KANTİL REGRESYON VE EN KÜÇÜK KARELER YÖNTEMLERİNİN KARŞILAŞTIRILMASI: BİR UYGULAMA DENEMESİ

Yüksek Lisans Tezi

Önder DORAK

Eskişehir, 2017

KANTİL REGRESYON VE EN KÜÇÜK KARELER YÖNTEMLERİNİN KARŞILAŞTIRILMASI: BİR UYGULAMA DENEMESİ

Önder DORAK

YÜKSEK LİSANS TEZİ

İşletme Anabilim Dalı

Danışman: Prof. Dr. Hasan DURUCASU

Eskişehir

Anadolu Üniversitesi

Sosyal Bilimler Enstitüsü

Mayıs, 2017

JÜRİ VE ENSTİTÜ ONAYI

Önder DORAK'ın "Kantil Regresyon ve En Küçük Kareler Yöntemlerinin Karşılaştırılması: Bir Uygulama Denemesi" başlıklı tezi **29 Mayıs 2017** tarihinde, aşağıdaki jüri tarafından Lisansüstü Eğitim Öğretim ve Sınav Yönetmeliğinin ilgili maddeleri uyarınca toplanan **İşletme (Sayısal Yöntemler)** Anabilim Dalında, **yüksek lisans tezi** olarak değerlendirilerek kabul edilmiştir.

İmza

Üye (Tez Danışmanı)	**: Prof.Dr.Hasan DURUCASU**	
Üye	**: Prof.Dr.Emel ŞIKLAR**	
Üye	**: Yrd.Doç.Dr.Hüseyin GÜRBÜZ**	

Prof.Dr.Kemal YILDIRIM
Anadolu Üniversitesi
Sosyal Bilimler Enstitüsü Müdürü

ÖZET

KANTİL REGRESYON ANALİZİ VE EN KÜÇÜK KARELER YÖNTEMLERİNİN KARŞILAŞTIRILMASI: BİR UYGULAMA DENEMESİ

Önder DORAK
İşletme Anabilim Dalı
Anadolu Üniversitesi, Sosyal Bilimler Enstitüsü, Mayıs, 2017
Danışman: Prof. Dr. Hasan DURUCASU

Bu çalışmada işletme kârlılığını etkileyen faktörler kantil regresyon analizi ve en küçük kareler yöntemi ile karşılaştırmalı olarak incelenmiş, asimetrik dağılıma sahip olduğu düşünülen işletme kazançlarının modellenmesinde, bu iki analiz ve yöntemin birbirlerine olan üstünlüklerinin ortaya konulması amaçlanmıştır. Bu amaç doğrultusunda, işletmelerin kârlılığını etkilediği düşünülen finansal kaldıraç ve işletme büyüklüğü faktörleri ele alınmış, bu faktörlerin işletme kârlılığının dağılımının farklı kantilleri üzerindeki etkileri araştırılmıştır. İşletme büyüklüğü, finansal kaldıraç ve kârlılığın ölçülmesinde çeşitli finansal oranlar ve finansal tablo kalemlerinden yararlanılmıştır.

Çalışmada Borsa İstanbul'da işlem gören imalat sektöründeki 183 firmanın 2013, 2014 ve 2015 yıllarına ait finansal tablolarından elde edilen veriler kullanılmış ve toplam 536 gözlemden oluşan bir örneklemden yararlanılmıştır.

Özellikle asimetrik dağılıma sahip değişkenlerin modellenmesinde, en küçük kareler yönteminden daha doğru ve yansız sonuçlar üreten kantil regresyon analizi sayesinde dağılımın koşullu ortalaması yerine, kantilleri modellenmiştir. Böylece işletme büyüklüğü ve finansal kaldıraç faktörlerinin düşük ve yüksek kârlı işletmeleri farklı yönlerde ve düzeylerde etkilediği sonucuna ulaşılmıştır.

Anahtar Kelimeler: Kantil regresyon, En küçük kareler, Finansal kaldıraç, İşletme büyüklüğü

ABSTRACT

THE COMPARISON OF QUANTILE REGRESSION ANALYSIS AND ORDINARY LEAST SQUARES METHODS: AN APPLICATION ATTEMPT

Department of Business Administration

Anadolu University, Graduate School of Social Sciences, June, 2017

Supervisor: Prof. Dr. Hasan DURUCASU

In this thesis, factors affecting firm earnings were analyzed via quantile regression analysis and ordinary least squares in comparison. It is aimed to reveal the superiorities of the methods in modeling the business earnings which are thought to have asymmetric distribution. For this purpose, financial leverage and firm size factors which are thought to affect the business profitability were held , and effects of these two factors on different quantiles of firm earnings were examined. Various financial ratios and financial table items were used in order to measure financial leverage, firm size and firm earnings.

In the study, the data of 183 firms traded on Istanbul Stock Exchange for the years 2013, 2014 and 2015 were collected, and thus sample of 536 observations were obtained.

Thanks to the quantile regression analysis, the method which produces more accurate and unbiased results in the modeling of asymmetrically distributed variables, quantiles of distribution were modeled instead of modeling the conditional mean. Thus, the conclusion that firm size and financial leverage affect poor and rich firms in different directions and levels was reached.

Keywords: Quantile regression, Ordinary Least Squares, Financial Leverage, Firm Size

29/05/2017

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

Bu tezin bana ait, özgün bir çalışma olduğunu; çalışmamın hazırlık, veri toplama, analiz ve bilgilerin sunumu olmak üzere tüm aşamalardan bilimsel etik ilke ve kurallara uygun davrandığımı; bu çalışma kapsamında elde edilemeyen tüm veri ve bilgiler için kaynak gösterdiğimi ve bu kaynaklara kaynakçada yer verdiğimi; bu çalışmanın Anadolu Üniversitesi tarafından kullanılan "bilimsel intihal tespit programı" ile tarandığını ve hiçbir şekilde "intihal içermediğini" beyan ederim. Herhangi bir zamanda, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçlara razı olduğumu bildiririm.

Önder DORAK

İÇİNDEKİLER

İKİNCİ BÖLÜM

ÜÇÜNCÜ BÖLÜM

TABLOLAR/ÇİZELGELER DİZİNİ

Sayfa

ŞEKİLLER DİZİNİ

SİMGELER VE KISALTMALAR DİZİNİ

α	: Alfa
σ	: Anakütle Standart Sapması
μ	: Anakütle Ortalaması
Y_i	: Bağımlı Değişken
X_i	: Bağımsız Değişken
β_i	: Regresyon Parametresi (Katsayısı)
ε_i	: Hata Terimi
n	: Gözlem sayısı
E( )	: Beklenen Değer
N( )	: Normal Dağılım
σ^2	: Varyans
$\hat{Y}_i$	: Tahmini Y Değeri
b_i	: Regresyon parametresinin tahmini değeri
$\hat{\beta}$	: Regresyon parametresinin tahmin vektörü (Örneklem)
e_i	: Örneklem hata terimi
Σ	: Toplam Sembolü
d	: Yön Vektörü
R^2	: Belirlilik Katsayısı
F(x)	: Kümülatif Dağılım Fonksiyonu
f(x)	: Olasılık Yoğunluk Fonksiyonu
Q(p)	: Kantil Fonksiyonu
I( )	: İşaret Fonksiyonu

$\rho_{(p)}$	: Kontrol Fonksiyonu
BORC	: Borç Oranı
BORCOZ	: Borç-Özkaynak Oranı
EKK	: En Küçük Kareler
EKMS	: En Küçük Mutlak Sapmalar (Medyan Regresyon)
Kov	: Kovaryans
lim	: Limit
MADDI	: Duran Varlıklar Oranı
min	: Minimum
$QSC^{(p)}$	: Kantil Tabanlı Basıklık Ölçüsü
$QSK^{(p)}$	:Kantile Tabanlı Çarpıklık Ölçüsü
ROE	: Özsermaye Kârlılığı
ROA	: Aktif Getiri
SATIS	: Toplam Satışlar
UOZ	: Uzun Dönem Borç Oranı/Devamlı Sermaye Oranı
UZUN	: Uzun Dönem Borç Oranı
Var	: Varyans
VBÇ	: Varyans Büyütme Çarpanı
VARLIK	: Toplam Varlıklar

GİRİŞ

İlk kez 19. yüzyılda ortaya atılan regresyon analizi, o günden bugüne kolay uygulanabilmesi ve anlaşılır olması sebebi ile gerek sosyal bilimlerde gerekse fen bilimlerinde sıklıkla başvurulan bir yöntem olmuştur. Regresyon analizi en temel anlamıyla, değişkenler arasındaki ilişkiyi matematiksel olarak modellemektedir. Böylece değişkenler arasındaki ilişkinin yönü ve varlığı belirlenirken, belirlenen regresyon modeli sayesinde açıklanmak istenen değişkenin değerleri, diğer değişkenlerin değerlerinin kullanılması ile tahmin edilebilmektedir.

Regresyon analizi ile ilgili çoğu kaynakta EKK yöntemi, regresyon analizinin önemli bir adımı olan parametre kestiriminin ayrılmaz ve alternatifi olmayan bir parçasıymış gibi aktarılmaktadır. Ancak EKK yöntemine alternatif başka yöntemler de mevcuttur. Parametrik olmayan, sağlam (robust) gibi çeşitli sınıflandırmalara tabi tutulan bu yöntemlerden bazıları EKK'nin bazı varsayımlarını göz ardı ederken, bazıları ise ek ve farklı varsayımlarla parametre kestirimlerini gerçekleştirmektedir. Alternatif parametre kestirimi yöntemlerinden biri olan EKMS yöntemi veya diğer adıyla medyan regresyon, EKK yöntemi kadar eski bir yöntemdir.

Koenker ve Bassett (1978), medyan regresyonu geliştirerek kantil regresyon yöntemini ortaya atmıştır. Bu yöntemde bağımlı değişkenin dağılımının çeşitli kantilleri için regresyon doğruları tahmin edilir. Bu tahminler gerçekleştirilirken, hata terimlerinin mutlak sapmaları kantil değerlerine bağlı olarak ağırlıklandırılarak enküçüklenmektedir. Kantil regresyon yöntemi, özellikle hata terimleri normal dağılmadığında EKK yönteminden daha doğru ve yansız sonuçlar üretir.

Bu tezde kantil regresyon yöntemini açıklamak ve kantil regresyon yöntemini EKK yöntemi ile karşılaştırarak hangisinin daha doğru sonuçlar ürettiğini belirlemek amaçlanmıştır. Bu amacı gerçekleştirmek için finansal kaldıraç ve işletme büyüklüğünün kârlılık üzerindeki etkileri uygulama konusu olarak seçilmiştir.

Çalışmanın birinci bölümünde doğrusal regresyon analizi genel olarak açıklanacaktır. Bu bölümde regresyon kavramı tanımlanacak, basit ve çoklu doğrusal regresyon modellerine yer verilecek ve doğrusal regresyon analizinin varsayımlarından bahsedilecektir. Bunun yanında doğrusal regresyonun parametre kestirim yöntemleri olan EKK ve EKMS yöntemleri açıklanacaktır.

İkinci bölümde ise kantil kavramı ve kantil regresyon yöntemi yer almaktadır. Bu bölümde kantil regresyon modeli ve kantil regresyon modelinin parametre tahminlerinin nasıl gerçekleştirildiğinden bahsedilecektir.

Üçüncü ve son bölümde ise finansal kaldıraç ve işletme büyüklüğünün Borsa İstanbul'da işlem gören imalat sektöründeki işletmeler üzerindeki etkileri kantil regresyon ve EKK yöntemi kullanılarak incelenmiştir.

