow N(0, \Lambda_\theta) \quad (3.16)$$

eşitlik 3.16'daki gibi gösterilebilir. Burada,

$$\Lambda_\theta = \theta(1-\theta) \left( \in [f_{u\theta}(0/x_i)x_i x_i'] \right)^{-1} \in [x_i x_i'] \left( \in [f_{u\theta}(0/x_i)x_i x_i'] \right)^{-1} \quad (3.17)$$

eşitlik 3.17'de olduğu gibi tanımlanır (Koenker ve Bassett, 1978:48).

Olasılık değeri 1 olduğunda ve  $f_{u\theta}(0/x_i) = f_{u\theta}(0)$  ise, yani hata teriminin yoğunluğu sıfır etrafındaysa ve  $x'$  ten bağımsızsa,  $\Lambda_\theta$ ,

$$\in (y_i / x_i) = \mu(x_i, \beta) = \mu_i \quad (3.18)$$

eşitlik 3.18'deki şekildeki gibi sadeleştirilebilir.  $f_{u\theta}(\cdot/x)$   $x'$  ten bağımsız olduğunda, tüm quantillerin parametre ve vektörleri sadece kesim noktalarında farklılık gösterir.

Kantil katsayılarını yorumlayabilmek için  $y$ 'nin  $k$  açıklayıcı değişkenine göre şartlı kantil 'nin kısmi türevi alınmaktadır. Türev alındığında,

$$\partial Quant_\theta(y_i / x_i) / \partial x_{ik} \quad (3.19)$$

eşitlik 3.19'daki gibi olacaktır. Bu türev,  $x'$  in  $k$ 'nci değerindeki marjinal değişime göre,  $\theta$ ' cı şartlı kantildeki marjinal değişimi vermektedir (Behr, 2008:570).

Kantil regresyon parametre tahminleri, açıklayıcı değişkendeki bir birim değişme karşısındaki  $y$ 'nin belli bir kantilindeki değişmeyi gösterir.

### 3.3.1. Kantil Regresyon Yönteminin Özellikleri

Kantil regresyonun en önemli özellikleri aşağıdaki gibi sıralanabilir (Leping, 2005:15).

- I. Çoklu Doğrusal Regresyon yöntemi  $y$  'nin koşullu dağılımının ortalaması hakkında bilgi vermekte, kantil regresyon ise farklı kantil değerleri için  $y$  'nin  $x$  'e göre koşullu dağılımının tümü hakkında bilgi vermektedir.
- II. Kantil Regresyonda;  $\min_b \frac{1}{n} \sum_{i=1}^n \rho_\theta(y_i - x_i'b)$  ifadesinin minimizasyonu, doğrusal programlama(LP) gösterimidir, bu durumda tahmin kolaylaşmaktadır.
- III. Kantiller monoton dönüşümlere olanak verirler. Herhangi  $h(\cdot)$  monoton fonksiyonu için  $Q_{h(y)/x}(\tau/x) = h(\theta_{y/x}(\tau/x))$  olur.
- IV. Kantiller bağımlı değişkendeki aşırı değerlere karşı kararlılırlar (robust).
- V. Hata terimi normal dağılmadığında, kantil regresyon tahmin edicileri çoklu doğrusal regresyon tahmin edicilerinden çok daha etkin olabilmektedir.
- VI. Kantil regresyon değişen varyansın belirlenmesine olanak vermektedir.
- VII. Kantil regresyon amaç fonksiyonu için tahmin edilen katsayı vektörü, bağımlı değişkendeki aşırı değerlere duyarlı değildir ve yerleşimin robust bir ölçüsüdür.
- VIII. Farklı kantillerde farklı sonuçlar çıkması, bağımlı değişkenin koşullu dağılımının farklı noktalarındaki bağımsız değişkenlerdeki değişikliklere farklı tepki vermesi olarak yorumlanabilir.
- IX. Kantil regresyon analizinde  $\tau = 0,5$  olması durumunda En Küçük Mutlak Sapmalar (LAD) regresyon analizi elde edilmektedir.

### 3.3.2. Kantil Regresyonun Doğrusal Programlama Gösterimi

Kantil regresyon tahmin edicileri doğrusal programlama problemi olarak formüle edilebilir ve artıkların iki parçalı doğrusal amaç fonksiyonu optimize edilerek simpleks veya sınır metodu gibi yöntemlerle çözülebilir (Koenker ve Hallock, 2001: 153).

$Y, F$  dağılım fonksiyonuna sahip bağımlı değişken ve  $b$ , tahmin edilecek katsayı vektörü ve  $e_i = y_i - x_i\beta$  hata değeri olmak üzere,  $\theta$ . regresyon kantili ( $0 < \theta < 1$ ) ;

$$\min_{\beta} \frac{1}{n} \left\{ \sum_{i: y_i \geq x_i\beta} \theta |y_i - x_i\beta| + \sum_{i: y_i < x_i\beta} (1-\theta) |y_i - x_i\beta| \right\}$$

$$\min \frac{1}{n} \left\{ \sum_{i=1}^n \rho_{\theta}(y_i - x_i \beta) \right\} \quad (3.20)$$

eşitlik 3.20'deki gibi minimizasyon ile tahmin edilir.  $y$ 'nin  $\theta$ . kantil olarak adlandırılır. Kantil regresyonun bu şekilde gösterimi doğrusal programlama gösterimidir.  $\beta$  parametre tahminçileri;

$$\hat{\beta}(\theta) = \arg \min_{\beta \in R^p} \left\{ \sum_{i=1}^n \rho_{\theta}(y_i - x_i \beta) \right\} \quad (3.21)$$

eşitlik 3.21'deki çözümü ile tahmin edilebilir. Burada  $0 < \theta < 1$ 'dir.  $Y_i$  eşitlik 3.22'deki gibi sadece pozitif elemanların fonksiyonu;

$$y_i = \sum_{j=1}^K x_{ij} \beta_{\theta_j} + u_{\theta_i} = \sum_{j=1}^K x_{ij} (\beta_{\theta_j}^1 - \beta_{\theta_j}^2) + (\varepsilon_{\theta_i} - \nu_{\theta_i}) \quad (3.22)$$

$$\beta_{\theta_j}^1 \geq 0, \beta_{\theta_j}^2 \geq 0 (j = 1, \dots, K) \text{ ve } \varepsilon_{\theta_i} \geq 0, \nu_{\theta_i} \geq 0 (i = 1, \dots, n) \quad (3.23)$$

eşitlik 3.23'de gösterildiği gibi yeniden yazılabilir (Buchinsky, 1998:91-92).

Daha etkili bir tahmin için;

$$\min_{\beta} \frac{1}{n} \left\{ \sum_{i=1}^n f_{u\theta}(\theta/x) \left( \theta - \frac{1}{2} + \frac{1}{2} \operatorname{sgn}(y_i - x_i \beta) \right) (y_i - x_i \beta) \right\} \quad (3.24)$$

eşitlik 3.24'teki yöntem uygulanabilir. Prosedür bilinmeyen yoğunluk  $f_{u\theta}(\theta/x)$  için bir tahmin kullanılmasını gerektirir (Newey ve Powel, 1990:295-300).

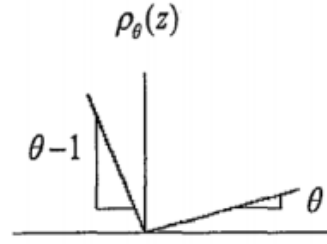
Burada  $I$  karakteristik fonksiyon,  $\rho_{\theta}$  ise kontrol fonksiyonudur ve

$$\rho_{\theta}(z) = z(\theta - I(z \leq 0))$$

veya

$$\rho_{\theta}(z) = \begin{cases} \theta z, & z \geq 0 \\ (\theta - 1)z, & z < 0 \end{cases} \quad (3.25)$$

eşitlik 3.25'deki gibi tanımlanır. Bu fonksiyon;



**Şekil 3. 4** Kantil Regresyon Amaç Fonksiyonu

şekil 3.4.'deki gibi gösterilir (Koenker ve Hallock, 2001:143).  $\theta=0,5$  olması durumunda kantil regresyon amaç fonksiyonu LAD amaç fonksiyonuna ve  $I_1$  regresyona eşittir (Lee, 2004:1). Kantil regresyon amaç fonksiyonu mutlak sapmaların ağırlıklandırılmış toplamıdır.

### 3.3.3. Kantil Regresyon Parametre Tahminleri

Kalıntıların kareleri toplamını minimize eden parametre değerlerini alarak eşitlik 3.26'da gösterildiği gibi  $\hat{\beta}_0$  ve  $\hat{\beta}_1$  için çoklu doğrusal regresyon parametreleri çözer.

$$\min \sum_i (y_i - (\beta_0 + \beta_1 x_i))^2 \quad (3.26)$$

Bu ifade  $y = \hat{\beta}_0 + \hat{\beta}_1 x_i$  doğrusu ve  $(x_i, y_i)$  veri noktaları arasındaki dikey mesafe karelerin toplamıdır. Bu denklemin sıfıra eşitlenerek kısmi türevi alındığında iki bilinmeyenli iki denklem elde edilir. Denklem çözüm sistemleri;

$$\hat{\beta}_1 = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_i^n (x_i - \bar{x})^2} \quad (3.27)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (3.28)$$

eşitlik 3.27 ve 3.28'deki gibidir.

Kantil regresyonda, çoklu doğrusal regresyon parametre tahminlerinin kantil regresyon parametre tahminlerinden farkı bir doğrudan noktaya olan uzaklığın dikey mesafenin toplam etkisi kullanılarak ölçülmesidir. Burada etki doğru üzerindeki noktalar için  $p$  ve tahmini doğru altındaki noktalar için  $(1-p)$  dir.  $p$  oranı için her seçim, örneğin;  $p = .10, .25, .50$  farklı tahmini koşullu- kantil fonksiyona neden olur. Amaç her  $p$  olasılığı için özellik tahmini bulmaktır.

Medyan regresyon modeli için tahmini düşünülürken medyan regresyonda da aynı durum söz konusudur. Saf kalıntıların toplamını minimize etme tercih edilir. Diğer bir ifade ile saf kalıntıların toplamını minimize ederek katsayıları bulunur.  $\beta_s$  için tahmini çözüm

$$\sum_i |y_i - \beta_0 - \beta_1 x_i| \quad (3.29)$$

eşitlik 3.29'da gösterildiği gibidir. Bu ifade minimize edildiğinde çözüm sonucu olarak medyan regresyon doğrusuna bakarız. Regresyon doğrusu üzerinde uydurma kalıntı verilerinin yarısı ve aşağı düşen diğer yarısı ile veri noktalarının bir çifti geçmelidir. Yani kalıntıların yarısı pozitif yarısı negatiftir (Hao ve Naiman, 2007:33).