BİRİNCİ BÖLÜM

DOĞRUSAL REGRESYON ANALİZİ

Regresyon kavramı ilk kez Francis Galton (1886) tarafından tanıtılmıştır. Galton çalışmasında uzun boylu ebeveynlerin uzun boylu çocuklara, kısa boylu ebeveynlerin de kısa boylu çocuklara sahip olma eğilimlerine karşın, çocukların boylarının anakütle ortalamasına yaklaştığını, kendi ifadesi ile gerilediğine/indirgendiğine (regress) dair bulgular elde etmiştir. Diğer bir ifadeyle çok uzun boylu ebeveynlerin ve çok kısa boylu ebeveynlerin çocuklarının boylarının ortalamaya yakın olduğu sonucuna ulaşmıştır. Bu çalışmayı daha da geliştiren Karl Pearson (1903) ise çalışmasında uzun boylu babaların çocuklarının babalarından daha kısa, kısa boylu babaların çocuklarının ise babalarından daha uzun olduklarını ortaya çıkarmıştır. Böylece uzun ve kısa çocukların boylarının benzer şekilde ortalamaya gerilediğini (regressing) öne sürmüştür. Galton bu durumu "sıradanlık regresyonu (sıradanlığa doğru çekilme) olarak adlandırmıştır (Gujarati, 2004, s. 17).

Günümüzde regresyon analizi daha farklı bir anlamda kullanılmaktadır. Bugünkü anlamda regresyon analizi, bir veya daha fazla değişkenin belirli bir değişken üzerinde nasıl bir etkiye sahip olduğunu açıklamaya çalışmaktadır.

Bu bölümde öncelikle regresyon ile ilgili kavram ve varsayımlara yer verilecek, ardından doğrusal regresyon modellerinden bahsedilecektir. Son olarak da doğrusal regresyon modellerinin parametre tahminleri hakkında bilgi verilecektir.

1.1. Regresyon Kavramı

Değişkenlerin mevcut olduğu her sistemde, değişkenler arasındaki ilişkileri incelemek her zaman bir ilgi konusu olmuştur. Kimi zaman değişkenler arasında basit bir fonksiyonel ilişki bulunurken, genelde değişkenler arasındaki ilişki karmaşık bir fonksiyonel yapıya sahiptir. Bu durumda gerçekte var olan karmaşık fonksiyonel ilişkiyi modelleyebilmek için çeşitli varsayımlara dayanan basit matematiksel modellerden yararlanılır. Araştırmada yer alan değişkenlerdeki değişimin incelenmesiyle elde edilen matematiksel model sayesinde, değişkenlerin ortak ve ayrı ayrı etkileri incelenebilir. Böylece aralarındaki ilişkinin yapısı ortaya çıkarılmış olur. Regresyon modeli de değişkenler arasındaki ilişkiyi açıklamaya çalışan böylesi matematiksel modellerden biridir (Draper ve Smith, 1998, s. 15).

Değişkenler, aralarındaki ilişki incelenirken bağımlı (açıklanan) değişkenler ve bağımsız (açıklayan) değişkenler olmak üzere iki gruba ayrılırlar. Bağımlı değişken anlaşılmaya, açıklanmaya ve modellenmeye çalışılan değişkendir ve genellikle Y harfi ile temsil edilir. Bağımsız değişken ise bağımlı değişkeni açıklamada ve modellemede kullanılan değişkendir ve genelde X harfi ile gösterilir (Chatterjee ve Simonoff, 2013, s. 3).

Regresyon, değişkenler arasındaki ilişkiyi modelleyerek açıklamak için kullanılan istatistiksel bir tekniktir. Regresyon analizi, bir bağımlı değişken ile bir veya daha fazla bağımsız değişken arasındaki ilişkinin var olup olmadığını, eğer bir ilişki varsa ilişkinin yönünü ve büyüklüğünü bir denklemle ifade etmektedir. Bağımsız değişkende gerçekleşen değişimlerin bağımlı değişken üzerindeki etkisi incelenir. Anlaşılması, uygulanması ve yorumlanması oldukça kolay olduğundan bilimsel çalışmalarda sıklıkla kullanılmaktadır (Montgomery, Peck, ve Vining, 2013, s. 1)

Regresyon modelinin genel yapısı (1.1)'de gösterilmiştir.

$$Bağımlı\ değişken = Model\ fonksiyonu + Hata\ terimi \qquad (1.1.)$$

Model fonksiyonu, bağımsız değişkeni ve veri kümesinin analizi ile elde edilen bilinmeyen parametrelerin kestirimlerini içermektedir(Draper ve Smith, 1998) Bilinmeyen parametre β ile gösterilir. β bir skaler veya vektördür.

Hata terimi ise bağımlı değişkenin gerçek değerleri ile gözlemlenen değerleri arasındaki farka eşittir. Ancak hata terimlerinin gözlemlenmesi hiçbir zaman mümkün olmadığından, regresyon modelleri kurulurken hata terimi ile ilgili bazı varsayımlarda bulunulur (Şıklar, 2000, s. 6). Bu varsayımlara ilerleyen kısımlarda değinilecektir.

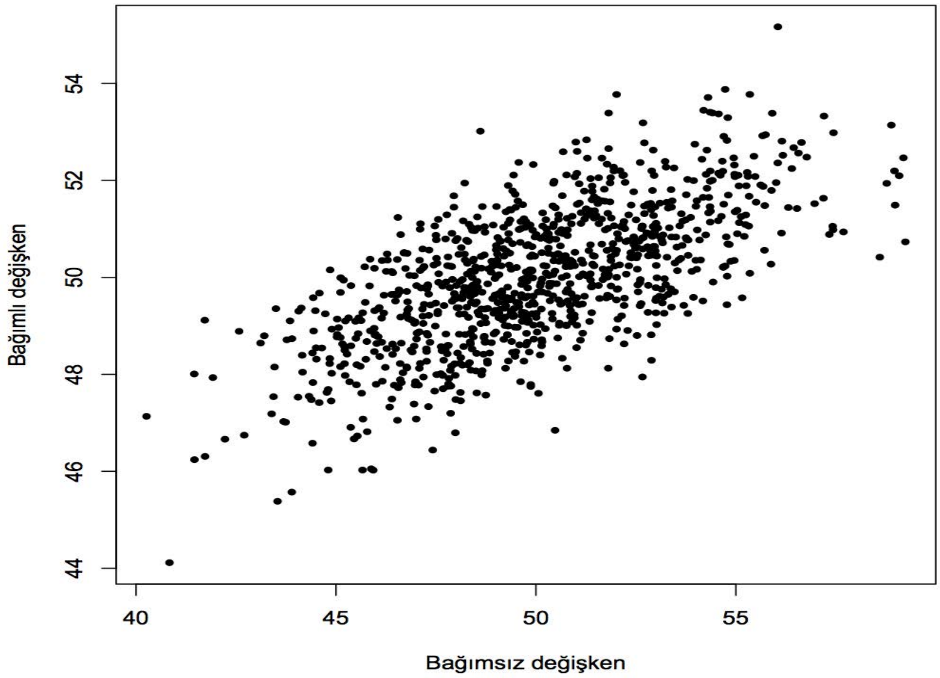
Bilinmeyen parametrenin (β) kestirilmesi ile bağımsız değişken ile bağımlı değişken arasındaki ilişkinin yönü ve varlığı belirlenir. İlişkinin yönü ile ifade edilmek istenen, bağımsız değişken ile bağımlı değişkenin aynı yönde mi yoksa zıt yönde mi hareket edeceğidir. Parametre değeri pozitif olduğunda ilişki aynı yönlü, negatif olduğunda ilişki zıt yönlüdür. İlişkinin derecesi ile değişkenler arasındaki ilişkinin kuvvetli veya zayıf olduğu ifade edilmektedir (Serper, 2014, s. 527).

İlişkinin fonksiyonel yapısının nasıl olduğunu anlayabilmek için serpilme diyagramlarından yararlanılır. Serpilme diyagramı, bağımlı değişkenin ordinatta, bağımsız değişkenin ise apsiste temsil edildiği ve bu iki değişkene ait gözlem değerlerinin düzlemde noktalar ile gösterildiği bir grafik türüdür (Friendly ve Denis, 2005, s. 103).

Tek bir bağımsız değişkenin yer aldığı regresyon problemlerinin analizinde serpilme diyagramı başlangıç noktasını oluşturmaktadır. Serpilme diyagramı, incelenerek bağımlı ve bağımsız değişken arasında anlamlı bir ilişkinin olup olmadığı; varsa ilişkinin yönünün ve fonksiyonel yapısının nasıl olduğu hakkında görsel bilgi edinilebilir. Ancak birden fazla bağımsız değişkenin olduğu problemlerde birçok serpilme diyagramının yer aldığı bir serpilme diyagramı matrisi oluşturulur. Bu matriste her bir bağımsız değişkenin bağımlı değişkenle ilişkisi ayrı grafiklerle gösterilir ve böylece bağımsız değişkenlerin bağımlı değişkenle ilişkileri ayrı ayrı incelenir (Weisberg, 2005, s. 2). Örnek bir serpilme diyagramı Şekil 1.1.'de verilmiştir.

Şekil 1.1. de görüldüğü üzere X ve Y gözlem ikilileri serpilme diyagramında birer nokta ile ifade edilmiştir. İlişkinin yönü ve derecesi hakkında noktaların oluşturduğu şekil incelenerek fikir sahibi olunabilir. Şekil 1.1. incelendiğinde X değişkeni arttığında Y değişkeninin değerinde de genel olarak bir artışın olduğu gözlemlenmektedir. Böylece X ve Y değişkenleri arasında aynı yönlü bir ilişkinin varlığından söz etmek mümkündür. İlişkinin şiddeti hakkında bilgi sahibi olabilmek için düzlemde yer alan noktalardan dışta kalanlar birleştirilerek bir elips çizilir. Elips ne kadar genişse ilişki o kadar zayıf, ne kadar darsa ilişki o kadar kuvvetlidir (Şıklar, 2000, s. 3).

Regresyon analizinde gerçekleştirilmek istenen amaç serpilme diyagramında noktalarla temsil edilen gözlem değerlerini en iyi açıklayacak matematiksel fonksiyonu belirlemektir. Bunun için noktaların arasından geçecek en iyi doğru veya eğri belirlenmeye çalışılır.



Şekil 1.1. *Serpilme Diyagramı*

1.1.2. Regresyon analizinin kullanım amaçları

Regresyon modeliyle bağımlı değişken ile bağımsız değişken(ler) arasındaki ilişkinin matematiksel olarak ifade edilmesi amaçlanır. Kurulan model sayesinde bağımlı değişkeni hangi bağımsız değişkenlerin etkilediği belirlenir. Bu sayede değişkenler arasındaki karmaşık ilişki basit bir matematiksel model yardımıyla daha kolay anlaşılır hale gelmiş olur. Ayrıca çoklu doğrusal regresyon analizinde hangi bağımsız değişkenin bağımlı değişken üzerinde daha çok etki sahibi olduğunu belirlemek de regresyon analizinin amaçları arasındadır. (Alpar, 2013, s. 417).

Regresyon analizinin bir diğer amacı da bağımlı değişkenin değerlerinin kestirilmesidir. Modelde yer alan parametre kestirimleri ve bağımsız değişken(ler)e ait deterministik veri kümesi sayesinde, bağımsız değişken(ler)in belirli değerleri için bağımlı değişken değerleri kestirilebilir (Alpar, 2013, s. 417).

Ancak regresyon analizinin en önemli amacı parametre kestiriminde bulunmaktır. Parametre kestirimi ile bağımsız değişkenlerin bağımlı değişkenleri ne yönde ve ne ölçüde etkilediği belirlenebilir. Bir parametre kestiriminin sıfır ya da anlamsız çıkması, ilgili bağımsız değişkenin bağımlı değişken üzerinde herhangi bir etkisi olmadığı anlamına gelir. Böylece hangi bağımsız değişkenlerin bağımlı değişkeni en iyi biçimde tanımladığı belirlenmiş olur (Şıklar, 2000, s. 4).