### 3.3.4. Belirlilik Katsayısı ve Düzeltilmiş Belirlilik Katsayısı

Açıklayıcı sayısı çok olduğu için çoklu korelasyon söz konusudur. Korelasyon katsayısının karesine çoklu belirlilik katsayısı ya da çoklu korelasyon katsayısının karesi adı verilir. Çoklu belirlilik katsayısı  $R^2$  ile eşitlik 3.30'da olduğu gibi gösterilir.

$$R^2_{Y|X_1, X_2, \dots, X_k} = \frac{\hat{\beta}_1 \sum yx_1 + \hat{\beta}_2 \sum yx_2 + \dots + \hat{\beta}_k \sum yx_k}{\sum y^2} \quad (3.30)$$

Fonksiyona yeni değişkenler eklendiği zaman azaldığı apaçık olan serbestlik derecesini hesaba katmak üzere,  $R^2$ 'yi düzeltilir. Düzeltilmiş çoklu belirlilik katsayısının ifadesi

$$R^2 = 1 - \frac{\min \sum_i (x_i \hat{\beta} - \bar{y})^2}{\min \sum_i (y_i - \bar{y})^2} \quad (3.31)$$

eşitlik 3.31'de olduğu gibi gösterilir.

Koenker ve Machado (1999) kantil regresyon modelleri için benzer ölçümü aşağıdaki eşitlik 3.32'deki gösterimi önerir.

$$R^1(\theta) = 1 - \frac{\min \sum_i (y_i - x_i \hat{\beta})^2}{\min \sum_i \rho_\theta(y_i - Q_\theta(y))} \quad (3.32)$$

$R^2$  gibi  $R^1(\theta)$  değeri de (0,1) arasındadır.

### 3.3.5. Eşit Varyans Özellikleri

Kantil regresyon eşit varyans özelliği hesaplama kolaylığı sağlayacak bazı özelliklere sahiptir. Aşağıdaki eşitlik 3.33'te gösterilen minimizasyonda tanımlanan soruna uygun çözümler kümesi  $\beta(\theta, Y, X)$  tarafından ifade edilir.

$$\min_{\beta} \frac{1}{n} \left\{ \sum_{i=1} \left( \theta - \frac{1}{2} + \frac{1}{2} \operatorname{sgn}(y_i - x_i \beta) \right) (y_i - x_i \beta) \right\} \quad (3.33)$$

Bu minimizasyon gösterimi eşitlik 3.34'te olduğu gibi ifade edildiğine göre;

$$\hat{\beta}_{\theta} \equiv \hat{\beta}(\theta, Y, X) \in \beta(\theta, Y, X) \quad (3.34)$$

$$\hat{\beta}(\theta, \lambda y, X) = \lambda \hat{\beta}(\theta, y, X) \quad \lambda \in [0, 1]$$

$$\hat{\beta}(1 - \theta, \lambda y, X) = \lambda \hat{\beta}(\theta, y, X) \quad \lambda \in [0, 1]$$

$$\hat{\beta}(\theta, y + X_{\gamma}, X) = \hat{\beta}(\theta, y, X) + \gamma \quad \gamma \in R^k$$

$$\hat{\beta}(\theta, y, XA) = A^{-1} \hat{\beta}(\theta, y, X) \quad A_{k \times k} \text{ tekil değil}$$

katsayılarının tahmininin eşit varyans özelliklerini gösterir.  $\hat{\beta}_{\theta}$  eşit varyans ölçümüdür. Aşağıda gösterildiği gibi kantil regresyon bir doğrusal programlama sorunudur. Simpleks iterasyonun önemli ölçüde azalan sayısı eşitlikleri ile kullanılabilir (Koenker ve Bassett, 1978:38).

### 3.3.6. Güven Aralıkları

Kantil regresyon parametresi  $\beta(\tau)$  kantil regresyonda güven aralığı hesaplamak için üç metod sağlar: Seyreklik, Rank Metodu ve Yeniden Örnekleme metodlarıdır.

Seyreklik metodu en doğrudan ve hızlı olanıdır. Fakat seyreklik fonksiyonunun tahminini içerir. Bu bağımsız ve özdeş olmayan veriler için robust değildir. Bu problemle uğraşarak kantil regresyon prosedürü seyreklik fonksiyonunun bölgesel tahminini kullanarak Huber Sandwich tahminini hesaplar.

Rank metodu, hata yoğunluklarının doğrudan tahminini önleyen bir metottur. Gutenbrunner, Jureckova, Koenker ve Portnoy tarafından 1993 yılında, bağımsız- özdeş dağılımlı hata terimli model için geliştirilmiştir. Rank metodu, rank skor testi ters çevrilerek

güven aralıklarını hesaplar. Simpleks algoritma kullanır ve büyük veri setleri ile hesaplama zordur. Bu model;

$$y_i = x_i' \beta + e_i \quad (3.35)$$

eşitlik 3.35'deki gibi gösterilir. Daha sonra ise, Koenker ve Machado (1999), bu metodu, konum-ölçek regresyon modelleri için geliştirmişlerdir.

Yeniden örnekleme, bu metot da varyans-kovaryans matrisinin doğrudan tahminini gerektirmemektedir. Yeniden örnekleme metotları; çift örnekleme, tahmin denklemlerini yeniden örnekleme ve Markov Zinciri Marjinal Yeniden Örnekleme şeklinde 3'e ayrılmaktadır.

### **3.3.7. Kovaryans ve Parametre Tahminlerinin Korelasyonu**

Kantil regresyon kovaryans ve korelasyon matrisleri hesaplamak için iki yöntem sağlar. Asimptotik metot ve bootstrap metodu. Yeniden örneklemede güven aralığı hesaplandığında bootstrap kovaryans ve korelasyon matrisi hesaplanır. Aksi takdirde, asimptotik kovaryans ve korelasyon matrisleri hesaplanır.

## BÖLÜM 4

### 4. BOOTSTRAP YÖNTEMİ VE TÜRLERİ

#### 4.1. Bootstrap Yöntemi

Bootstrap yöntemi ilk defa 1979 yılında Efron tarafından ortaya atılmıştır. Bu yöntemde temel düşünce eldeki örnekleme, yığın olarak varsayıp buradan belirli sayıda tekrarlı örnekleme yaparak ilgilenilen tahmincinin yapay bir örnekleme dağılımını oluşturmaktır. Özellikle söz konusu tahmincinin örnekleme dağılımını asimptotik teori ile elde etmek zor ya da olanaksızsa bootstrap yöntemi güçlü bir potansiyel oluşturmaktadır (Aktükün, 2002). Bootstrap yöntemiyle oluşturulan örneklemler, ana örneklemin istatistiksel özelliklerini yansıtacaktır (Smeeke, 2009).

Literatüre bakıldığı zaman birkaç farklı bootstrapa rastlanılabilir. Bunlar, basit bootstrap, çift bootstrap, ağırlıklı bootstrap, tekrarlamalı bootstrap, doğal bootstrap, ardışık bootstrap, ve daha bir çoğu sayılabilmektedir.

$x_1, x_2, \dots, x_n$  birikimli dağılım fonksiyonu  $F(x) = P(x_i \leq x)$  den rastgele çekilmiş örneklem ve  $F$  sonlu (ya da sonsuz) boyutlu dağılımlar ailesi  $F'$ ye ait olsun. Örneklemin bir fonksiyonu  $T_n(x_1, x_2, \dots, x_n)$  istatistiği ele alındığında,  $T_n$  istatistiği genellikle örneğin çekildiği

dağılım olan,  $F$  in bilinmeyen parametresine tahmin olarak kullanılır. Örneğin  $T_n = n^{-1} \sum_{i=1}^n x_i$

örnek ortalaması,  $F$  dağılımının ortalaması  $\mu$  nün tahmin edicisi olarak kullanılır.  $T_n$  istatistiği rastgele verilere bağlı olduğundan bir rastgele değişkendir. Dolayısıyla bir dağılıma sahiptir.

İstatistiğin bu dağılımı  $G_n(x) = P(T_n \leq x)$  ile gösterildiğinde,  $G_n(x)$  dağılımı genellikle  $F(x)$

dağılımına bağlıdır. Bu ilişki genellikle  $G_n(x, F)$  gösterimi ile ifade edilir. Eğer  $F_1, F_2 \in F$  için

$G_n(x, F_1) = G_n(x, F_2)$  oluyorsa  $G_n$  dağılımı  $F$  dağılımına bağlı değildir. Bu şekildeki  $T_n$

istatistiğine  $F$  ailesinde pivotal denir. Örneğin yığın ortalamasına ilişkin hipotez testinde

kullanılan  $t$  istatistiği bilinmeyen yığın parametresinden bağımsız olduğundan pivotal istatistik denir. Fakat ekonometride kullanılan istatistiklerin birçoğu pivotal değildir (Horowitz, 2001).

$T_n$  istatistiğinin dağılımının belirlenebilmesi için  $F$  dağılımının bilinmesi gerekmektedir.

Uygulamada  $F$  dağılımı bilinmediğinden  $G_n$  hesaplanamaz. Dolayısıyla  $G_n$  dağılımı farklı yollarla yaklaşık olarak belirlenmeye çalışılır (Smeeke, 2009). Bu yollardan biri

asimptotik dağılım teorisidir. Ekonometride kullanılan birçok istatistiğin asimptotik dağılımı standart normal ya da ki-kare dağılımına sahiptir(Horowitz, 2001). Böyle istatistiklere asimptotik olarak pivotal denir. Başka bir ifadeyle bu istatistiklerin asimptotik dağılımı bilinmeyen yığın parametresine bağlı değildir.  $T_n$  istatistiğinin asimptotik dağılımı  $G_\infty(x, F) \equiv G_\infty(x)$  şeklinde ifade edilir ve  $F$  dağılımdan bağımsızdır. Bu sebeple  $n$  örnek çapı yeterince büyükse  $G_n(x, F)$  dağılımının tahmini olarak  $G_\infty(x)$  dağılımı kullanılabilir.

Uygulamada kullanılan örnek çapları genellikle, asimptotik dağılımın iyi yaklaşımlar vermesi için yeterince büyük değildir(Smeekes, 2009).  $G_n$  dağılımının bulunmasında kullanılabilir yollardan bir diğeri bootstrap yöntemidir.  $G_n$  dağılımının bulunabilmesi için  $F$  dağılımına ihtiyaç vardır. Bootstrap yönteminin temelindeki düşünce bilinmeyen  $F$  dağılımı yerine tahmini olan  $\hat{F}_n$  dağılımını kullanmaktır. Başka bir ifadeyle  $T_n$  istatistiğinin dağılımı  $G_n(x, F)$  yerine  $G_n(x, \hat{F}_n)$  dağılımı kullanılır. Burada  $G_n(x, \hat{F}_n)$  dağılımına bootstrap dağılımı adı verilmektedir ve genellikle  $G^*(x)$  ile gösterilmektedir. Genel olarak bootstrap kullanılarak elde edilen tahminler üst simge “\*” ifadesiyle gösterilir.

$F$  dağılımının tahmini olan  $\hat{F}_n$  dağılımının nasıl elde edileceği konusunda iki durum mevcuttur. Birincisi, parametrik bootstrap adı verilen ve  $F$  dağılımının hangi dağılıma ait olduğunun bilinmediği ama yığın parametrelerinin tahmininin gerektiği durumdur. Örneğin  $F \approx N(\mu, \sigma^2)$  olduğu bilinmekte fakat parametreleri bilinmemektedir. İkincisi ise parametrik olmayan bootstrap adı verilen ve  $F$  dağılımına ait hiçbir bilginin bilinmediği durumdur. Bu durumda  $F$  dağılımı ampirik dağılım fonksiyonu kullanarak tahmin edilir. Başka bir ifadeyle;

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n I(x_i \leq x) \quad (4.1)$$

eşitlik 4.1’de olduğu gibidir. Burada  $I$  işaret fonksiyonudur. Ampirik dağılım fonksiyonu  $\hat{F}_n(x)$  hiçbir parametrik varsayım olmadığında  $F(x)$  dağılımının en çok olabilirlik tahmin edicisidir (Chernick, 2008). Genellikle  $G_n(x, \hat{F}_n)$  dağılımı analitik yollarla hesaplanamaz ve  $\hat{F}_n$  dağılımı kullanılarak simülasyonlarla yeni örneklemeler elde edilir. Bu şekilde oluşturulan örneklemelere bootstrap örneklemeleri denir ve  $x_1^*, x_2^*, \dots, x_n^*$  ile gösterilir. Bu örneklemeler yardımıyla  $G_n(x, \hat{F}_n)$  dağılımı, bootstrap algoritmasında tarif edildiği gibi elde edilir.

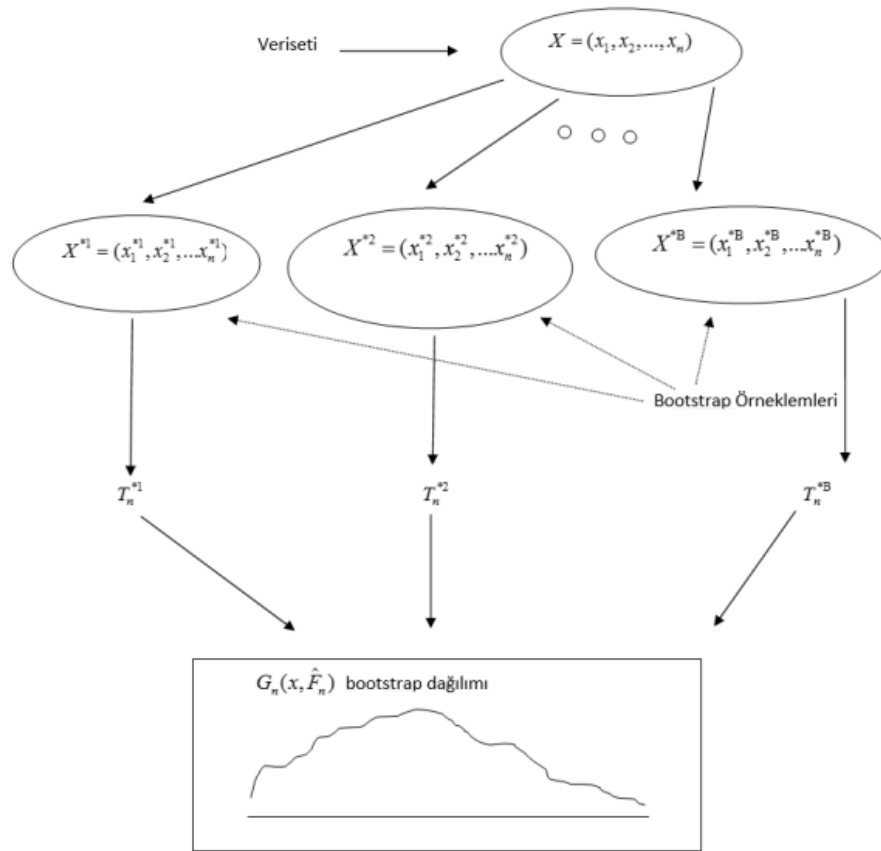
## Genel Bootstrap Algoritması

**Adım 1.**  $x_1, x_2, \dots, x_n$  örnekleminde rastgele örnek çekerek bootstrap örneklemleri  $x_1^*, x_2^*, \dots, x_n^*$  ları elde edilir.

**Adım 2.**  $T_n^* = T_n(x_1^*, x_2^*, \dots, x_n^*)$  istatistiği hesaplanır.

**Adım 3.** Adım 1 ve 2 B kez tekrarlanır. B tekrardan elde edilen  $T_n^{*1}, T_n^{*2}, \dots, T_n^{*B}$  istatistikleri kullanarak  $G_n(x, \hat{F}_n)$  dağılımı elde edilir.

Algoritması verilen bootstrap yöntemini şekil 4.1.deki gibi şemalandırmak mümkündür.



**Şekil 4. 1** Bilinmeyen Parametrenin Dağılımına İlişkin Tahminin Bootstrap Süreci

Bootstrap örneklemlerinin nasıl oluşturulduğunu daha iyi göstermek adına basit bir örnek ele alırsak  $X = (x_1, x_2, x_3, x_4, x_5, x_6, x_7)$  bilinmeyen  $F$  dağılımından rastgele çekilen veriler olsun.  $X = (3;8;1;5;7;9;2)$  değerlerini alsın ve bu dağılıma ait yığın ortalamasının tahmini bootstrap yöntemi ile elde edilmek istensin.

Bootstrap tekrar sayısı  $B = 3$  alınmak üzere bootstrap örneklemleri ve her örnekleme ait yığın ortalaması tahmini şekil 4.1. de verildiği gibidir. Bootstrap örneklemleri oluşturulurken bazı veriler birden fazla kez gözlenebilir. Ayrıca oluşturulabilecek toplam bootstrap

örneklemi sayısı  $n^n$ 'dir. Bu örnek için toplam bootstrap örneklem sayısı  $7^7$ 'dir. Burada sadece 3 bootstrap örnekleme oluşturulmuştur. Oluşturulan bu bootstrap örneklemi tablo 4.1. gösterilmektedir.

**Tablo 4. 1.** Bootstrap Örneklemi

Mevcut Veriler	Bootstrap Örneklemi		
	Örneklem 1	Örneklem 2	Örneklem 3
3	3	7	3
8	1	1	1
1	1	2	9
5	5	2	9
7	7	9	5
9	8	5	8
2	8	5	1
<b>Ortalama</b>	9,42	4,42	5,14

Örnekleme işlemi, yerine koyarak ve rastgele yapıldığından her bir bootstrap örnekleme, mevcut verilerin farklı bir kombinasyonundan oluşur. Dolayısıyla bir bootstrap örnekleme bir veri değerini hiç içermeyen, diğerleri bir, iki ya da daha çok kez içerebilir. Tekrar sayısı artırıldığında örnek ortalamaları bir dağılım oluşturacaktır(Gayaker,2015).

#### 4.1.1. Parametrik Bootstrap Yöntemi

$F$  dağılımına sahip bir anakitleden alınan  $x$  tesadüfi değişkeninin birbirinden bağımsız ve aynı şekilde dağılmış değerlerinin oluşturduğu  $x_1, x_2, \dots, x_n$  örneğine sahip olduğumuzu varsayalım.  $x$ 'in değerleri gerçek sayılar veya sayıların oluşturduğu vektörler olabilir. Anakitleyi tanımlamanın en genel yolu, eşitlik 4.2'de olduğu gibi gösterilen kümülatif dağılım fonksiyonunu yazmaktır.

$$F(x) = Prob(X \leq x) \quad (4.2)$$

Bu fonksiyonun  $x$ 'e göre türevi olasılık yoğunluk fonksiyonunu ( $f(x)$ ) verir ve  $f(x)$  aşağıdaki eşitlik 4.3'teki gibi gösterilir.

$$f(x) = \frac{dF(x)}{dx} \quad (4.3)$$

Gözlemlerimiz için tanımladığımız olasılık yoğunluk fonksiyonu aşağıdaki eşitlik 4.4'te olduğu gibi gösterilir;

$$X : f_{(\theta)}(x) \quad (4.4)$$

Bu ifadede  $\theta$ , daha önce de ifade edildiği gibi  $X$  dağılımından elde edilen bir veya birden fazla bilinmeyen parametreyi gösterir. Bu ifade ayrıca,  $X$  için parametrik model olarak adlandırılır.  $\theta$ 'nın elemanlarının sayısını  $\rho$  ile gösterirsek; örneğin  $X, \mu$  ortalama ve  $\sigma^2$  varyanslı normal dağılıma sahip ise bu durumda,

$$\Theta = (\mu, \sigma^2)$$

olacak ve  $\rho = 2$  elde edilecektir. Dağılımın olasılık yoğunluk fonksiyonu,

$$f_{\theta}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \ell^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (4.5)$$

eşitlik 4.5'te olduğu gibi gösterilir.

Maksimum olabilirlik fonksiyonuna dayandırılır ve aşağıdaki eşitlik 4.6'da olduğu gibi gösterilir:

$$L(\theta; x) = \prod_1^n f_{\theta}(x_i) \quad (4.6)$$

$L(\theta; x)$ ,  $\theta$ 'nin bir fonksiyonu olarak düşünülebilir. Veri kesikli olduğunda  $L(\theta; x)$  gözlemlenen örneğin olasılığıdır. Süreklilik durumunda  $L(\theta; x)\Delta, [x, x+\Delta]$  gibi küçük bir aralıkta yer alan örneğin olasılığı şeklinde açıklanabilir.  $I(\theta; x)$ 'nin algoritmasını,

$$I(\theta; x) = \sum_1^n I(\theta; x_i) \quad (4.7)$$

eşitlik 4.7'deki gibi yazmak mümkündür. Bu  $l(\theta)$  şeklinde özetlenebilir. Bu ifade logaritmik olabilirlik olarak anılır ve her bir değer  $I(\theta; x) = \log f_{\theta}(x_i)$ , bir logaritmik olabilirlik bileşeni olarak adlandırılır. Ayrıca maksimum olabilirlik yönteminde  $I(\theta; x)$ 'i maksimize etmek için  $\theta = \hat{\theta}$  olduğu düşünülür.  $\hat{\theta}$ 'nin örnekleme dağılımını ve varyansını tahmin etmenin en doğru şekli olarak tanımlanan parametrik yonteme göre, öncelikle  $\hat{f}_{\theta}(x)$  olasılık yoğunluk fonksiyonuna göre  $n$  büyüklüğünde  $B$  tane örnek çekilir ve her bir örnek varyansı,  $\hat{\theta}$ 'nin varyansını tahmin etmeye yarar. Bu süreç ise parametrik bootstrap olarak adlandırılır. Standart hatanın parametrik bootstrap tahmini eşitlik 4.8'de gösterildiği gibi şöyledir:

$$se_{\hat{F}}(\hat{\theta}^*) \quad (4.8)$$

Burada  $\hat{F}_{par}$  veriler için elde edilen parametrik modelden türetilen  $F$ 'in tahminidir.  $F$  dağılımının sürekli olduğu varsayıldığında, düzeltilmiş bootstrap dönüşümü yapılmaktadır. Bu aşama bir adım daha ilerletildiğinde, örneğin  $F$  dağılımının Gaussian dağılım gibi parametrik bir formda olduğu varsayıldığında, bu durumda  $F$ 'nin yaklaşık tahmin edicisi de  $\mu$  ve  $\sigma$  'nın maksimum olabilirlik tahminlerine sahip Gaussian dağılım gösterecektir.  $F$  in parametrik tahmininden yapılan iadeli örnekleme sayesinde, Fisher'in teorisi ile ilişkili maksimum olabilirlik tahminlerinin yapıldığı bootstrap tahminlerine ulaşılabilecektir. Parametrik bootstrap yöntemine Monte Carlo simülasyonu ile bakıldığında ise, bunun maksimum olabilirlik tahminine ulaşmaktan başka bir şey olmadığı gözlenir.