1.1.3. Koşullu beklenen değer ve koşullu varyans

Bağımsız değişken *X*'te gerçekleşecek bir değişimin, bağımlı değişken *Y*'nin dağılımında nasıl bir etki oluşturduğunu incelemek için ortalama fonksiyonu kavramından yararlanılır. Ortalama fonksiyonu (1.2)'de gösterilmiştir:

$$\mathrm{E}(\mathrm{Y}|\mathrm{X} = \mathrm{x}) = x'in\ bir\ fonksiyonu \qquad (1.2)$$

E(Y|X = x), "*Sabit bir X=x değeri için Y'nin beklenen değeri*" biçiminde okunmaktadır. Diğer bir ifadeyle *Y*'nin koşullu ortalama değeri, *X*'in bir fonksiyonu yardımıyla belirlenmeye çalışılır. Eşitliğin sağ tarafı regresyon probleminin yapısına göre değişir. Aynı zamanda eşitliğin sağ tarafı değişkenler arasındaki ilişkinin fonksiyonel yapısını belirler. Eğer değişkenler arasında ilişki doğrusal ise eşitliğin sağ tarafı doğrusal bir fonksiyon olur.

Koşullu varyans ise bağımlı değişkenin dağılımının bir başka özelliğidir. Varyans fonksiyonu olarak da adlandırılır. Var(Y|X=x) biçiminde ifade edilir ve "*X=x değeri için Y'nin varyansı*" olarak okunur. Genellikle regresyon modelleri kurulurken koşullu varyans hakkında varsayımlarda bulunulur (Weisberg, 2005, s. 9–11).

Regresyon doğrusu ortalamayı temsil eder. Bu durum genellikle göz ardı edildiğinden, regresyon analizinin doğru olarak yorumlanamamasına yol açmaktadır. Bu güçlüğü aşabilmek için regresyon doğrusunun ortalamayı temsil etmesi, ortalama fonksiyonu ve varyans fonksiyonu kullanılarak aşamalı olarak aşağıda açıklanmaya çalışılmıştır (Freund, Wilson, ve Sa, 2006, s. 38).

Bağımlı değişkenin değeri, anakütle ortalaması ile varyansın doğrusal bir modeline dayanılarak hesaplanır. Bu durum (1.3)'te gösterilmiştir:

$$Y_i = \mu + \varepsilon_i \qquad (1.3)$$

Modelde yer alan μ ortalamayı, ε_i hata terimi de ortalamadan sapmayı diğer bir ifadeyle varyansı temsil eder. Regresyon modelinde ise *Y*'nin ortalama değeri, *X* bağımsız değişkeni ile ilişkilendirildiğinde bir fonksiyon ile ifade edilmektedir

(Freund, Wilson ve Sa 2006, s. 38). Buna göre örneğin basit doğrusal regresyon modelinin ortalama fonksiyonu ve varyans fonksiyonu sırasıyla (1.4) ve (1.5)'te verilmiştir.

$$\mu = \mathrm{E}(\mathrm{Y}|\mathrm{X} = \mathrm{x}) = \beta_0 + \beta_1 X \tag{1.4}$$

$$Var(Y|X = x) = \sigma^2 \tag{1.5}$$

Görüldüğü gibi Y değişkeninin ortalama değeri artık X değişkenin bir fonksiyonu ile ifade edilmektedir. β_0 ve β_1 ise modelin bilinmeyen parametreleridir. Bu parametrelerin anlamlarına basit doğrusal regresyon modeli başlığı altında değinilecektir.

Varyans $\sigma^2>0$ olduğu varsayılmaktadır. Varyansın pozitif olduğu düşünüldüğünden, *i.* gözlem değerine dayanılarak hesaplanan $\hat{Y_i}$ gözlem değeri, $\mathrm{E}(\mathrm{Y}|\mathrm{X} = \mathrm{x})$ beklenen değerine hiçbir zaman eşit olamaz. Bu sebeple, beklenen değer ile gözlemlenen değer arasındaki farkı hesaba katabilmek için hata terimi ε_i kavramı geliştirilmiştir (Weisberg, 2005, s. 21).

1.2. Doğrusal Regresyon Modelleri

Regresyon modelleri doğrusal ve doğrusal olmayan olmak üzere iki ana başlıkta sınıflandırılır. Bu çalışmada yalnızca doğrusal regresyon modelleri üzerinde durulmuştur. Değişkenler arasındaki ilişkinin fonksiyonel yapısı doğrusal ise bu değişkenler arasındaki ilişki doğrusal regresyon modeli ile analiz edilir (Weisberg, 2005, s. 21). Doğrusal regresyon, iki veya daha fazla değişken arasındaki ilişkiyi tanımlamaya çalışan matematiksel bir tekniktir (Marill, 2004, s. 87).

Doğrusal regresyon modelleri veri kümelerinden yararlanarak değişkenler arasındaki ilişkiler hakkında teorik sonuçlar üretirler. Ancak bu sonuçların gerçek hayattaki karmaşık ilişkileri, kesin bir doğrulukla açıklayabildiğini söylemek yanlış olur. Bunun için doğrusal regresyon modelleri kurulurken çeşitli varsayımlarda bulunularak, elde edilen sonuçların gerçek hayat sonuçlarına olabildiğince yakın olmasına çalışılır. Doğrusallık varsayımı ile değişkenler arasındaki ilişkinin doğrusal olduğu ve bu ilişkinin grafiksel olarak düz bir çizgi ile gösterilebileceği ifade edilmek istenmektedir. Böylece değişkenler arasındaki ilişkilerin kesin bir doğruluk ile olmasa da, daha kolay ve kullanışlı matematiksel bir teknikle açıklanması amaçlanmıştır (Birkes ve Dodge, 1993, s. 4).

Doğrusal regresyon modeli, parametrelerine göre doğrusaldır (Gürsakal, 2013, s. 361) Diğer bir ifadeyle bir modelin doğrusal olup olmadığına, parametrelerinin doğrusallığı incelenerek karar verilir. Örneğin;

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \varepsilon_i \tag{1.6}$$

modeli, X değişkeninin 2. kuvveti yer aldığından dolayı kuadratik gibi görünmektedir. Ancak model parametreleri olan β değerleri doğrusaldır ve modelin doğrusal olduğunu ifade etmek için yeterlidir. Çünkü $X=X_1$ ve $X^2=X_2$ biçiminde iki farklı değişken olarak ifade edildiğinde model şu biçimi alır:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon_i \tag{1.7}$$

Böylece model doğrusal tanımına uymuş olur (Draper ve Smith, 1998, s. 21)

Bir bağımsız değişkenin olduğu regresyon modelleri basit doğrusal regresyon, birden fazla bağımsız değişkenin olduğu regresyon modelleri ise çoklu doğrusal regresyon olarak adlandırılır.

1.2.1. Doğrusal regresyon analizinin varsayımları

Doğrusal regresyon analizinin varsayımları; hata terimi ile ilgili varsayımlar, bağımsız değişken ile ilgili varsayımlar ve diğer varsayımlar olmak üzere üç alt başlıkta incelenir (Şıklar, 2000, s. 6).

1.2.1.1. Hata terimi ile ilgili varsayımlar

1. Varsayım: Hata terimi ε_i'nin aldığı değerler tesadüfidir ve X_i değerleri için sıfır, negatif veya pozitif değerler alabilir ancak asla önceden kestirilemez. Alacağı değerler belirli bir olasılığa göre her bir gözlem değeri için farklılık gösterir. Bu yüzden hata terimleri, rassal hatalar olarak da anılırlar (Şıklar, 2000, s. 7).

2. Varsayım: Tüm koşullu dağılımları için hataların beklenen değeri (ortalaması) sıfıra eşittir:

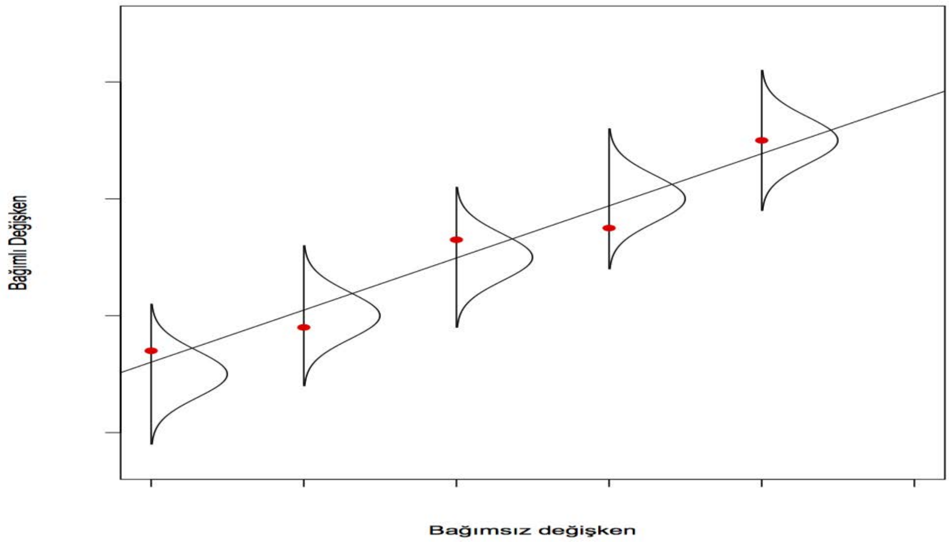
$$E(\varepsilon_i|X) = 0, i = 1, \dots, n \tag{1.8}$$

Varsayım, saçılım grafiğinde regresyon doğrusunun her iki yanında (altında ve üstünde) yer alan noktaların sayısının birbirine eşit olması anlamına gelir. Diğer bir ifadeyle regresyon doğrusu, gözlem değerlerini temsil eden noktaların tam ortasından geçer. Böylece veri kümesi iki eşit parçaya bölünmüş olur. (Pardoe, 2012, s. 60). Bu varsayım, belirli aşırı değerlere sahip gözlemlerin modelin çok yüksek veya çok düşük değerler üretmesini engeller. Varsayım göz ardı edildiğinde β_0'ın hesaplanmasında hatalar meydana gelebilir (Chatterjee ve Simonoff, 2013, s. 8). Hataların beklenen değerinin sıfır olduğu varsayıldığında Y'nin koşullu beklenen değeri şu biçimde ifade edilir:

$$E(Y_i|X_i) = \beta_0 + \beta_1 X_i \tag{1.9}$$

Bu sayede bilinmeyen regresyon parametreleri kestirildiğinde, sabit X değerleri için Y'nin koşullu ortalaması da hata terimleri göz ardı edilerek kestirilmiş olur.

3. Varsayım: Her X_i değeri için hata terimleri normal dağılır. Normal dağılım varsayımı, güven aralığı ve hipotez testlerinin hesaplanması gibi istatistiksel çıkarımlarda gereklidir (Şıklar, 2000, s. 8). Saçılım grafiğindeki noktaların regresyon doğrusuna yakın olduğu ve doğrudan uzak olma ihtimalinin az olduğu varsayılır (Marill, 2004, s. 88). Şekil 1.2.'de gösterildiği üzere gözlem noktaları regresyon doğrusuna oldukça yakındır.



Şekil 1.2. *Rassal Hataların Olasılık Dağılımları*

4. Varsayım: Seçilmesi mümkün tüm örneklemlerden elde edilen gözlemlerden çıkarılan varyans değerleri anakütle değerleri ile ilgili olarak her örneklemde sabittir (Şıklar, 2000, s. 10). Buna göre, her X değeri için hata teriminin olasılık dağılımı sabit varyansa sahiptir. Bu durum sabit varyanslılık (homoscedasticity) olarak adlandırılır. Sabit varyanslılık, modelin anakütlenin bazı kısımlarında yüksek, bazı kısımlarında düşük varyansa sahip olmadığını varsayar (Chatterjee ve Simonoff, 2013, s. 8). Gerçekte çoğu zaman sabit varyanslılık varsayımı ihlal edilir. Sabit varyanslılığın olmadığı zaman değişen varyanslılıktan (heteroscedasticity) söz edilir. Değişen varyanslılık hata terimlerinin asimetrik dağılması anlamına gelmektedir. Bu durumda ise bağımlı değişkenin ortalamasında meydana gelen artışlar, dağılımın daha çarpık olmasına sebep olur (Freund, Wilson ve Sa, 2006, s. 63). Saçılım grafiği kullanılarak sabit varyanslılık açıklanmak istendiğinde, gözlem değerlerini temsil eden noktaların regresyon doğrusuna olan uzaklıklarının birbirine yakın oldukları kabul edilir. Böylece rassal hataların değişkenliği birbirine benzer olur. Bazı noktalar diğerlerine göre regresyon doğrusundan daha uzakta olsalar da; sabit varyanslılık varsayımı, noktaların regresyon doğrusuna göre ortalama değişkenliklerinin aynı kaldığını kabul etmektedir (Marill, 2004, s. 88). Şekil 1.2. incelendiğinde her gözlem değeri için varyansın çan eğrisi biçiminde ve sabit olduğu görülmektedir. Her gözlemin regresyon doğrusuna olan dikey uzaklıkları birbirine benzerdir (Pardoe, 2012, s. 60).

5. Varsayım: Hata terimleri birbirlerinden bağımsızdırlar. Herhangi bir hata teriminin değerinin bilinmesi, bir başka hata teriminin değeri hakkında bir bilgi vermez. Diğer bir ifadeyle, regresyon modelinin Y'yi herhangi bir gözlem değeri için hatalı kestirdiğinin bilinmesi, herhangi bir gözlem değeri ile ilgili bir bilgi sağlamaz. Hata terimlerinin arasında bir ilişkinin var olması "otokorelasyon" olarak adlandırılır (Poole ve O'Farrell, 1971, s. 148).

Hata terimi ile ilgili varsayımlar (1.10)'daki gibi ifade edilir.

$$\varepsilon_i \sim N(0, \sigma^2) \qquad (1.10)$$

Burada yer alan ifadede 0, hata teriminin sıfır ortalamalı; σ^2, hata teriminin sabit varyansa sahip olduğu ve bu varyansın σ^2'ye eşit olduğu açıklanmaktadır. N harfi ise rassal hataların normal dağıldığının varsayıldığı anlamına gelmektedir.

1.2.1.2. Bağımsız değişkenle ilgili varsayımlar

Bağımsız değişkenle ilgili dört adet varsayım bulunmaktadır.

6. Varsayım: Bağımsız değişkenler X_i'ler ile bağımlı değişken Y değerleri arasında doğrusal bir ilişki vardır. Noktalar saçılım grafiği üzerinde eğri vb. bir şeklin değil, bir doğrunun etrafında toplanırlar (Marill, 2004, s. 87).

7. Varsayım: Doğrusal regresyon analizi ayrıca bağımsız değişkenler arasında doğrusal bir ilişkinin olmadığını varsayar. Eğer bağımsız değişkenler arasında doğrusal bir ilişki mevcutsa bu duruma çoklu bağıntı adı verilir. Bu varsayım yalnızca çoklu doğrusal regresyon için geçerlidir. Basit doğrusal regresyonda yalnızca bir bağımsız değişken yer aldığından bu varsayıma gerek bulunmamaktadır. (Poole ve O'Farrell, 1971, s. 148) Çoklu bağıntının var olması parametre kestirimlerinin beklentinin aksine sonuçlar vermesine yol açabilir (Şıklar, 2000, s. 81).

8. Varsayım: Hata terimleri ile bağımsız değişken arasındaki kovaryans sıfıra eşittir.

$$Kov(\varepsilon_i X_i) = 0 \tag{1.11}$$

Diğer bir ifadeyle bağımsız değişkenler ile hata terimi arasında bir ilişki yoktur. Bu yüzden, bağımsız değişkenin değerleri stokastik değil sabittir. Diğer bir ifadeyle X'in gözlemlenen değerlerinin bilindiği ve sabit olduğu varsayılır (Alpar, 2013).

9. Varsayım: Bağımsız değişkenlerin değerleri ölçülürken hata yapılmamaktadır. Ölçme hatasının varlığı bağımsız değişken ile hata terimleri arasında bir ilişkinin olmasına yol açar. Bu durum ise parametre kestirimlerinin yanlış hesaplanmasına neden olur (Şıklar, 2000, s. 12).

1.2.1.3. Diğer varsayımlar

10. Varsayım: Doğrusal regresyon analizi ile ilgili ele alınacak son varsayım model belirlenme hatası ile ilgilidir. Buna göre doğrusal regresyon analizinde modelin belirlenmesinde hata olmadığı varsayılır. Model belirlenme hatası; regresyon modelinin fonksiyonel yapısı doğru seçilmediğinde, modele yanlış bağımsız değişkenler eklendiğinde ve herhangi bir varsayım doğru olmadığında gerçekleşir. Model belirleme hatasının olup olmadığı, model çıktılarının regresyon analizine konu olan araştırma ile ilgili teorileri ne kadar iyi yansıtabildiği göz önüne alınarak belirlenir. Model sonuçları teori ile çelişiyorsa genellikle bir belirleme hatasından söz edilir. Bu durumda modelde hangi değişkenlerin yer alacağı, hangi varsayımların kabul edildiği ve modelin fonksiyonel yapısının nasıl olduğu gözden geçirilmelidir (Berry, 1993, s. 30)

1.2.2. Basit doğrusal regresyon modeli

Basit doğrusal regresyon modeli bir bağımsız değişken ile bir bağımlı değişken arasındaki ilişkiyi tanımlayan matematiksel bir tekniktir. "Basit doğrusal regresyon modeli" teriminde geçen "basit" ibaresi modelde bir tane bağımsız değişkenin olduğunu; "doğrusal" ibaresi ise değişkenler arasındaki ilişkinin doğrusal olduğunun varsayıldığı anlamına gelmektedir.

Basit doğrusal regresyon modeli (1.12)'de verildiği gibi ifade edilir:

$$Y = \beta_0 + \beta_1 X_1 + \varepsilon_i \tag{1.12}$$

β_0, X=0 olduğunda regresyon doğrusunun dikey ekseni kestiği noktayı gösterir ve $E(Y|X = x)$ beklenen değerine eşittir. β_1 ise doğrunun eğimini verir ve X bağımsız değişkeninde gerçekleşecek bir birimlik değişimin Y'de nasıl bir değişime yol açacağını gösterir. Basit doğrusal regresyonda varyans fonksiyonunun pozitif σ^2'ye eşit ve sabit olduğu varsayılır (Weisberg, 2005, s. 21).

Yukarıdaki açıklamaları temel alarak (1.3)'te yer alan μ yerine (1.4)'te yer alan eşitlik koyulduğunda, basit doğrusal regresyon modeli (1.13)'deki biçimde elde edilir.

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad i = 1,2, \dots, n \tag{1.13}$$

Modelde bulunan ε_i, *i.* gözleme karşılık gelen hata terimini temsil etmektedir. β_0 ve β_1 bilinmeyen parametrelerdir. Daha önceki kısımlarda ortalama fonksiyonu olarak adlandırılan $\beta_0 + \beta_1 X_i$ modelin deterministik bölümünü oluşturur. Söz konusu deterministik bölüm bağımlı değişkeninin anakütle ortalamasının, X'in değerlerine bağlı olarak doğrusal bir fonksiyon ile tanımlandığını ifade etmektedir. Basit doğrusal regresyonda bağımlı değişkenin ortalaması bağımsız değişkenin değerinin artıp azalmasına bağlı olarak sabit oranda değişir (Rawlings, Pantula, ve Dickey, 1998). Varyans fonksiyonu olarak da adlandırılan ε_i ise modelin rassal bölümünü oluşturur. ε_i bağımlı değişkenin ortalama ile ilgili değişkenliğini açıklamaktadır (Freund, Wilson ve Sa, 2006, s. 37).

Regresyon parametreleri genellikle bilinmez ve veri kümesinin incelenmesi sonunda ancak kestirilebilirler. Parametre değerlerinin hesaplanmasında tüm anakütledeki birimlere ulaşılması ve değerlerinin hesaplanması mümkün olmadığından örneklemeye başvurulur (Şıklar, 2000, s. 6).

1.2.3. Çoklu doğrusal regresyon modeli

Çoklu doğrusal regresyon modelinde bir adet bağımlı değişken bulunurken, basit doğrusal regresyon modelinin aksine birden fazla bağımsız değişken yer alır. Çoklu doğrusal regresyon analizi ile bağımsız değişkenlerin kendi aralarındaki doğrusal ilişkiler göz önüne alınmaksızın, bağımsız değişkenler ile bağımlı değişken arasındaki ilişkinin yapısı doğrusal olarak ortaya konulmaya çalışılır.

Basit doğrusal regresyon modelinde koşullu beklenen değeri veren ortalama fonksiyonu (1.4)'teki biçimde ifade edilmekteydi. Çoklu doğrusal regresyon modelinde ise ortalama fonksiyonu (1.15)'te gösterildiği gibi ifade edilir.

$$\mu = \mathrm{E}(\mathrm{Y}|\mathrm{X} = \mathrm{x}) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_i X_i \qquad (i=1,\dots,n) \qquad (1.15)$$

Varyans fonksiyonu ise basit doğrusal regresyon modelindeki ile aynıdır. Ortalama fonksiyonu ve varyans fonksiyonunu kullanarak çoklu doğrusal regresyon modeli (1.16)'deki haliyle ifade edilir.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_i X_i + \varepsilon_i \qquad (i=1,\dots,n) \qquad (1.16)$$

Modelde yer alan Y_i bağımlı değişkeni; X_i bağımsız değişkenleri ifade etmektedir. β_0 kesişim noktasını göstermektedir ve tüm X_i değerleri sıfıra eşit olduğunda, Y_i bağımlı değişkeninin alacağı değeri gösterir. β_i'ler ise bilinmeyen regresyon parametreleridir ve ait oldukları bağımsız değişkenlerin bağımlı değişkeni ne ölçüde ve ne yönde etkilediklerini belirtirler.

1.3. Parametre Kestirimi

Bağımlı değişkenin koşullu ortalama değerini ve regresyon modelindeki bilinmeyen parametreleri kestirmek için en sık kullanılan yöntemler En Küçük Kareler (EKK) Yöntemi ve En Küçük Mutlak Sapmalar (EKMS) Yöntemi'dir.

1.3.1. En küçük kareler yöntemi

β_0 ve β_1 bilinmeyen parametrelerinin kesin değerlerini belirleyebilmek için anakütledeki tüm birimlerin değerlerine ulaşmak gereklidir. Anakütledeki tüm birimlere ulaşmak çoğu zaman çok zor, hatta bazen imkansız, ve/veya yüksek maliyetli olduğundan dolayı örneklemeye başvurulur. Örnekleme sonucunda elde edilen veri kümesi incelenerek $\beta_0, \beta_1,\ldots, \beta_i$ parametrelerinin kestirimleri olan $b_0, b_1,\ldots,$ b_i elde edilir.