#### 4.1.2. Parametrik Olmayan Bootstrap Yöntemi

Parametrik ve parametrik olmayan bootstrap arasındaki en önemli fark, parametrik bootstrap için bir parametrik modelin var olması, parametrik olmayan bootstrap için doğal modelin var olmadığıdır. Bilinmeyen bir  $F$  dağılımından alınan bağımsız ve aynı şekilde dağılmış  $x_1, x_2, \dots, x_n$  değerlerine sahip olduğumuzu ve herhangi bir parametrik modelin var olmadığını düşünelim. Bilinmeyen  $F$  dağılımının kümülatif dağılım fonksiyonunu elde etmek için  $F$  deneysel dağılımını kullanır. Ancak daha önceki açıklamalara dayanarak,  $F$  i sadece parametrik bir model var olduğu takdirde kullanmanın mümkün olduğu söylenebilir. Aksi takdirde, verilerin simülasyonu ve gerekli özelliklerin deneysel hesaplamaları yapılmalıdır. Parametrik olmayan bootstrap yöntemi ile ilgili bir örnek açıklamaları anlaşılır kılacaktır. Ortalama hesaplarken deneysel dağılım fonksiyonundan yapılan örnekleme yardımıyla momentler kolayca bulunabilir. Örneğin,

$$E^*(\bar{X}^*) = E^*(X^*) = \sum_{j=1}^n \frac{1}{n} x_j = \bar{x} \quad (4.9)$$

eşitlik 4.9' da olduğu gibi olur. Benzer şekilde;

$$\begin{aligned} \text{var}^*(\bar{X}^*) &= \frac{1}{n} \text{var}^*(X^*) = \frac{1}{n} E^* \{ X^* - E^*(X^*) \}^2 \\ &= \frac{1}{n} \sum_{j=1}^n \frac{1}{n} (x_j - \bar{x})^2 = \frac{(n-1)}{n} \frac{1}{n(n-1)} \sum_{j=1}^n (x_j - \bar{x})^2 \end{aligned} \quad (4.10)$$

eşitlik 4.10'da gösterildiği gibi yazılır.

Yukarıda verilen şekil 4.9.'da, ilk çarpan hariç,  $X$  'in tahmin edilmiş varyans sonucudur. Bu noktada, deneysel dağılım fonksiyonuna sahip simülasyon uygulaması yapılır.

Deneysel dağılım fonksiyonu orijinal veri grubu olan  $x_1, x_2, \dots, x_n$  kümesindeki değerlerin her birine eşit olasılık verdiği için; her bir  $X^*$ , orijinal örnekten tesadüfi olarak örneklenmiş bağımsız değerler olacaktır. Bu nedenle, simülasyon örneği olan  $x^*_1, x^*_2, \dots, x^*_n$ , orijinal verilerden iadeli olarak alınan tesadüfi bir örnek olacaktır. Burada kolaylık sağlayan, verilerin homojen olmasıdır. Bu yeniden örnekleme yöntemi, parametrik olmayan bootstrap olarak bilinir.

## 4.2. Regresyon Analizinde Kullanılan Bootstrap Yöntemleri

En küçük kareler regresyon analizinin temelleri, hata teriminin ( $\epsilon$ ) analizine dayanır. En küçük kareler regresyon analizinde hata terimlerinin tekrarlanması ile uygulanan bu yöntem, 1979 yılında Bradley Efron tarafından ileri sürülmüş ve klasik en küçük kareler yönteminden daha etkili parametre tahminleri elde etmek amacı ile geliştirilmiştir ve hata teriminin yeniden örneklenmesi olarak bilinir.

$\beta$ 'nin bir bootstrap tahmin edicisini elde etmek için algoritma aşağıdaki şekilde izlenir:

- I. Populasyondan şansa bağlı olarak  $n$  sayıda bir örnek seçilir.
- II. Seçilen bu örneğe ait EKK regresyon doğrusu oluşturulur.
- III. Bu modelden  $e_i$  değerleri hesaplanır.
- IV. Elde edilen  $e_i$  değerlerine  $1/n$  olasılığı verilerek her biri  $n$  hacminde  $B$  tane bootstrap hata alt örnekleri oluşturulur. Böylece deneysel dağılım fonksiyonu  $(\hat{F}_\epsilon(x)), \hat{F}_\epsilon(x) = \left\{ \frac{\hat{e}_i^2 \leq x}{n} \right\}$  şeklinde elde edilir.
- V. Oluşan bu deneysel dağılım fonksiyonundan bootstrap hata değerlerinin

ortalaması,  $\bar{\hat{\epsilon}}_i^* = \frac{\sum_{b=1}^B \hat{\epsilon}_{bi}}{B}$  şeklinde hesaplanır. Burada,

$\bar{\hat{\epsilon}}_i^*$  :  $i$ 'nci bootstrap hata tahmin edicisi

$\hat{\epsilon}_{bi}$  :  $b$ 'nci bootstrap örneğine ait  $i$ 'nci hata tahmin edicisi

- VI. Elde edilen  $\bar{\hat{\epsilon}}_i^*$  değerleri 2. adımda oluşturulan modeldeki  $e_i$ 'ler yerine konarak  $Y_i^* = \hat{\beta}X + \bar{\hat{\epsilon}}_i^*$  şeklinde bootstrap  $Y^*$  değerleri hesaplanır.
- VII.  $Y^*$  ve  $X$ 'den hareketle  $\beta$ 'nin bootstrap tahmin edicisi, EKK yöntemi ile,  $\hat{\beta}^* = (X'X)^{-1} X'Y^*$

şeklinde hesaplanır(Topuz,2002). Elde edilen bu tahmin edici sapmasız olup;

$$E(\hat{\beta}^*) = (X'X)^{-1} X'E(Y^*) = \hat{\beta} \quad (4.11)$$

eşitlik 4.11'deki gibi olmaktadır. Bootstrap yönteminin bu uygulamasında hata terimlerinden tekrarlı örnekler seçilerek, tahmin değerlerine  $(\hat{Y}_i)$  eklendiği için hata terimlerinin normal dağılıma sahip olduğu varsayılmaktadır.

Bootstrap yönteminde, hata terimlerinin dağılımına ilişkin varsayım yapılmamasına rağmen,  $Y^*$  'lar doğrusal modele göre oluşturularak, modelin fonksiyonel biçiminin doğrusal olduğu varsayılır. Bunun yanında hata terimlerinin normalliği varsayımından hareket edildiği için gerçek hatalar sabit varyansa sahip değilse, bu özellik alt örnek hata terimlerine de yansımacaktır(Fox,1997).

En küçük kareler tahmincileri üzerinde herhangi bir üstünlüğe sahip olmamasına rağmen, regresyon analizinde bootstrap tahmin edicilerinin etkinlikleri varsayımından sapmalardan etkilenmez. Bootstrap yöntemlerinin güçlüğü, sabit varyans varsayımının varsayıp varsayılmadığına bağlıdır(Altaş,2000).

Bootstrap yönteminin doğrusal regresyon analizine uygulanmasında iki farklı yöntem izlenmektedir. Hangi yöntemin daha etkili olduğu ise farklı durumlar için değişmektedir. Bunlarda birincisi tahmin edilen modelden elde edilen hata terimleri kullanılarak yapılan bootstrap, ikincisi ise bağımsız değişken ve bağımlı değişken çiftine uygulanan bootstrap yöntemidir.

#### **4.2.1. Hata Terimlerine Dayanan Bootstrap Yöntemi**

Basit doğrusal regresyon modeli eşitlik 4.1.' de parametrelerin tahminine ilişkin güven aralığı oluşturmak için modelin artıklarına bootstrap yöntemi uygulanmaktadır. Bu yöntemin uygulanabilmesi için hataların birbirinden bağımsız ve aynı dağılıma sahip olması gerekmektedir. Eğer hata terimlerinde değişen varyans sorunu varsa artıklara dayanan bootstrap yöntemiyle oluşturulan örneklem gerçekte örneklem karakterini yansıtmayacaktır. Bu durumda çift bootstrap yönteminin uygulanması daha doğru sonuçlar verecektir (Cameron ve Trivedi, 2005).

#### **4.2.2. Çift Bootstrap Yöntemi**

Bu yöntem eldeki veri setine direkt bootstrap uygulanmasına dayanmaktadır. Doğrusal regresyon analizinde bootstrap uygulanırken iki farklı yöntem izlenebileceği görülmektedir.

Hangi yöntemin daha etkili olduğu ise farklı durumlar için değişmektedir. Regresyon analizinde açıklayıcı değişken  $x$  in sabit olduğu varsayımı vardır. Çift(pairs) bootstrap yönteminde ise algoritma her çalıştığında açıklayıcı değişken  $x$  değişmektedir. Bu ise  $x$ 'i sabit olmasından çok rastgele bir duruma getirmektedir (Cameron ve Trivedi, 2005).



## BÖLÜM 5

### 5. UYGULAMA

Uygulamada 2000-2017 yılları arası aylık Üfe, Üfe(-2) dönem gecikmesi, beklenti anket sonuçları ve dolar kuru verileri kullanılmıştır. Analizler için STATA 14 ve GRETL paket programlarından yararlanılmıştır. Uygulama verilerine bootstrap yöntemi kullanılarak gözlem sayısı % 25 'den % 200 'e kadar arttırılmış olup oluşan yeni verilere Doğrusal Regresyon ve Kantil Regresyon Modeli uygulanmıştır. Verilerin kantil regresyona uygun olup olmadığını belirlemek için Kolmogorov Smirnov Testi uygulanmış ve verilerin normal dağılmadığı( $\rho=0.00$ ) belirlenmiştir. Uygulamadan ilk olarak, belirlenen dört değişken için doğrusal regresyon yöntemiyle katsayıların anlamlı olup olmadığı Tablo 5.1.'de verilmiştir. Daha sonra hangi modelin daha iyi sonuçlar verdiğini Ortalama Mutlak Sapma(OMS) ve Hata Karelerinin Ortalaması Karekökü(HKOK) sonuçlarını karşılaştırarak uygun model belirlenmiştir.

**Tablo 5. 1.** Doğrusal Regresyon Modeli Analiz Sonuçları

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ÜFE_2	-.1867573	.0559931	-3.34	0.001	-.2971348 -.0763797
BEKLENTİ	1.229728	.0778522	15.80	0.000	1.076261 1.383196
DOLAR	-.0543506	.0839559	-0.65	0.518	-.2198504 .1111492
cons	.1349631	.1829093	0.74	0.461	-.2256006 .4955268

Tablo 5.1.'e baktığımızda Üfe(-2) ve Beklenti değişkenlerinin katsayıları anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.1'de verilmiştir.

$$\dot{UFE} = 0,135 - 0,1868\dot{UFE}(-2) + 1,2297BEKLENTİ - 0,0544DOLAR \quad (5.1)$$

Dolar değişkeni bootstrap uygulanmadan önce ve sonra da hem Kantil( $Q_1, Q_2, Q_3$  için) hem de Doğrusal Regresyon hesaplamalarındaki bütün sonuçlarda anlamsız çıktığı için uygulamadan çıkartılmıştır.

Bootstrap yöntemi ile veri sayısını arttırmadan önce verilere Doğrusal ve Kantil Regresyon yöntemleriyle hangi modelin daha iyi sonuç verdiğini OMS ve HKOK değerlerini karşılaştırarak başlangıçta hangi modelin tercih edildiği gösterilmiştir.