1.3.1.1. Basit doğrusal regresyon modeli

Basit doğrusal regresyon modelinin parametreleri olan β_0, β_1 kestirimlerinin yer aldığı ve Y_i bağımlı değişkeninin kestirim değerlerini veren basit doğrusal regresyon denklemi (1.17)'de verilen biçimde ifade edilir:

$$\hat{Y}_i = b_0 + b_1 X_i \tag{1.17}$$

$\hat{Y}_i$, b_0 ve b_1 belirlendikten sonra verilen X değeri için bağımlı değişkenin ortalamasının tahmin edilen değerini ifade etmektedir.

En küçük kareler yönteminde, b_0 ve b_1 kestirimlerini elde edebilmek için hata terimlerinin karelerinin toplamı enküçüklenir. $\hat{Y}_i$ teorik değeri ile Y_i gözlem değeri arasındaki farklar hata terimlerini oluşturur. Bu amaçla hata terimi

$$e_i = Y_i - \hat{Y}_i \tag{1.18}$$

biçiminde ifade edilebilir. Hata terimi e_i'lerin cebirsel toplamı sıfıra eşittir.

$$\sum_{i=1}^{n} e_i = \sum_{i=1}^{n} (Y_i - \hat{Y}_i) = 0 \tag{1.19}$$

EKK yöntemi temel olarak (1.19)'daki eşitliğin karesinin enküçüklenmesi ilkesine dayanır. Bu amaçla eşitliğin her iki tarafının karesi alınır:

$$min \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 \tag{1.20}$$

(1.20)'de verilen eşitlikte $\hat{Y}_i$ yerine (1.17) yazıldığında denklem

$$min \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{n}(Y_i - b_0 - b_1X_i)^2 \quad (1.21)$$

halini alır. (1.21)'in enküçüklenmesi amacıyla denklem b_0 ve b_1'e göre türevleri alınarak sıfıra eşitlenir:

$$\frac{\partial e}{\partial b_0} = -2\sum_{i=1}^{n}(Y_i - b_0 - b_1X_i = 0 \quad (1.22)$$

$$\frac{\partial e}{\partial b_1} = -2\sum_{i=1}^{n}X_i(Y_i - b_0 - b_1X_i) = 0 \quad (1.23)$$

Elde edilen bu iki eşitlik sadeleştirildiğinde normal denklemler adı verilen denklemler elde edilir:

$$\sum_{i=1}^{n}Y_i = b_0n + b_1\sum_{i=1}^{n}X_i \quad (1.24)$$

$$\sum_{i=1}^{n}X_iY_i = b_0\sum_{i=1}^{n}X_i + b_1\sum_{i=1}^{n}X_i^2 \quad (1.25)$$

Normal denklemlerin çözülmesi ile b_0 ve b_1 katsayıları elde edilir:

$$b_0 = \bar{Y} - b_1\bar{X} \quad (1.26)$$

$$b_1 = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2} \quad (1.27)$$

b_0 ve b_1 elde edildiğinde bağımlı değişkenin kestirimlerini veren regresyon doğrusunun denklemi (1.28)'de verilen biçimde ifade edilir:

$$\hat{Y} = b_0 + b_1X \quad (1.28)$$

1.3.1.2. Çoklu doğrusal regresyon modeli

Daha önce de ifade edildiği gibi bir bağımlı değişken ve birden fazla bağımsız değişkenin olduğu regresyon modelleri çoklu doğrusal regresyon modeli olarak adlandırılmaktadır. Çoklu doğrusal regresyon modelinin parametre kestirimlerinin EKK yöntemi ile yapılabilmesi için (1.16)'da verilen model, matrislerle ifade edilmelidir. Çoklu doğrusal modelin matrislerle gösterimi (1.29)'da gösterildiği biçimde ifade edilir:

$$Y = X\beta + \varepsilon \quad (1.29)$$

(1.29) denkleminde yer alan matrisler (1.30)'daki gibi yazılır:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \tag{1.30}$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{bmatrix}$$

Y, *n* x 1 boyutlu bağımlı değişken vektörü; *X*, *n* x *p*'lik bağımsız değişken matrisi; *β*, *p* x 1 boyutlu regresyon katsayıları vektörü ve *ε* ise *n* x 1 boyutlu hata terimleri vektörüdür. (1.29)'daki regresyon modelinin kestirimlerini veren regresyon denklemi matrislerle (1.31)'deki biçimde tanımlanır:

$$\hat{Y} = X\hat{\beta} \tag{1.31}$$

$\hat{\beta}$, *p* x 1 boyutlu regresyon kestirimlerini veren vektördür:

$$\hat{\beta} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_p \end{bmatrix} \tag{1.32}$$

Basit doğrusal regresyonda olduğu gibi çoklu doğrusal regresyonun parametreleri kestirilirken de hata kareleri en küçüklenir:

$$\min S(\beta) = \sum_{i=1}^{n} e_i^2 = \varepsilon'\varepsilon = (Y - X\beta)'(Y - X\beta) \tag{1.33}$$

$$S(\beta) = y'y - \beta'X'Y - y'X\beta + \beta'X'X\beta \tag{1.34}$$

$$= y'y - 2\beta'X'Y + \beta'X'X\beta$$

S(β) fonksiyonu en küçüklendiğinde parametre kestirimini veren $\hat{\beta}$ matrisi de elde edilmiş olur. Buna göre *S(β)* fonksiyonunun *β*'ya göre türevi alınıp sıfıra eşitlenir:

$$\frac{\partial S}{\partial \beta} = -2X'Y + 2X'X\hat{\beta} = 0 \tag{1.35}$$

Basitleştirildiğinde en küçük kareler kestiricisi elde edilir:

$$\hat{\beta} = (X'X)^{-1}X'Y \tag{1.36}$$

Buradan Y gözlem değerlerinin kestirimlerini veren $\hat{Y}$ kestirim vektörü şu hali alır:

$$\hat{Y} = X\hat{\beta} = X(X'X)^{-1}X'Y \tag{1.37}$$

$$\hat{Y_i} = b_0 + b_1X_{ip} + b_2X_{ip} + \cdots + b_pX_{ip} + e_i \quad (1.38)$$

1.3.2. En küçük mutlak sapmalar yöntemi

EKMS yöntemi, en küçük kareler yöntemi gibi regresyon modelindeki bilinmeyen parametrelerin kestiriminde kullanılan alternatif bir yöntemdir. Söz konusu parametrelerin kestirimi için en küçük kareler yönteminde hata karelerinin toplamı ($\sum e_i^2$)enküçüklenmeye çalışılırken, en küçük mutlak sapmalar yönteminde ise hataların mutlak değerlerinin toplamı ($\sum|e_i|$)enküçüklenir:

$$min \sum\left|Y_i - \hat{Y}_i\right| \quad (1.39)$$

$$min \sum|Y_i - (b_0 + b_1X_i)| \quad (1.40)$$

EKK yöntemine kıyasla en küçük mutlak sapmalar yönteminin en büyük dezavantajı, belirli bir formüle sahip olmayışıdır. Enküçüklenmek istenen denklem 1. dereceden bir fonksiyon olduğu için EKK yöntemindeki gibi türevi alınarak çözülemez. Bu yüzden fonksiyonun çözümü için çeşitli algoritmalardan yararlanılır.

1.3.2.1. Basit doğrusal regresyon modeli

EKMS yöntemi ile hesaplanan regresyon doğrusunun iki gözlem noktasından geçtiği varsayılmaktadır. Bu yüzden geliştirilen algoritmanın amacı da iki gözlem noktasından geçen en küçük mutlak sapmaya sahip doğruyu belirlemektir. Algoritma herhangi bir (x_0, y_0) gözlem noktasından başlar ve bu noktadan geçen en küçük mutlak sapmalı doğru belirlenir. Belirlenen doğru bir başka gözlem noktasından (x_1, y_1) da geçmektedir. Bu işlemden sonra (x_1, y_1) noktasından geçen doğru belirlenir. Belirlenen bu yeni doğru da bir başka gözlem noktası olan (x_2, y_2) noktasından geçmektedir. Daha sonra aynı işlem (x_2, y_2) noktası için de yapılır ve aynı doğru iki defa arka arkaya bulunduğunda algoritma durur. En son belirlenen doğru EKMS yöntemi ile hesaplanan regresyon doğrusudur (Birkes ve Dodge, 1993, s. 58).

Basit doğrusal regresyon doğrusunun EKMS kestirimleri için geliştirilen algoritmanın adımları aşağıdaki gibidir:

1) *n* tane gözlem noktasından herhangi biri başlangıç noktası (x_0, y_0) olarak seçilir.

2) Her gözlem noktası (x_i, y_i) ile (x_0, y_0) arasındaki eğim (1.41)'de verilen işlem ile hesaplanır:

$$\frac{y_i - y_0}{x_i - x_0} \tag{1.41}$$

Eğer $x_i = x_0$ ise nokta göz ardı edilir.

3) Bulunan eğimler,

$$\frac{y_1 - y_0}{x_1 - x_0} \leq \frac{y_2 - y_0}{x_2 - x_0} \leq \cdots \leq \frac{y_n - y_0}{x_n - x_0} \tag{1.42}$$

olacak biçimde sıralanır.

4) Kritik *T* değeri, her bir x_i değerinin x_0 değerinden mutlak sapmalarının toplamı biçiminde elde edilir:

$$T = \sum_{i=1}^{n} |x_i - x_0| \tag{1.43}$$

5) *T*'nin yarısından büyük olan k değeri koşulu sağlayan değerdir. Bu durum (1.44)'te verilen biçimde ifade edilir:

$$|x_1 - x_0| + \cdots + |x_{k-1} - x_0| < \frac{1}{2}T \tag{1.44}$$

$$|x_1 - x_0| + \cdots + |x_k - x_0| > \frac{1}{2}T$$

6) k noktası (x_k, y_k) yeni göz önüne alınacak noktadır ve (x_0, y_0) nokta çiftinden geçen en küçük sapmalı doğru bu noktadan geçmektedir. Doğrunun eğimi β_1 katsayısının kestirimini verirken, ordinatı kestiği nokta ise β_0 katsayısının kestirimini verir:

$$b_1 = \frac{y_k - y_0}{x_k - x_0} \tag{1.45}$$

$$b_0 = y_0 - \beta_1 x_0 \tag{1.46}$$

$$\hat{Y} = b_0 + b_1 X \tag{1.47}$$

7) 1. adıma dönülerek (x_0, y_0) yerine (x_k, y_k) yazılır ve aynı işlemler k noktası için uygulanır. Arka arkaya iki iterasyonda aynı sonuç bulunduğunda algoritma durur (Vural, 2007, s. 35)

Algoritmanın iki tane sorunu vardır. Bunlar tek olmama ve bozulma sorunlarıdır. Tek olmama ile ifade edilmek istenen sorun, bir gözlem noktasından birden fazla en iyi doğrunun geçmesidir. Bozulma ise bir noktadan geçen en iyi doğrunun başka iki veya daha fazla noktadan da geçmesidir. Bu sorunların mevcut olduğu bir veri kümesi için parametre kestirimi yapılmak istendiğinde algoritma kısır bir döngüye girer. Tek olmama sorunu ile karşılaşıldığında doğrulardan herhangi biri tercih edilerek algoritma devam edebilirken; bozulma durumunda algoritma ilerleyemez (Akyol, 2013, s. 15)

1.3.2.2.Çoklu doğrusal regresyon modeli

EKMS yöntemi ile çoklu doğrusal regresyon modelinin parametre kestirimleri yapılırken de toplam mutlak sapmalar enküçüklenir:

$$min \sum |Y_i - (b_0 + b_1 X_{i1} + b_2 X_{i2} + \cdots + b_k X_{ik})| \qquad (1.48)$$

Bunun için öncelikle regresyon katsayı kestiricileri ve X değerleri vektör gösterimi yardımıyla ifade edilmelidir:

$$b = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_p \end{bmatrix} \text{ ve } X_i = \begin{bmatrix} 1 \\ X_{i1} \\ X_{i2} \\ \vdots \\ X_{ip} \end{bmatrix} \qquad (1.49)$$