**Tablo 5. 2.** Doğrusal Regresyon Modeli Analiz Sonuçları

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ÜFE_2	-.1821466	.0554622	-3.28	0.001	-.2914746 - .0728187
BEKLENTİ	1.235237	.0772797	15.98	0.000	1.082902 1.387572
cons	.0293358	.0825514	0.36	0.723	-.1333908 .1920624

Tablo.5.2.'ye baktığımızda Üfe(-2) ve Beklenti değişkenlerinin katsayıları anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.2'de verilmiştir.

$$\ddot{ÜFE}=0,029-0,1821\ddot{ÜFE}(-2)+1,2352BEKLENTİ \quad (5.2)$$

**Tablo 5. 3.** Kantil Regresyon( $Q_1$ ) Modeli Analiz Sonuçları

0.25 Quantile regression	Number of obs =	215
Raw sum of deviations	67.41 (about .30000001)	
Min sum of deviations	43.94736	Pseudo R2 = 0.3481

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ÜFE_2	-.2135467	.0509083	-4.19	0.000	-.3138979 - .1131955
BEKLENTİ	1.150641	.0709344	16.22	0.000	1.010814 1.290468
cons	-.3386122	.0757732	-4.47	0.000	-.4879777 - .1892468

Tablo 5.3.'e baktığımızda  $Q_1$  kantiline göre bütün değişkenlerin katsayıları anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.3'te verilmiştir.

$$Q_1(\ddot{ÜFE})= -0,339-0,2135\ddot{ÜFE}(-2)+1,1506BEKLENTİ \quad (5.3)$$

**Tablo 5. 4.** Kantil Regresyon( $Q_2$ ) Modeli Analiz Sonuçları

Median regression		Number of obs =	215
Raw sum of deviations	100.14 (about .75999999)		
Min sum of deviations	59.68624	Pseudo R2 =	0.4040

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ÜFE_2	-.1416481	.0634155	-2.23	0.027	-.2666538 - .0166425
BEKLENTİ	1.129918	.0883616	12.79	0.000	.9557385 1.304098
cons	-.0081515	.0943893	-0.09	0.931	-.1942132 .1779103

Tablo 5.4.'te baktığımızda  $Q_2$ (Medyan) kantiline göre ise Üfe(-2) ve Beklenti değişkenlerinin anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.4'te verilmiştir.

$$Q_2(\ddot{U}FE) = -0,008 - 0,1416\ddot{U}FE(-2) + 1,1299BEKLENTİ \quad (5.4)$$

**Tablo 5. 5.** Kantil Regresyon( $Q_3$ ) Modeli Analiz Sonuçları

0.75 Quantile regression		Number of obs =	215
Raw sum of deviations	96.72 (about 1.7)		
Min sum of deviations	50.62294	Pseudo R2 =	0.4766

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ÜFE_2	-.1309154	.0556457	-2.35	0.020	-.2406051 - .0212256
BEKLENTİ	1.425496	.0775354	18.39	0.000	1.272656 1.578335
cons	.1852988	.0828245	2.24	0.026	.0220336 .3485639

Tablo 5.5.'e baktığımızda  $Q_3$  kantil değerine göre bütün değişkenlerin katsayıları anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.5'te verilmiştir.

$$Q_3(\ddot{U}FE) = 0,185 - 0,1309\ddot{U}FE(-2) + 1,4254BEKLENTİ \quad (5.5)$$

**Tablo 5. 6.** Doğrusal Regresyon ve Kantil Regresyon Modeli OMS ve HKOK Analiz Sonuçları

REGRESYON	OMS	HKOK
Doğrusal	<b><u>0.695829</u></b>	<b><u>0.834164</u></b>
Kantil( $Q_1$ )	0.955382	0.977436
Kantil( $Q_2$ )	0.71272	0.844227
Kantil( $Q_3$ )	0.934517	0.966704

Tablo 5.6.'ya bakıldığında OMS ve HKOK değerleri karşılaştırıldığında en küçük değere sahip olan Doğrusal Regresyon yöntemiyle hesaplanan modelin en iyi model olarak seçilebileceği görülmektedir.

Bundan sonraki uygulamalarda ise Bootstrap yardımıyla veri setini %25, %50, %75, %100 ve %200 arttırarak uygulamaları yapıp hangi modelin daha iyi sonuçlar verdiği incelenmiştir.

Veri setimizi **%25** arttırdığımız durumdaki sonuçları aşağıda incelenmiştir.

**Tablo 5. 7.** Bootstrap(%25) - Doğrusal Regresyon Modeli Analiz Sonuçları

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ÜFE_2	.0476234	.0381275	1.25	0.213	-.0274453 .1226922
BEKLENTİ	1.038761	.0512604	20.26	0.000	.9378353 1.139687
cons	-.0428114	.0782394	-0.55	0.585	-.1968561 .1112332

Tablo.5.7.'ye baktığımızda sadece Beklenti değişkeninin katsayısının anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.6'da verilmiştir.

$$\hat{ÜFE} = -0,043 - 0,0476\hat{ÜFE}(-2) + 1,0387\hat{BEKLENTİ} \quad (5.6)$$

**Tablo 5. 8.** Bootstrap(%25) - Kantil Regresyon( $Q_1$ ) Modeli Analiz Sonuçları

0.25 Quantile regression		Number of obs =	270
Raw sum of deviations	78.6 (about .28999999)		
Min sum of deviations	55.8307	Pseudo R2 =	0.2897

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ÜFE_2	.0511755	.0443477	1.15	0.250	-.0361401 .1384911
BEKLENTİ	.9179069	.0596231	15.40	0.000	.8005156 1.035298
cons	-.4341981	.0910036	-4.77	0.000	-.613374 -.2550223

Tablo 5.8.'e baktığımızda  $Q_1$  kantil değerine göre Sabit katsayının ve Beklenti değişkeninin katsayıları anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.7'de verilmiştir.

$$Q_1(\ddot{U}FE) = -0,434 + 0,0512\ddot{U}FE(-2) + 0,9179BEKLENTİ \quad (5.7)$$

**Tablo 5. 9.** Bootstrap(%25) - Kantil Regresyon( $Q_2$ ) Modeli Analiz Sonuçları

Median regression		Number of obs =	270
Raw sum of deviations	113.87 (about .75999999)		
Min sum of deviations	74.38752	Pseudo R2 =	0.3467

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ÜFE_2	.0332344	.0546014	0.61	0.543	-.0742697 .1407385
BEKLENTİ	1.047764	.0734087	14.27	0.000	.9032302 1.192297
cons	-.1004928	.1120447	-0.90	0.371	-.3210964 .1201108

Tablo 5.9.'a baktığımızda  $Q_2$  kantil değerine göre sadece Beklenti değişkeninin katsayısının anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.8'de verilmiştir.

$$Q_2(\ddot{U}FE) = -0,101 + 0,0332\ddot{U}FE(-2) + 1,0477BEKLENTİ \quad (5.8)$$



Veri setimizi **%50** arttırdığımız durumdaki sonuçları aşağıda incelenmiştir.

**Tablo 5. 12.** Bootstrap(%50) - Doğrusal Regresyon Modeli Analiz Sonuçları

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ÜFE_2	.0141685	.0329135	0.43	0.667	-.0505841	.0789211
BEKLENTİ	.9979871	.040782	24.47	0.000	.9177544	1.07822
cons	.0696399	.0737469	0.94	0.346	-.0754466	.2147264

Tablo 5.12.'ye baktığımızda sadece Beklenti değişkeninin katsayısının anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.10'da verilmiştir.

$$\dot{ÜFE} = 0,069 + 0,0141\dot{ÜFE}(-2) + 0,9978BEKLENTİ \quad (5.10)$$

**Tablo 5. 13.** Bootstrap(%50) - Kantil Regresyon( $Q_1$ ) Modeli Analiz Sonuçları

0.25 Quantile regression	Number of obs =	325
Raw sum of deviations 112.1225 (about .28999999)		
Min sum of deviations 74.07721	Pseudo R2 =	0.3393

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ÜFE_2	.0354445	.0371901	0.95	0.341	-.0377219	.1086109
BEKLENTİ	.8802394	.046081	19.10	0.000	.7895815	.9708973
cons	-.339613	.0833293	-4.08	0.000	-.5035516	-.1756745

Tablo 5.13.'te baktığımızda  $Q_1$  kantil değerine göre sabit parametre ve Beklenti değişkeninin katsayısının anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.11'de verilmiştir.

$$Q_1(\dot{ÜFE}) = -0,339 + 0,0354\dot{ÜFE}(-2) + 0,8802BEKLENTİ \quad (5.11)$$

**Tablo 5. 14. Bootstrap(%50) - Kantil Regresyon( $Q_2$ ) Modeli Analiz Sonuçları**

Median regression					Number of obs =	325
Raw sum of deviations	162.215	(about .82999998)				
Min sum of deviations	95.11556				Pseudo R2 =	0.4136

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ÜFE_2	.013289	.0376273	0.35	0.724	-.0607373 .0873154
BEKLENTİ	.9242525	.0466227	19.82	0.000	.832529 1.015976
cons	.1275748	.0843088	1.51	0.131	-.0382908 .2934403

Tablo 5.14.'e baktığımızda  $Q_2$  kantil değerine göre sadece beklenti değişkeninin katsayısının anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.12'de verilmiştir.

$$Q_2(\dot{U}FE) = 0,128 + 0,0132\dot{U}FE(-2) + 0,9242BEKLENTİ \quad (5.12)$$

**Tablo 5. 15. Bootstrap(%50) - Kantil Regresyon( $Q_3$ ) Modeli Analiz Sonuçları**

0.75 Quantile regression					Number of obs =	325
Raw sum of deviations	149.57	(about 1.8200001)				
Min sum of deviations	78.25403				Pseudo R2 =	0.4768

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ÜFE_2	-.0069116	.0373204	-0.19	0.853	-.0803342 .0665109
BEKLENTİ	1.239214	.0462424	26.80	0.000	1.148239 1.33019
cons	.3251124	.0836211	3.89	0.000	.1605998 .4896251

Tablo 5.15.'e baktığımızda  $Q_3$  kantil değerine göre sabit parametre ve beklenti değişkeninin katsayısının anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.13'te verilmiştir.

$$Q_3(\dot{U}FE) = 0,325 - 0,0069\dot{U}FE(-2) + 1,2392BEKLENTİ \quad (5.13)$$

**Tablo 5. 16.** Bootstrap(%50) - Doğrusal Regresyon ve Kantil Regresyon Modeli OMS ve HKOK Analiz Sonuçları

REGRESYON	OMS	HKOK
Doğrusal	<u>0.692931</u>	<u>0.832425</u>
Kantil( $Q_1$ )	0.974558	0.987197
Kantil( $Q_2$ )	0.700581	0.837007
Kantil( $Q_3$ )	1.013169	1.006563

Tablo 5.16.'ya bakıldığımızda veri setinin %50 arttırdığımız takdirde OMS ve HKOK değerleri karşılaştırıldığında en küçük değere sahip olan Doğrusal Regresyon yöntemiyle hesaplanan modelin en iyi model olarak seçilebileceği görülmektedir.

Veri setimizi **%75** arttırdığımız durumdaki sonuçları aşağıda incelenmiştir.

**Tablo 5. 17.** Bootstrap(%75) - Doğrusal Regresyon Modeli Analiz Sonuçları

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ÜFE_2	-.0057486	.0281088	-0.20	0.838	-.0610206	.0495234
BEKLENTİ	1.013847	.0339111	29.90	0.000	.9471657	1.080529
cons	-.0472798	.0627502	-0.75	0.452	-.1706694	.0761098

Tablo 5.17.'ye baktığımızda sadece beklenti değişkeninin katsayısının anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.14'te verilmiştir.

$$\hat{ÜFE} = -0,047 - 0,0057\hat{ÜFE}(-2) + 1,0138\hat{BEKLENTİ} \quad (5.14)$$

**Tablo 5. 18. Bootstrap(%75) - Kantil Regresyon( $Q_1$ ) Modeli Analiz Sonuçları**

0.25 Quantile regression	Number of obs =	375
Raw sum of deviations 119.2275 (about .22)		
Min sum of deviations 75.80109	Pseudo R2 =	0.3642

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ÜFE_2	-.0143146	.0386154	-0.37	0.711	-.0902465 .0616172
BEKLENTİ	.9042894	.0465865	19.41	0.000	.8126834 .9958953
cons	-.4349029	.0862053	-5.04	0.000	-.6044136 -.2653922

Tablo 5.18.'e baktığımızda  $Q_1$  kantil değerine göre sabit parametre ve beklenti değişkeninin katsayısının anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.15'te verilmiştir.

$$Q_1(\ddot{U}FE) = -0,435 - 0,0143\ddot{U}FE(-2) + 0,9042BEKLENTİ \quad (5.15)$$

**Tablo 5. 19. Bootstrap(%75) - Kantil Regresyon( $Q_2$ ) Modeli Analiz Sonuçları**

Median regression	Number of obs =	375
Raw sum of deviations 179.305 (about .69999999)		
Min sum of deviations 99.15719	Pseudo R2 =	0.4470

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ÜFE_2	-.0232165	.0303652	-0.76	0.445	-.0829255 .0364926
BEKLENTİ	1.078696	.0366334	29.45	0.000	1.006661 1.15073
cons	-.1557098	.0677876	-2.30	0.022	-.2890046 -.0224149

Tablo 5.19.'a baktığımızda  $Q_2$  kantil değerine göre sabit parametre ve beklenti değişkeninin katsayısının anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.16'da verilmiştir.

$$Q_2(\ddot{U}FE) = -0,156 - 0,0232\ddot{U}FE(-2) + 1,0787BEKLENTİ \quad (5.16)$$



Veri setimizi **%100** arttırdığımız durumdaki sonuçları aşağıda incelenmiştir.

**Tablo 5. 22.** Bootstrap(%100) - Doğrusal Regresyon Modeli Analiz Sonuçları

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ÜFE_2	.0035855	.0282677	0.13	0.899	-.0519756	.0591467
BEKLENTİ	1.007525	.0350757	28.72	0.000	.9385826	1.076468
cons	-.0019507	.0645515	-0.03	0.976	-.1288289	.1249275

Tablo 5.22.'ye baktığımızda sadece beklenti değişkeninin katsayısının anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.18'de verilmiştir.

$$\dot{ÜFE} = -0,002 + 0,0036\dot{ÜFE}(-2) + 1,0075BEKLENTİ \quad (5.18)$$

**Tablo 5. 23.** Bootstrap(%100) - Kantil Regresyon( $Q_1$ ) Modeli Analiz Sonuçları

0.25 Quantile regression	Number of obs =	430
Raw sum of deviations 134.3925 (about .30000001)		
Min sum of deviations 86.82341	Pseudo R2 =	0.3540

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ÜFE_2	-.0067293	.0341031	-0.20	0.844	-.0737601	.0603015
BEKLENTİ	.8646917	.0423165	20.43	0.000	.7815172	.9478663
cons	-.3007904	.077877	-3.86	0.000	-.4538604	-.1477204

Tablo 5.23.'e baktığımızda  $Q_1$  kantil değerine göre sabit parametre ve beklenti değişkeninin katsayısının anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.19'da verilmiştir.

$$Q_1(\dot{ÜFE}) = -0,301 - 0,0067\dot{ÜFE}(-2) + 0,8647BEKLENTİ \quad (5.19)$$

**Tablo 5. 24. Bootstrap(%100) - Kantil Regresyon( $Q_2$ ) Modeli Analiz Sonuçları**

Median regression	Number of obs =	430
Raw sum of deviations	200.205 (about .73000002)	
Min sum of deviations	112.3893	Pseudo R2 = 0.4386

---

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ÜFE_2	-.0015892	.0262697	-0.06	0.952	-.0532232 .0500448
BEKLENTİ	1.043559	.0325965	32.01	0.000	.9794893 1.107628
cons	-.0858464	.0599889	-1.43	0.153	-.2037566 .0320638

---

Tablo 5.24.'e baktığımızda  $Q_2$  kantil değerine göre sadece beklenti değişkeninin katsayısının anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.20'de verilmiştir.

$$Q_2(\dot{U}FE) = -0,086 - 0,0016\dot{U}FE(-2) + 1,0436BEKLENTİ \quad (5.20)$$

**Tablo 5. 25. Bootstrap(%100) - Kantil Regresyon( $Q_3$ ) Modeli Analiz Sonuçları**

0.75 Quantile regression	Number of obs =	430
Raw sum of deviations	193.9175 (about 1.7)	
Min sum of deviations	98.40054	Pseudo R2 = 0.4926

---

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ÜFE_2	-.0072201	.0346729	-0.21	0.835	-.0753709 .0609308
BEKLENTİ	1.213605	.0430236	28.21	0.000	1.129041 1.298169
cons	.1586572	.0791783	2.00	0.046	.0030294 .3142849

---

Tablo 5.25.'e baktığımızda  $Q_3$  kantil değerine göre sabit parametre ve beklenti değişkeninin katsayısının anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.21'de verilmiştir.

$$Q_3(\dot{U}FE) = 0,157 - 0,0072\dot{U}FE(-2) + 1,2136BEKLENTİ \quad (5.21)$$

**Tablo 5. 26.** Bootstrap(%100) - Doğrusal Regresyon ve Kantil Regresyon Modeli OMS ve HKOK Analiz Sonuçları

REGRESYON	OMS	HKOK
Doğrusal	<b><u>0.661196</u></b>	<b><u>0.81314</u></b>
Kantil( $Q_1$ )	0.907959	0.952869
Kantil( $Q_2$ )	0.665344	0.815686
Kantil( $Q_3$ )	0.858653	0.926636

Tablo 5.26.'ya bakıldığımızda veri setinin %100 arttırdığımız takdirde OMS ve HKOK değerleri karşılaştırıldığında en küçük değere sahip olan Doğrusal Regresyon yöntemiyle hesaplanan modelin en iyi model olarak seçilebileceği görülmektedir.

Veri setimizi son olarak **%200** arttırdığımız durumdaki sonuçları aşağıda incelenmiştir.

**Tablo 5. 27.** Bootstrap(%200) - Doğrusal Regresyon Modeli Analiz Sonuçları

ÜFE	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ÜFE_2	.0118092	.0228794	0.52	0.606	-.0331184	.0567367
BEKLENTİ	1.155121	.0323508	35.71	0.000	1.091595	1.218647
cons	-.0760223	.0558389	-1.36	0.174	-.1856712	.0336265

Tablo 5.27.'ye baktığımızda sadece beklenti değişkeninin katsayısının anlamlı olduğu görülmektedir. Elde edilen model eşitlik 5.22'de verilmiştir.

$$\hat{ÜFE} = -0,076 + 0,0118\hat{ÜFE}(-2) + 1,1551\hat{BEKLENTİ} \quad (5.22)$$





## BÖLÜM 6

### 6. SONUÇ VE ÖNERİLER

Ekonometrik ve İstatistiksel çalışmalarda değişkenler arasındaki ilişkiyi daha iyi açıklayabilmek için farklı regresyon modelleri, verinin yapısına göre tercih edilmektedir. Bu çalışmada ise Doğrusal Regresyon ile Kantil Regresyon modeli karşılaştırması yapılmıştır. Gretl programı yardımı ile de Bootstrap yöntemi uygulanarak veri sayısı değiştirilerek her veri için Doğrusal ve Kantil Regresyon modeli sonuçlarından hangi modelin ilişkiyi daha iyi açıkladığı incelenmiştir. Elde edilen sonuçlar aşağıda özetlenmiştir.

İlk olarak gerçek veri değerlerimize gerçekleştirmiş olduğumuz Doğrusal ve Kantil Regresyon sonuçlarımıza göre;

Doğrusal Regresyon Modelin' de Üfe(-2) dönem gecikmesi parametresi ile Beklenti parametresi katsayıları anlamlı iken sabit katsayı parametresi anlamsız çıkmıştır.

Kantil Regresyon Modelin' de ise  $Q_1, Q_2$  ve  $Q_3$  kantil değerlerinde;

$Q_1$  Kantil için bütün parametreler anlamlı çıkmıştır.

$Q_2$  Kantil için ise Doğrusal Model'de olduğu gibi Üfe(-2) dönem gecikmesi parametresi ile Beklenti Anketi parametresi katsayıları anlamlı iken Sabit katsayı parametresi anlamsız çıkmıştır.

$Q_3$  Kantil için ise bütün parametreler anlamlı çıkmıştır.

Sonuç olarak modellerin OMS ve HKOK değerleri karşılaştırıldığında ise en küçük ve birbirlerine çok yakın iki sonuç olan Doğrusal Regresyon ve Kantil Regresyon( $Q_2$ ) Modeli en uygun modeller olarak analiz edilmiştir.

Veriyi Bootstrap ile %25 oranında arttırdığımız zaman Doğrusal ve Kantil Regresyon sonuçlarımıza göre;

Doğrusal Regresyon Modelin' de sadece Beklenti Anketi parametresi katsayısı anlamlı iken Üfe(-2) dönem gecikmesi parametresi ile sabit katsayı parametresi anlamsız çıkmıştır.

Kantil Regresyon Modelin' de ise  $Q_1, Q_2$  ve  $Q_3$  kantil değerlerinde;

$Q_1$  kantil için Sabit katsayı parametresi ve Beklenti Anketi parametresi katsayısı anlamlı iken Üfe(-2) dönem gecikmesi parametresi anlamsız çıkmıştır.

$Q_2$  ve  $Q_3$  kantilleri için ise Doğrusal Modelde olduğu gibi sadece Beklenti Anketi parametresi katsayısı anlamlı iken Üfe(-2) dönem gecikmesi parametresi ile Sabit katsayı parametresi anlamsız çıkmıştır.

Sonuç olarak modellerin OMS ve HKOK değerleri karşılaştırıldığı da ise yine en küçük ve birbirlerine çok yakın iki sonuç olan Doğrusal Regresyon ve Kantil Regresyon( $Q_2$ ) Modeli en uygun modeller olarak analiz edilmiştir.

Veriyi Bootstrap ile %50 oranın da arttırdığımız zaman Doğrusal ve Kantil Regresyon sonuçlarımıza göre;

Doğrusal Regresyon Modelin' de sadece Beklenti Anketi parametresi katsayısı anlamlı iken Üfe(-2) dönem gecikmesi parametresi ile Sabit katsayı parametresi anlamsız çıkmıştır.

Kantil Regresyon Modelin' de ise  $Q_1, Q_2$  ve  $Q_3$  kantil değerlerinde;

$Q_1$  kantil için Sabit katsayı parametresi ve Beklenti Anketi parametresi katsayısı anlamlı iken Üfe(-2) dönem gecikmesi parametresi anlamsız çıkmıştır.