Buna göre mutlak sapmalar toplamı $\sum |Y_i - b' X_i|$ biçiminde yazılır. Mutlak sapmaların enküçüklenmesi için b' vektörü kestirilmelidir. Basit doğrusal regresyon kestirimleri için kullanılan algoritmanın her iterasyonunda daha iyi bir doğru belirlenmekteyken; çoklu doğrusal regresyonun parametrelerinin kestirilmesi için algoritmanın her adımında daha az mutlak sapmalı bir vektör belirlenir. Bu amaçla algoritma bir b başlangıç vektörü ile başlar ve daha az sapmaya sahip bir vektör olan b^* elde edilir. b^* vektörü d yön vektörü ile t uzunluğunun bulunması ile (1.50)'de verildiği gibi hesaplanır:

$$b^* = b + td \qquad (1.50)$$

Buna göre enküçüklenecek fonksiyon (1.51)'deki hali alır:

$$\sum |Y_i - (b + td)' X_i| \qquad (1.51)$$

Öncelikle t'yi enküçükleyen değer belirlenir. Fonksiyonda $Z_i = Y_i - b' X_i$ ve $W_i = d' X_i$ dönüşümü gerçekleştirildiğinde

$$\sum |Z_i - t W_i| \qquad (1.52)$$

elde edilir. Sonraki adımda ${Z_i}/{W_i}$ oranları hesaplanır ve küçükten büyüğe doğru sıralanır. Küçükten büyüğe doğru sıralanan değerlere karşılık gelen $|W_i|$ değerleri kümülatif olarak toplanır. Buna göre aşağıdaki koşulu sağlayan k. indis bulunur.

$$|W_1| + |W_2| + \cdots + |W_{k-1}| < \tfrac{1}{2}T \tag{1.53}$$

$$|W_1| + |W_2| + \cdots + |W_{k-1}| + |W_k| > \tfrac{1}{2}T$$

$T = \sum |W_i|$'dir. Buradan t'yi enküçükleyen değer ${Z_k}/{W_k}$ olur.

Algoritmanın her adımında, p tane bağımsız değişkenin olduğu bir model için (p+1) tane yön vektörü göz önüne alınır. Her vektörün negatif değerleri de göz önüne alındığında 2(p+1) tane d_j yön vektörü elde edilmiş olur. Yön vektörlerini belirlemek için öncelikle örneklem regresyon denklemi $\hat{Y}_i = b'X_i$ matris notasyonu ile ifade edilir:

$$A = \begin{bmatrix} X'_{i1} \\ X'_{i2} \\ \vdots \\ X'_{ip+1} \end{bmatrix}, c = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{p+1} \end{bmatrix} \tag{1.54}$$

Buradan $b = A^{-1}c$ olur. A^{-1} vektörünün sütunları yön vektörlerini verir. b'den daha iyi vektörü belirlemek için

$$b^* = b + td_j \tag{1.55}$$

hesaplanır.

Bu yön vektörlerinden en iyisi (1.50)'yi t=0 noktasına en hızlı biçimde indirendir. Bunu belirleyebilmek için ise denklemin sağ tarafının t=0 noktasındaki türevi alınır. (1.50) denkleminin t = 0 noktasındaki türevi $W_- + W_0 + W_+$ olur. Türev 2(p+1) tane olan her yön için hesaplanır. En küçük negatif türeve sahip olan yön en uygun olandır. Tüm türevler pozitif olduğunda b vektörü en iyi kestirimleri veren türev olarak belirlenir ve algoritma durur (Birkes ve Dodge, 1993, s. 71).

Algoritmanın adımları:

1) A matrisi ile c vektörü belirlenir.
2) b başlangıç tahmin vektörü hesaplanır.
3) Başlangıç yön vektörleri A^{-1}'den oluşturulur.
4) (1.49) denklemine Z_i, w_i dönüşümü gerçekleştirilir.

5) ${}^{z_i}/_{w_i}$'ler oluşturulur.

6) (1.50) denkleminin türevi alınır ve en küçük negatif değere sahip olan yön vektörü d_j seçilir.

7) Küçükten büyüğe doğru sıralanan ${}^{z_i}/_{w_i}$'ler arasından k. indis seçilerek t'nin en küçük değeri bulunur.

8) Bulunan t ve d_j değerlerinden yararlanılarak (1.53) denkleminden b^* değeri bulunur.

9) A matrisinin j. satırı bulunan X'_k değeri ile değiştirilir. Oluşan yeni matris A^* olarak adlandırılır. Yeni yön vektörleri A^{*-1} matrisinin sütunlarından elde edilir.

Bulunan k indisi tekrarlandığında algoritma durur (Altındağ, 2010, s. 36).

İKİNCİ BÖLÜM

KANTİL REGRESYON

EKK yöntemini kullanarak parametre kestirimlerini gerçekleştiren geleneksel regresyon yaklaşımlarında bağımlı değişkenin koşullu ortalaması, bağımsız değişkenin bir fonksiyonu ile tanımlanmaktadır. Sabit X değerlerinden oluşan bir kümeden yararlanılarak, bağımlı değişken Y'nin koşullu ortalama değerleri kestirilmektedir. EKK yöntemi ile örneklem regresyon denkleminin oluşturulması ve de denklemin yorumlanması oldukça kolaydır. Bu avantajına rağmen, bağımlı değişkenin dağılımı asimetrik olduğunda, EKK yönteminin sabit varyanslılık varsayımı ihlal edilmiş olur. Bu durumda değişen varyanslılık (heteroscedasticity) adı verilen durum ortaya çıkar. Bunun sebebi ortalamanın aşırı düşük veya yüksek değerlere karşı hassas olmasıdır.

Kantil regresyon analizi, Koenker ve Bassett (1978) tarafından bağımlı değişken normal dağılmadığı durumlarda ortaya çıkan söz konusu güçlüğü aşmak için medyan regresyondan yola çıkarak geliştirilmiştir. Medyan regresyon, bağımsız değişkende gerçekleşen bir değişimin, bağımlı değişkenin dağılımının 0,50. kantilindeki, diğer bir ifadeyle medyandaki, değişimi açıklamaktaydı. Kantil regresyon ise bağımlı değişkenin dağılımının çeşitli kantillerindeki değerlerini tahmin etmektedir. Bağımsız değişkende gerçekleşen bir değişimin bağımlı değişkene ait koşullu kantillerde nasıl bir değişime yol açtığı incelenmektedir. Dağılımın her iki ucundaki kantiller incelendiğinden ve koşullu kantil değerleri tahmin edildiğinden dolayı, asimetrik dağılıma sahip bağımlı değişkenin değerleri daha doğru ve yansız olarak kestirilebilir.

2.1. Kantil Kavramı

Seriler değişkenlik ve asimetri konularında çeşitli ölçülerin hesabında kullanılmak üzere, küçükten büyüğe doğru sıralanarak 2, 4, 10 ve 100 eşit parçaya bölünürler. Seriyi bölen bu değerler kantil olarak adlandırılır. Kantillerden bir seriyi iki eşit parçaya bölen değerler medyan, dörde eşit parçaya bölenler kartil, on eşit parçaya bölenler desil ve yüz eşit parçaya bölenler de santil olarak adlandırılırlar

(Serper, 2014). Bir seride bir adet medyan, üç adet kartil, dokuz adet desil ve doksan dokuz santil bulunur.

2.2. Kantil Fonksiyonları

Bir dağılımın şeklini ve konumunu belirleyebilmek için istatistiksel araştırmalarda yaygın olarak aritmetik ortalama, medyan, mod gibi merkezi eğilim ölçülerinden ve standart sapma, varyans gibi merkezi yayılım ölçülerinden faydalanılır. Ancak çarpık ve basık dağılımların şeklini ve konumunu belirlemek için bu ölçüler yeterli olamamaktadır. Bu yüzden izleyen kesimde anakütlenin dağılımının şeklini ve konumunu tanımlamada kullanılan kümülatif dağılım fonksiyonu, olasılık yoğunluk fonksiyonu, kantil fonksiyonu ve kantil yoğunluk fonksiyonu açıklanacaktır.

2.2.1. Kümülatif dağılım fonksiyonu

Bir bağımlı değişkenin dağılımını açıklamak için dağılımın kümülatif dağılım fonksiyonundan yararlanılır (Hao, Naiman, 2007, s. 7). Kümülatif dağılım fonksiyonu, bir X değişkeninin verilen bir x değerinden daha küçük olma olasılığını veren, monotonik artan ve sürekli bir fonksiyondur. *F(x)* ile gösterilir:

$$F(x) = Olasılık(rassal\ değişken\ X \leq x)$$

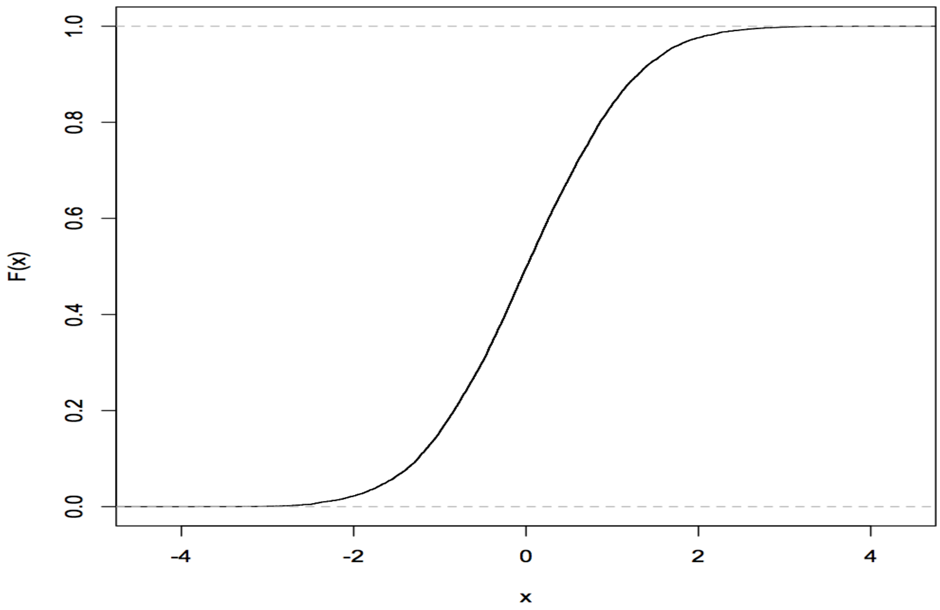
$$F(x) = P(X \leq x) \tag{2.1}$$

$$0 \leq F(x) \leq 1, x \in R$$

Şekil 2.1'de görüldüğü gibi kümülatif dağılım fonksiyonu *F(x)*, her zaman 0 ile 1 arasında değerler alır. Ayrıca *F(x)* soldan sıfır değeriyle başlar ve sağa doğru gittikçe 1 değerine yaklaşır. x değeri $-\infty$'a yaklaştıkça *F(x)* fonksiyonunun değeri 0'a, x $+\infty$'a yaklaştıkça da *F(x)* fonksiyonunun değeri 1'e yaklaşır (Huang, 2009, s. 6):

$$\lim_{x \to -\infty} F(x) = 0$$

$$\lim_{x \to +\infty} F(x) = 1 \tag{2.2}$$



Şekil 2.1. *Kümülatif Dağılım Fonksiyonu*

2.2.2. Olasılık yoğunluk fonksiyonu

Olasılık yoğunluk fonksiyonu, bir değişkenin alabileceği değerlerle bu değerleri alma olasılıkları arasındaki bağıntıyı gösteren fonksiyondur.

$$f(x)dx = P(x \leq X \leq x + dx) \qquad (2.3)$$

Olasılık yoğunluk fonksiyonu, kümülatif dağılım fonksiyonunun türevine eşittir. Kümülatif dağılım fonksiyonu ile olasılık yoğunluk fonksiyonu arasındaki ilişki aşağıdaki biçimde ifade edilir:

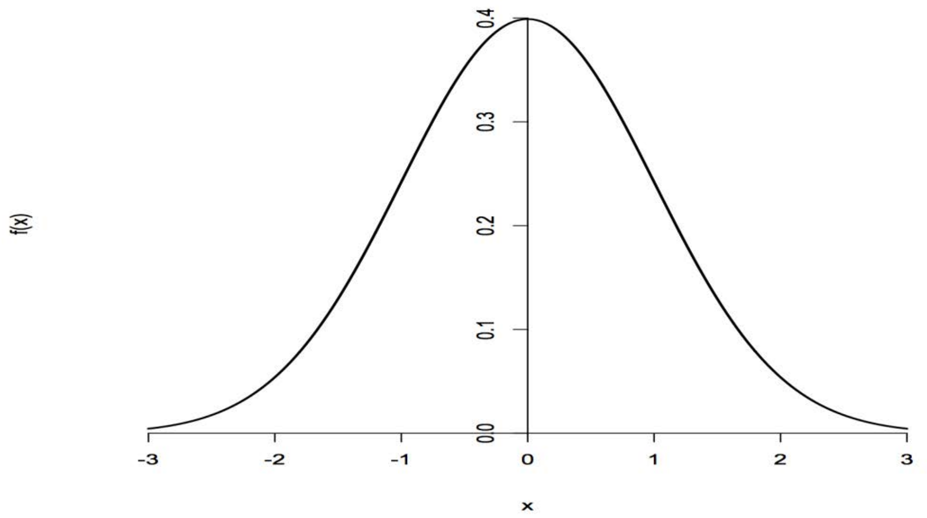
$$f(x)dx = F(x + dx) - F(x) = dF(x) \qquad (2.4)$$

Şekil 2.2'de gösterildiği üzere *f(x)* eğrisinin altında kalan alan, gözlemlenen değerlerin alabileceği olasılık değerlerinin toplamını göstermektedir ve 1'e eşittir. Bu durum, eğrinin altında kalan alanın gözlemlenen değerlerin alabileceği olasılık değerlerinin toplamını gösterdiği anlamına gelir. Olasılık yoğunluk fonksiyonu;

$$f(x)dx = Olasılık(x \leq rassal\ değişken\ X \leq x + dx)$$

$$f(x)dx = P(x \leq X \leq x + dx) \qquad (2.5)$$

biçiminde tanımlanır.



Şekil 2.2. *Olasılık Yoğunluk Fonksiyonu*

2.2.3. Kantil fonksiyonu

X reel sayılar kümesinde tanımlı rassal bir değişken, x_p, anakütlenin p. kantilini gösteren değer olsun. Bu durumda $x_p = Q(p)$ kantil fonksiyonu, X değerinin x_p'den küçük olma olasılığını belirtir.

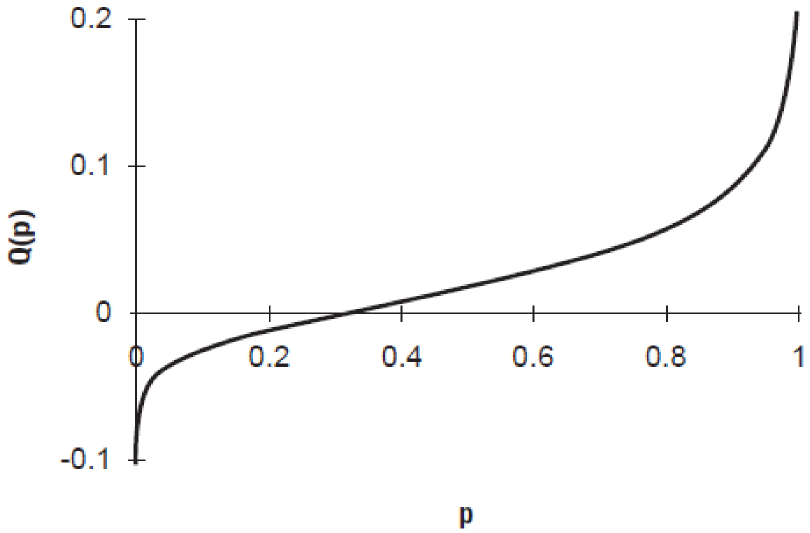
$$Q(p) = P(X \leq x) = p \tag{2.6}$$

$Q(p)$ fonksiyonu F kümülatif dağılım fonksiyonunun kantil fonksiyonu olarak adlandırılır. Kümülatif dağılım fonksiyonu ile kantil fonksiyon birbirilerinin tersine eşittir (Parzen, 1979, s. 106):

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

$$Q(p) = F^{-1}(p) = \inf\{x: F(x) \geq p\} \quad 0 < p < 1 \tag{2.7}$$

Örnek bir kantil fonksiyonu grafiği Şekil 2.3'te verilmiştir.

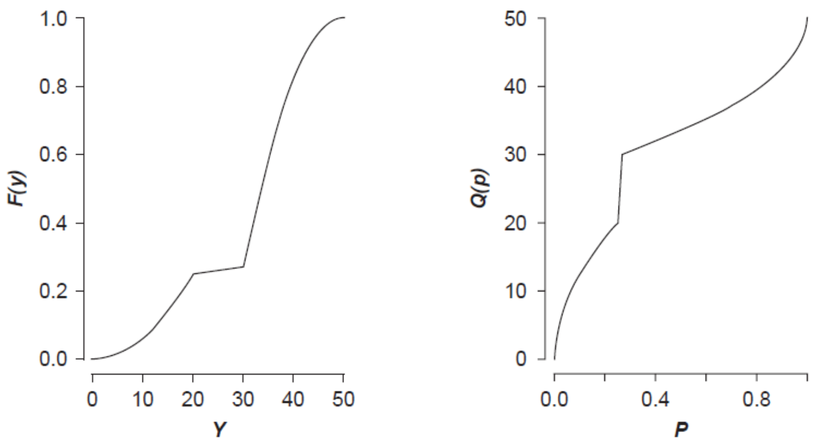


Şekil 2.3. *Kantil Fonksiyonu*

Kaynak: *Gilchrist, 2000, s. 13*

(2.7)'ye göre *Q(p) X*'in *p.* kantili olarak adlandırılır. Bu durumda örneğin medyan değeri Q(0.5) ile 1. ve 3. kartiller de sırasıyla Q(0.25) ve Q(0.75) ile ifade edilir (Koenker, 2005, s. 5).

Kümülatif dağılım fonksiyonunun kantil fonksiyonun tersi olması durumu Şekil 2.4.'te belirtilmiştir. Şeklin sol tarafında kümülatif dağılım fonksiyonu, sağ tarafındaysa kümülatif dağılım fonksiyonunun tersi olan kantil fonksiyonu yer almaktadır.



Şekil 2.4. *Kümülatif Dağılım Fonksiyonu ve Kantil Fonksiyonu*

Kaynak: *Hao, L; Naiman, 2007, s. 11*

Q(p) kantil fonksiyonu p'nin tüm olasılıkları için $0 < p < 1$ kantil değerlerini verir. Örneğin; $Q(0.5)$ medyan değerini, $Q(0.25)$ ve $Q(0.75)$ diğer kartil değerlerini vermektedir.

2.2.4. Kantil yoğunluk fonksiyonu

Olasılık yoğunluk fonksiyonu, kümülatif dağılım fonksiyonunun türevine eşit olduğu gibi kantil yoğunluk fonksiyonu da kantil fonksiyonun türevine eşittir. Dağılımları tanımlamak ve modellemek için gerekli değerleri elde etmede kullanılır.

$$q(p) = dQ(p)/dp \tag{2.8}$$

Kantil fonksiyon azalmayan bir fonksiyon olduğu için eğimi hiçbir zaman negatif olamaz. Bu yüzden *q(p)* kantil yoğunluk fonksiyonu her zaman 0 ile 1 arasında yer alır, $0 \leq p \leq 1$ (Altındağ, 2010, s. 20).

Bu konuda olasılık yoğunluk fonksiyonundan farklılaşır. Çünkü olasılık yoğunluk fonksiyonu sonsuz tanım aralığına sahiptir. Kantil yoğunluk fonksiyonu ve olasılık yoğunluk fonksiyonu arasındaki ilişki (2.9)'da tanımlanmıştır (Jones, 1992, s. 722):

$$q(p) = \frac{1}{f(F^{-1}(p))}, \qquad f(x) = \frac{1}{q(Q^{-1}(x))} \tag{2.9}$$

2.3. Kantillere Dayanan Konum ve Şekil Ölçüleri

Bir dağılımın şekli basıklık ve çarpıklık özellikleri ile incelenir. Genellikle basıklığın ölçülmesinde standart sapmadan yararlanılır. Gözlem değerlerinin ortalamadan uzaklıklarının bir ölçüsü olan standart sapma, simetrik dağılımların basıklığını tanımlamada oldukça etkili iken asimetrik dağılımlarda yanlış sonuçlar vermektedir. Gerçek hayatta çoğu dağılımın asimetrik olduğu göz önüne alındığında standart sapmanın basıklığı tanımlamak için kullanılması yanlış değerlendirmelere yol açar. Bu yüzden, özellikle asimetrik dağılımların basıklığını ölçmek için kantil tabanlı basıklık ölçüsü (QSC) geliştirilmiştir:

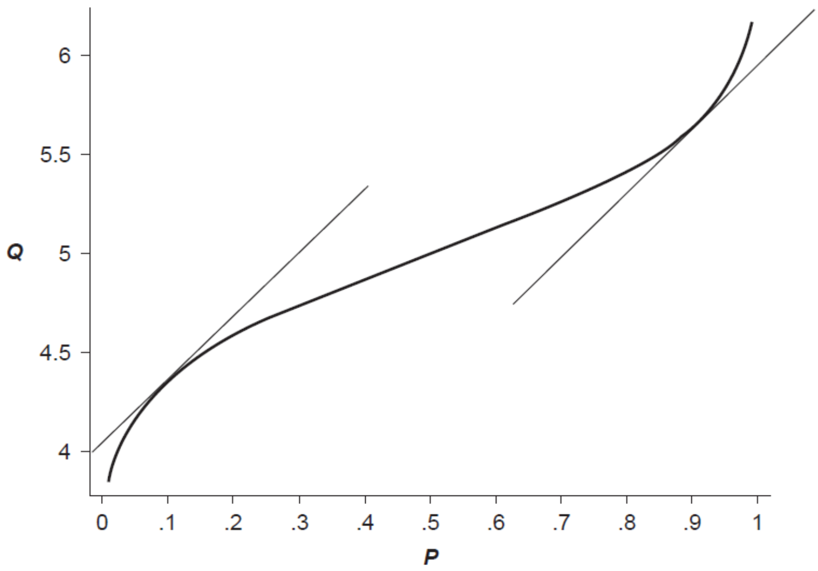
$$QSC^{(p)} = Q^{(1-p)} - Q^{(p)} \quad p < 0.5 \tag{2.10}$$

(2.10)'da görüldüğü üzere dağılımın orta kısmında kalan belirli bir kısmın basıklığını ölçmek için birbirilerinin tümleyeni olan iki kantil fonksiyon değeri birbirinden çıkarılır. Örneğin; dağılımın %90'lık orta kısmının basıklığı açıklanmak istendiğinde $Q^{(0.95)}$ değerinden $Q^{(0.05)}$ değeri çıkarılır.