$Q_2$  kantil için ise Doğrusal Modelde olduğu gibi sadece Beklenti Anketi parametresi katsayısı anlamlı iken Üfe(-2) dönem gecikmesi parametresi ile Sabit katsayı parametresi anlamsız çıkmıştır.

$Q_3$  kantilde ise  $Q_1$  kantilde olduğu gibi Sabit katsayı parametresi ve Beklenti Anketi parametresi katsayısı anlamlı iken Üfe(-2) dönem gecikmesi parametresi anlamsız çıkmıştır.

Sonuç olarak modellerin OMS ve HKOK değerleri karşılaştırıldığı da ise en küçük ve birbirlerine çok yakın iki sonuç olan Doğrusal Regresyon ve Kantil Regresyon( $Q_2$ ) Modeli en uygun modeller olarak analiz edilmiştir.

Veriyi Bootstrap ile %75 oranın da arttırdığımız zaman Doğrusal ve Kantil Regresyon sonuçlarımıza göre;

Doğrusal Regresyon Modelin' de sadece Beklenti Anketi parametresi katsayısı anlamlı iken Üfe(-2) dönem gecikmesi parametresi ile Sabit katsayı parametresi anlamsız çıkmıştır.

Kantil Regresyon Modelin' de ise  $Q_1, Q_2$  ve  $Q_3$  kantil değerlerinde;

$Q_1$  kantil için Sabit katsayı parametresi ve Beklenti Anketi parametresi katsayısı anlamlı iken Üfe(-2) dönem gecikmesi parametresi anlamsız çıkmıştır.

$Q_2$  kantilde  $Q_1$  kantildeki gibi Sabit katsayı parametresi ve Beklenti Anketi parametresi katsayısı anlamlı iken Üfe(-2) dönem gecikmesi parametresi anlamsız çıkmıştır.

$Q_3$  kantil için ise Doğrusal Modelde olduğu gibi sadece Beklenti Anketi parametresi katsayısı anlamlı iken Üfe(-2) dönem gecikmesi parametresi ile Sabit katsayı parametresi anlamsız çıkmıştır.

Sonuç olarak modellerin OMS ve HKOK değerleri karşılaştırıldığı da ise en küçük ve birbirlerine çok yakın iki sonuç olan Doğrusal Regresyon ve Kantil Regresyon( $Q_2$ ) Modeli en uygun modeller olarak analiz edilmiştir.

Veriyi Bootstrap ile %100 oranın da arttırdığımız zaman Doğrusal ve Kantil Regresyon sonuçlarımıza göre;

Doğrusal Regresyon Modelin' de sadece Beklenti Anketi parametresi katsayısı anlamlı iken Üfe(-2) dönem gecikmesi parametresi ile Sabit katsayı parametresi anlamsız çıkmıştır.

Kantil Regresyon Modelin' de ise  $Q_1, Q_2$  ve  $Q_3$  kantil değerlerinde;

$Q_1$  kantil için Sabit katsayı parametresi ve Beklenti Anketi parametresi katsayısı anlamlı iken Üfe(-2) dönem gecikmesi parametresi anlamsız çıkmıştır.

$Q_2$  kantil için ise Doğrusal Modelde olduğu gibi sadece Beklenti Anketi parametresi katsayısı anlamlı iken Üfe(-2) dönem gecikmesi parametresi ile Sabit katsayı parametresi anlamsız çıkmıştır.

$Q_3$  kantilde ise  $Q_1$  kantil' de olduğu gibi Sabit katsayı parametresi ve Beklenti Anketi parametresi katsayısı anlamlı iken Üfe(-2) dönem gecikmesi parametresi anlamsız çıkmıştır.

Sonuç olarak modellerin OMS ve HKOK değerleri karşılaştırıldığı da ise en küçük ve birbirlerine çok yakın iki sonuç olan Doğrusal Regresyon ve Kantil Regresyon( $Q_2$ ) Modeli en uygun modeller olarak analiz edilmiştir.

Veriyi Bootstrap ile %200 oranın da arttırdığımız zaman Doğrusal ve Kantil Regresyon sonuçlarımıza göre;

Doğrusal Regresyon Modelin' de sadece Beklenti Anketi parametresi katsayısı anlamlı iken Üfe(-2) dönem gecikmesi parametresi ile Sabit katsayı parametresi anlamsız çıkmıştır.

Kantil Regresyon Modelin' de ise  $Q_1, Q_2$  ve  $Q_3$  kantil değerlerinde;

$Q_1$  kantil için Sabit katsayı parametresi ve Beklenti Anketi parametresi katsayısı anlamlı iken Üfe(-2) dönem gecikmesi parametresi anlamsız çıkmıştır.

$Q_2$  kantil için ise Doğrusal Modelde olduğu gibi sadece Beklenti Anketi parametresi katsayısı anlamlı iken Üfe(-2) dönem gecikmesi parametresi ile Sabit katsayı parametresi anlamsız çıkmıştır.

$Q_3$  kantilde ise  $Q_1$  kantilde olduğu gibi Sabit katsayı parametresi ve Beklenti Anketi parametresi katsayısı anlamlı iken Üfe(-2) dönem gecikmesi parametresi anlamsız çıkmıştır.

Sonuç olarak modellerin OMS ve HKOK değerleri karşılaştırıldığında ise en küçük ve birbirlerine çok yakın iki sonuç olan Doğrusal Regresyon ve Kantil Regresyon( $Q_2$ ) Modeli en uygun modeller olarak analiz edilmiştir.

Son olarak ise Üfe(-2) dönem gecikmesi parametresi katsayısı hesaplanan bütün modeller için anlamsız sonucu çıkmıştır.

Parametrik regresyon yöntemlerinin varsayımlarından biri olan normal dağılım varsayımının geçerli olmadığı durumlarda parametrik olmayan yöntemlerin daha etkin sonuçlar verdiği bilinmektedir. Uygulamada kullanılan veri set için normallik testi yapılmış ve verilerin normal dağılmadığı görülmüştür. Normal dağılmayan veri setlerinde başarılı sonuçlar elde edebilen kantil regresyon yöntemi bu veri seti uygulanmış, doğrusal regresyon ve kantil regresyon yöntemi sonuçları ise bölüm beşte incelenmiştir.

Yapılan analizlerde, yeniden örnekleme ile veri sayısının artması, verilerin dağılımlarının da normal dağılıma doğru yaklaşmasını sağlamıştır. Bu tip verilerde beklendiği gibi doğrusal regresyon analizi daha başarılı sonuçlar verdiği görülmüştür. Tüm veri yapılarından elde edilen sonuçlara bakıldığında, kantil regresyon analizi sonuçlarından sadece kantil regresyon( $Q_2$ ) modeli en iyi sonucu vermiştir. Diğer tüm modellerde, veri sayısının artışına bağlı olarak, doğrusal regresyon modeli en iyi tahmin sonuçlarını vermektedir. Ayrıca bootstrap uygulanmadan önce Tablo 5.3.'te görüldüğü gibi kantil regresyon( $Q_1$ ) değeri için bütün parametreler anlamlı sonuç vermiştir.

Uygulama sonuçlarına baktığımızda; Bootstrap yöntemi uygulanarak elde edilen verilerde parametrelerin anlamlılıkları bakımından kantil regresyonun, doğrusal regresyona göre daha anlamlı katsayılar elde ettiği söylenebilir. Fakat bu iki yöntemi OMS ve HKOK değerleri bakımından karşılaştırıldığında ise doğrusal regresyon yönteminin daha iyi model

belirlediđi görülmüştür. Bunun en önemli sebebi ise yeniden örnekleme yöntemi ile genişletilen veri yapısının, normal dağılıma yaklaşması gösterilebilir.

Bu çalışmada, farklı oranlarda kantil regresyon modelleri, farklı bootstrap sonuçları ile değerlendirilmiştir. Çalışmada kullanılan veriler ile yapılan uygulamalar sonucunda beklenenin aksine bir sonucun çıkması, deđişen varyans ya da otokorelasyon gibi EKK varsayımlarının göz ardı edilmesine de bağlanabilir. İleriki çalışmalarda, bootstrap veya benzer yapılardaki yeniden örnekleme teknikleri ile elde edilen yeni veri yapılarında, EKK varsayımlarının göz önüne alınarak karşılaştırma yapılması ile doğrusal – kantil regresyon farklarının daha net anlaşılması mümkün olacaktır.



## KAYNAKÇA

Akar, Cüneyt, (2013). “Gelişmekte Olan Piyasalarda Finansal Piyasa İstikrarının Kantil regresyon Yöntemiyle Test Edilmesi”. Doğu Üniversitesi Dergisi. 14(1): 1-9

Atakan, Cemal, (2003). “Anadolu Üniversitesi Bilim ve Teknoloji Dergisi”. 4(1) : 9-66

Atukeren, Erdal. (2008). “Çok Değişkenli Granger Nedensellik Sınamalarında Sonuçların Yorumu Sorunu”. 9. Ekonometri ve İstatistik Sempozyumu, Dokuz Eylül Üniversitesi. İzmir.

Aktükün, Aylın. (2002). “Asal Bileşenler Analizine Bootstrap Yaklaşımı”, İstanbul Üniversitesi İktisat Fakültesi Ekonometri ve İstatistik e-Dergisi, [eidergisi.istanbul.edu.tr/sayil1/ieuis1m1.pdf](http://eidergisi.istanbul.edu.tr/sayil1/ieuis1m1.pdf)

Altaş, Dilek. (2000), Yeniden Örnekleme Yaklaşımı Olarak Bootstrap Yöntemi ve Türkiye İthalat Modeline Uygulanması. Ekonometri Anabilim Dalı Doktora Tezi(Yayınlanmamış). İstanbul.

Behr, A. (2008). “Kantil Regression For Robust Bank Efficiency Score Estimation”. European Journal of Operational Research. 200: 568–581.

Buchinsky, M. (1998). “Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research”. The Journal of Human Resources. 33(1): 90-101.

Baur D., Saisana M., Niel S.N. (2004). “Modelling The Effects of Meteorological Variables on Ozone Concentration a Quantile Regression Approach”. Atmospheric Environment. 38(28): 4689–4699.

Birkes, D. ve Dodge Y. (1993). Alternative Methods of Regression, John Wiley Sons. New York. 80–140.

Baskan, Ş. (1993). Uygulamalı İstatistik. İzmir: Basımevi.

Berry, W. D. (1993). Understanding Regression Assumptions.

Bickel, P. J. ve Freedman, D. A. (1981). Some Asymptotic Theory For The Bootstrap. The Annals of Statistics. 9: 1196-1217

Beran, R. J. Ve Srivastava, M. S. (1985). Bootstrap Tests and Confidence Regions for Functions of a Covariance Matrix. The Annals of Statistics. 13: 95-115.

Ciner C., Gurdgiev C. ve Lucey B. M. (2013). "Hedges and Safe Havens: An Examination of Stocks, Bonds, Gold, Oil and Exchange Rates". *International Review of Financial Analysis*. (29): 202-211.

Chen, C. ve Wei, Y. (2005). Computational Issues for Quantile Regression. *Special Issue on Quantile Regression and Related Methods*. 67(2): 399-417.

Chen L. (2005). "An Introduction to Quantil Regression and the Quantreg Procedure". *Statistics and Data Analysis*. 213-230

Chernick, M. R. (1999). *Bootstrap Methods (Second Edition)*. Canada: John Wiley and Sons.

Cameron, C.A. ve Trivedi P.K. (2005). *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.

Çelik, Orkun. ve Selim, Sibel. ( 2013). Temel İnsan Sermayesi Modeli. "Europe and Global Economic Rebalancing." 24-25

Dicio, T. ve Tibshirani, R. (1987). Bootstrap Confidence Intervals and Bootstrap Approximations. *JASSA*. 82(397): 163-170.

Efron, B. (1979). Bootstrap Methods: Another Look at The Jackknife. *The Annals of Statistics*. 7: 1-26.

Efron, B. (1982). *The Jackknife, The Bootstrap and Other Resampling Plans*, Society for Industrial and Applied Mathematics. Philadelphia

Erfon, B. ve Gong, G., (1983). "Aleisurely Look at The Bootstrap, The Jackknife and CrossValidation ". *The American Statistician*. 37(1): 06-48.

Erfon, B, ve Tibshlram, R. (1986). "Bootstrap Methods for Standart Errors, Confidence Intervals and Other Measures of Statistical Accuracy". *Statistical Science*. 1: 54-77.

Efron B., (1990). More Efficient Bootstrap Computations. *JASA*. 85(409)

Efron, B. (1998). "R.A. Fisher in the 21 st Century". *Statistical Science*. 13(2): 95- 114

Eide, E. ve Showalter, H. M. (1998). "The Effect of School Quality on Student Performance: A Kantil Regression Approach". *Economics Letters*. 58: 345–350.

Eubank, R.A. (1988). *Spline Smoothing and Nonparametric Regression*. North Holland/ Amsterdam.

Eubank R. L. (1990). Nonparametric Regression And Spline Smoothing. Second Edition.

Frangos, C. C. ve Schucany, W. R. (1990). Jackknife eEtimation of the Bootstrap Acceleration Constant. Comp. Statist. Data Anal. 9: 271-282.

Fox, J. (1997). Appiled Regession Analysiss. Linear Models and Related Methods. London. 494-520.

Fox, J. (2000). Nonparametric Simple Regression: Smoothing Scatterplots. Sage University Paper. 130: 2.

Giil, Indermit S. (2002). Crafting Labor Policy: Technigues and Lessons from Latin America. Washington. World Bank.

Güriş Selahattin, Çağlayan Ebru ve Saçıldı İrem. (2010). Schooling and Wage Distribution for Male and Female Employees in the Public and Private Sector. International Research Journal of Finance and Economics, ISSN 1450-2887. 40:146- 154.

Gilchrist, Warren G. (2000). Statistical Modelling with Quantile Functions. Florida: Chapman and Hall/ CRC.

Gayeker, Savaş. (2015). Durağan Olmayan Var Sistemlerinde Bootstrap Yöntemi ile Granger Nedensellik Sınaması. Ekonometri Anabilim Dalı Yüksek Lisans Tezi(Yayınlanmamış). Gazi Üniversitesi. Ankara

Gujarati, D. N. (2004). Basic Econometrics. The Mc-Graw Hill Companies. s. 18.

Galton, F. (1886). “Regression Towards Mediocrity in Hereditary Stature”, Journal of Anthropological Institute of Great Britain and Ireland. (15):246–263

Güriş Selahattin ve Çağlayan Ebru (2000). Ekonometri Temel Kavramlar. İstanbul. s. 94.

Hahn, G. J. ve Meeker, W. Q. (1991). Statistical Intervals: A Guide for Practitioners. Wiley, New York.

Hao, L. ve Naiman, D. Q. (2007). Quantile Regression. The United States of America: Sage Publication.

Horowitz, J. L. (2001). The bootstrap, In J, J, Heckman and E, E, Leamer (Eds), Handbook of Econometrics. Amsterdam: North Holland Publishing.

Hardle, W., Müller, M., Sperlich, S. Ve Werwartz, A. (2004). Nonparametric and Semiparametric Models: An Introduction. Springer Series in Statistics. s.115.

Hardle W. (1999). Applied Nonparametric Regression. Cambridge University Press.

Jeremy, B. (1996). Recent Developments in Bootstrapping Times Series. Federal reserve Board, international Finance Discussion Paper Series. (25):2-34

Krieger, A. M. ve Bickel, P. J. (1989). Confidence bands for a distribution function using the bootstrap, J. Am. Statist. Assoc. (84):95-100.

Koenker, R. (2005). Quantile Regression, Econometric Society Monographs, Cambridge University Press.

Kurtoğlu, Fikriye. (2011). Quantile Regresyon: Teorisi ve Uygulamaları. Çukurova Üniversitesi Sosyal Bilimler Enstitüsü İstatistik Anabilim Dalı Ekonometri. Yüksek Lisans Tezi(Yayınlanmamış). Adana.

Koenker, R. ve Zhijie X. (2002). Inference on the Quantile Regression Process. Econometrica. 70(4):1583-1612

Kan, K. ve W. D. Tsai. (2004). “Obesity and Risk Knowledge in Taiwan: A Quantile Regression Analysis”. Journal of Health Economics. 23(5):907-34.

Koenker, R. (2005). Kantil Regression, USA. Cambridge University Press.

Koenker, R. ve Bassett G. (1978). “Regression Quantiles”, The Econometric Society. 46(1):33-50.

Koenker, R. ve Hallock, K. F. (2001). “Quantile Regression: An Introduction”. Journal of Economic Perspectives. 15: 143–156

Koenker, R. ve Machado, A. F. (1999). “Goodness of Fit and Related Inference Processes for Quantile Regression”. Journal of the American Statistical Association. 94: 1296–1310

Lahiri, S. N. (1992). On Bootstrapping M-estimators. Sankhya A. 54:157-170

Leng C, ve Tong X. (2013). A Quantile Regression Estimator for Censored Data. Bernoulli. 344-361.

Lingren, A. (1997). “Kantil Regression With Censored Data Using Generalized  $L_1$  Minimization”. Computational Statistics & Data Analysis. 23: 509-524.

Leping K-O. (2005). “Public-Private Sector Wage Differential in Estonia: Evidence From Quantile Regression, Tartu University”. Faculty of Economics And Business Administration. Tartu University Press. Orden. No:431

Lee, S. (2004). “Quantile Regression”. Lecture Notes for MECTI. s.1

McLachlan, G. J. ve Krishnan, T. (1997). The EM Algorithm and Extensions. Wiley, New York.

Mooney C. Z ve Duval, R. D. (1993). Bostatrapping: A Nonparametric Approach to Statistical Inferencece Un: Papers.

Meligkotsidou A., Vrontos D., L. ve Vrontos, D., V., (2009). “Kantil Regression Analysis of Hedge Fund Strategies”. Journal of Empirical Finance. 16: 264–279.

Maddala, G.S. (2001). Introduction to Econometrics, Third Edition. John Wiley&Sons. s.60.

Moussa M. ve Cheema M. (1992). “Non-Parametrik Regression in Curve Fitting”. The Statistician. 41(42): 209-225.

Newey, W. K. ve Powell. J. L. (1990). “Efficient Estimation of Linear and Type I Censored Regression Models Under Conditional Quantile Restrictions”. Econometric Theory. 6(3): 295–317.

Okutan, Duygu. (2009). Bootstrap Yönteminin Regresyon Analizinde Kullanımı ve Diğer Yöntemlerle Karşılaştırılması. Uygulamalı İstatistik Anabilim Dalı Yüksek Lisans Tezi(Yayınlanmamış). Ege Üniversitesi.

Orhunbilge, Ayşe Neyran. (2000). Uygulamalı Regresyon ve Korelasyon Analizi. 2. Baskı, İstanbul. Nobel Akademik Yayıncılık.

Özel Hasan Alp ve Sezgin Funda (2011). Ticari Serbestleşme-Ekonomik Büyüme İlişkisinin Bootstrap Kantil Regresyon Yardımıyla Analizi. Çanakkale Onsekiz Mart Üniversitesi ve İstanbul Üniversitesi.

Pagan, A. (1999). Nonparametric Econometrics. USA, Ullah. Aman: Cambridge University Press.

Ruppert D.,Wand M.P. ve Carroll, R.J. (2003). Semiparametric Regression. Cambridge University Press, s.57

Singh, K. (1981). On the Asymptotic Accuracy of Efron's Bootstrap. The Annals of Statistics. 9:1187-1195.

Staudte R. G. ve Sheather S. J. (1990). Robust Estimation and Testing. Wiley, New York.

Shao, J. (1996). Bootstrap Model Selection. JASA. 91(434): 55-665.

Stine, R. (1985). Bootstrap Prediction Intervals for Regression Jour of the Amer. Stat. Assoç. 80(392):1027

Strawderman R. L. ve Wells, M. T. (1997). Accurate Bootstrap Confidence Limits for the Cumulative Hazard and Survivor Functions Under Random Censoring. Journal of the American Statistical Association. 92:1356-1374.

Saçaklı, İrem. (2005). Kantil Regresyon ve Alternatif Regresyon Modelleri ile Karşılaştırılması. Ekonometri Anabilim Dalı Yüksek Lisans Tezi (Yayınlanmamış). Marmara Üniversitesi

Stifel D.C. ve Averett S.L. (2009). "Childhood Overweight In The United States: A Quantile Regression Approach". Economics & Human Biology. 7( 3): 387-397

Serper, Özer. (2004). Uygulamalı İstatistik. Bursa: Ezgi Kitabevi.

Saraçoğlu Bedriye ve Çevik F. (1995). Matematiksel İstatistik. Ankara: Gazi Büro Kitabevi.

Smeekees, S. (2009). Bootstrapping Nonstationary Time Series (Doctoral Thesis). Maastricht University.

Topuz, Derviş. (2002). Regresyonda Yeniden Örnekleme Yöntemlerinin Karşılaştırmalı Olarak İncelenmesi. Biyoloji Anabilim Dalı Yüksek Lisan Tezi (Yayınlanmamış).Niğde Üniversitesi.

Topal, Mehmet. (1999). Nonparametrik Regresyon Metodlarının İncelenmesi. s.3-4.

Zeng Q. ve Davidan M. (1997). Bootstrap-Adjusted Calibration Confidence Intervals for Immunoassay. Journal of the American Statistical Association. 92: 278-290.

Wu C. ve F J. (1986). Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis. Am. Of Stat. 14(4): 1261-1295.

Wang, H. (2007). "Quantile Regression: Overview and Applications to Risk Assessment". North Carolina State University. 1-26



# ÖZGEÇMİŞ

## KİŞİSEL BİLGİLER

Adı Soyadı: Seçkin ÇAMURLU

Uyruğu: T.C.

Doğum Tarihi ve Yeri: 25/09/1991 MALATYA

e-posta: sckncmrl@gmail.com

## EĞİTİM

Derece	Kurum	Mezuniyet Yılı
Lisans	Cumhuriyet Üniversitesi	2011-2015
Yüksek Lisans	Cumhuriyet Üniversitesi	2015-2018

## İŞ TECRÜBESİ

Tarih	Kurum	Görev
-------	-------	-------

## YABANCI DİL BİLGİSİ

KPDS ( )	ÜDS ( )	TOEFL ( )	EILTS ( )
----------	---------	-----------	-----------