Dağılımın şeklini tanımlamada kullanılan bir diğer ölçü olan çarpıklığı tanımlamak için ise gözlem noktalarının ortalamadan sapmalarının kübik bir fonksiyonu ile ölçüm yapılır. Simetrik bir dağılımın çarpıklığı sıfıra eşittir. Ölçünün negatif olması dağılımın sola çarpık, pozitif olması da dağılımın sağa çarpık olduğu anlamına gelir.

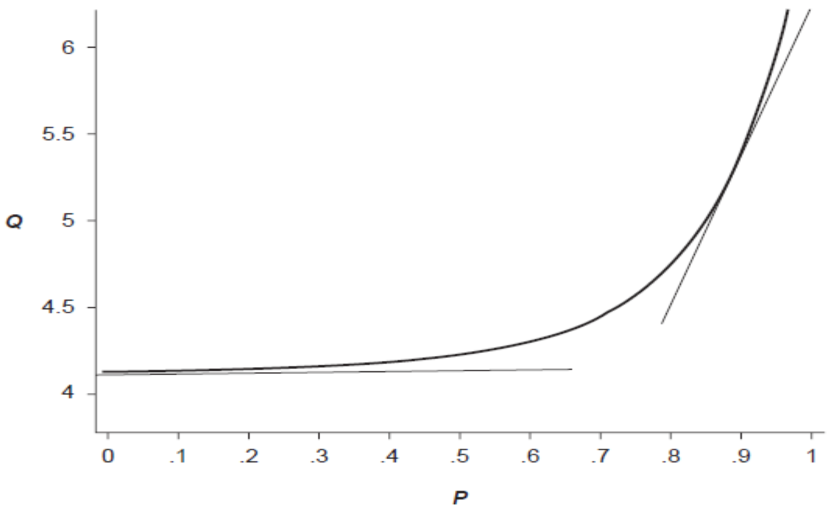
Normal dağılımlarda birbirinin simetriği olan iki kantil fonksiyonun eğimleri birbirine eşittir. Şekil 2.5 incelendiğinde 0,10 ve 0,90 kantillerinin eğimlerinin birbirlerine eşit olduğu gözlemlenebilir. Buna karşın Şekil 2.6.'da gösterilen sağa çarpık bir dağılımda kantil fonksiyon değerleri asimetrik olur. Bu yüzden kantil tabanlı çarpıklık ölçüsü (QSK) geliştirilmiştir (Hao, L; Naiman, 2007, s. 13):

$$QSK^{(p)} = \left(\frac{Q^{(1-p)} - Q^{(0.5)}}{Q^{(0.5)} - Q^{(p)}}\right) - 1 \quad p < 0.5 \qquad (2.11)$$



Şekil 2.5. *Normal Kantil Fonksiyonu*

Kaynakça: *Hao, L; Naiman, 2007, s. 15*



Şekil 2.6. *Çarpık Kantil Fonksiyonu*

Kaynakça: *Hao, L; Naiman, 2007, s. 15*

2.4. Kantil Regresyon Modeli

Kantil regresyon modeli, bağımlı değişkenin dağılımındaki koşullu kantilleri tahmin etmek için kullanılan doğrusal bir modeldir. Mosteller ve Tukey (1977), bağımlı değişkenin dağılımının çeşitli kantilleri için regresyon eğrileri geliştirilerek değişkenler arasındaki ilişkinin daha doğru ve eksiksiz biçimde açıklanabileceğini öne sürmüştür. Kantil regresyon modeli ilk olarak Boskovich'in 18. yüzyılda koşullu-medyan regresyon üzerine yaptığı çalışmaları geliştiren Koenker ve Bassett, (1978) tarafından ortaya atılmıştır (Hao, L; Naiman, 2007, s. 29).

Özellikle değişen varyansın olduğu dağılımların modellenmesinde etkin ve yansız sonuçlar üreten bir yöntemdir. EKK yöntemi ile kestirilen bir doğrusal regresyon modeli ile değişen varyansa sahip bir bağımlı değişkenin dağılımının koşullu ortalaması, anakütle parametresinden fazla veya az tahmin edilebilir. Hatta gerçekte bağımlı değişken ile bağımsız değişken arasında istatistiksel olarak anlamlı olmayan bir ilişkinin varlığından söz edilebilir (Cade ve Noon, 2003, s. 412).

EKK yönteminde bağımlı değişkenin koşullu ortalaması tek bir regresyon doğrusu ile modellenmekteyken, kantil regresyonda ise çeşitli kantiller için regresyon doğruları hesaplanabilmektedir. Böylece koşullu kantilleri tahminleyen birden fazla regresyon doğrusu sayesinde bağımlı değişkenin dağılımının tamamını inceleme ve yorumlama imkanı vermektedir (Uyar, Kangallı Uyar ve Gökçe, 2016, s. 590).

Kantil regresyon modeli (2.12)'de verildiği gibi ifade edilir:

$$Y_i = \beta_0^{(p)} + \beta_1^{(p)} X_i + \varepsilon_i^{(p)} \qquad 0 < p < 1 \tag{2.12}$$

p değeri kantil değerini ifade etmektedir ve p. kantilin altında kalan gözlem sayısının toplam anakütle hacmine oranını verir. Doğrusal regresyon analizinde Y'nin koşullu ortalaması, verilen bir X_i değeri için

$$E(Y_i|X_i) = \beta_0 + \beta_1 X_i \tag{2.13}$$

formülü ile hesaplanır ve hata terimi ε_i'nin beklenen değeri sıfıra eşittir. Buna karşın kantil regresyon modelinde verilen bir X_i değeri için p. koşullu kantilin beklenen değeri

$$Q^{(p)}(Y_i|X_i) = \beta_0^{(p)} + \beta_1^{(p)} X_i \tag{2.14}$$

eşitliği ile hesaplanır. Böylece p. koşullu kantil değeri, kantil parametreleri $\beta_0^{(p)}, \beta_1^{(p)}$ ile X_i değerine bağlı olarak değişir. Doğrusal regresyon modelinde olduğu gibi hata terimlerinin beklenen değeri sıfıra eşit olarak kabul edilmektedir (Hao, L; Naiman, 2007, s. 29).

i. gözlem değeri için hata terimi $\varepsilon_i^{(p)}$, kantiller arasında farklı değerlere sahip olur. Örneğin belirli bir *i.* gözlem için p ve q koşullu kantil regresyon doğruları hesaplandığında p. Kantil için (2.14) modeli elde edilirken, q. kantil için (2.15) elde edilir.

$$Q^{(q)}(Y_i|X_i) = \beta_0^{(q)} + \beta_1^{(q)} X_i \tag{2.15}$$

İki denkleme ait hata terimleri birbirinden çıkartıldığında;

$$\varepsilon_i^{(p)} - \varepsilon_i^{(q)} = \beta_0^{(p)} - \beta_0^{(q)} + (\beta_1^{(p)} - \beta_1^{(q)}) X_i \tag{2.16}$$

elde edilir. Görüldüğü üzere belirli bir *i.* gözlemin p ve q gibi iki farklı koşullu kantil regresyon denklemlerine ait hata terimleri arasındaki fark sıfıra eşit değildir. Buradan hata terimlerinin *i.* gözlem için farklı kantillerde farklı değerlere sahip olacağı sonucu çıkartılabilir. Diğer bir ifadeyle hata terimleri gözlem değerlerine değil, kantillere bağlıdır. Ancak kantil regresyon yönteminde $\varepsilon_i^{(p)}$ hata teriminin aynı ve bağımsız olarak dağıldığı ve hata teriminin beklenen değerinin sıfır olduğu varsayılır. EKMS yönteminde olduğu gibi kantil regresyonda da hata terimleri gözlem değerlerinin $Q^{(p)}$ kantil doğrusuna dikey uzaklıkları olarak tanımlanır (Hao, L; Naiman, 2007, s. 29).

2.4.1. Parametre kestirimi

Kantil regresyonda parametre kestirimi, EKMS yönteminde olduğu gibi gözlem değerlerinin kantil regresyon doğrusundan dikey uzaklıklarının mutlak değerleri göz önüne alınarak gerçekleştirilir. Ancak bu kez dikey uzaklıklarının ağırlıklı toplamı enküçüklenmeye çalışılır. Doğrunun altında kalan gözlem noktalarının doğruya uzaklıklarının mutlak değerleri *p*, üstünde kalan gözlem değerlerinin uzaklıklarının mutlak değerleri *(1 - p)* ile çarpılıp toplanırlar. Daha sonra bu toplam enküçüklenerek kantil regresyon parametreleri olan $\beta_0^{(p)}$ ve $\beta_1^{(p)}$'in kestirimleri elde edilir.

Çoklu doğrusal regresyon modeli, $Y_i = X_i'\beta \;\; (i = 1, \dots, n)$ biçiminde ifade edilmekteydi. Hata terimlerinin beklenen değeri sıfır kabul edildiğinde, Y_i bağımlı değişkeninin koşullu ortalaması $E(Y_i|X_i) = X_i{}'\beta$ biçiminde ifade edilir. Bu durumda EKK yöntemi ile β'nın kestirimi sırasıyla (2.17) ve (2.18)'de verildiği biçimde elde edilir:

$$\hat{\beta} = \min_{\beta} \textstyle\sum_i (Y_i - X_i'\beta)^2 \tag{2.17}$$

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

Benzer biçimde kantil regresyonun hesaplanması için de *p*. kantildeki hata terimlerinin beklentisinin sıfıra eşit olduğu varsayılır:

$$Q_{(p)}\left(\varepsilon_{i,p}|X_i\right) = 0 \tag{2.19}$$

Buna göre *p.* koşullu kantil değeri verilen bir X_i değerine göre (2.20)'deki biçimde elde edilir.

$$Q_{(p)}(Y_i|X_i) = X_i{}'\beta_{(p)} \tag{2.20}$$

Buna göre koşullu kantilin ağırlıklı sapmalarının toplamı

$$\hat{\beta}_{(p)} = \min_{\beta} \textstyle\sum_{i=1}^{n} \rho_{(p)}(Y_i - X_i'\beta) \tag{2.21}$$

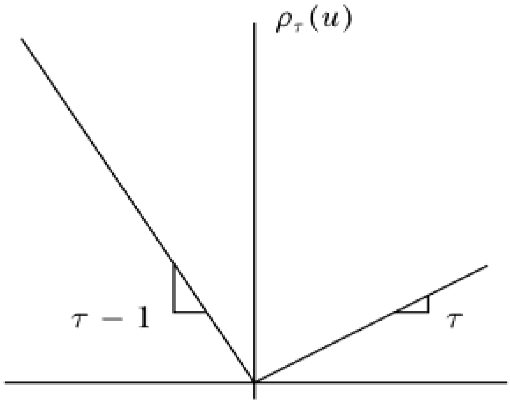
olacak biçimde enküçüklenir. Burada Y_i bağımlı değişkeninin koşullu dağılımı, X_i değerinin *p.* kantil değerleri ile tanımlanmaktadır (Chiang ve Li, 2012, s. 38). Denklemde yer alan $\rho_{(p)}$ise kontrol fonksiyonu olarak adlandırılan mutlak ağırlık faktörüdür (Akyol, 2013, s. 38). Kontrol fonksiyonu, *p.* Kantil regresyon doğrusunun üstünde kalan gözlem değerlerini *p*, altında kalan gözlem değerlerini ise (1 – *p*) ile ağırlıklandırır. Bu durum;

$$\rho_{(p)}(\varepsilon) = \varepsilon(p - I(\varepsilon < 0)) \qquad (2.22)$$

biçiminde ifade edilir. *I*(A) fonksiyonu işaret fonksiyonudur ve *I*(*A*) doğru olduğunda 1, yanlış olduğunda 0 değerini alır (Shulze, 2004, s. 20). Kontrol fonksiyonu

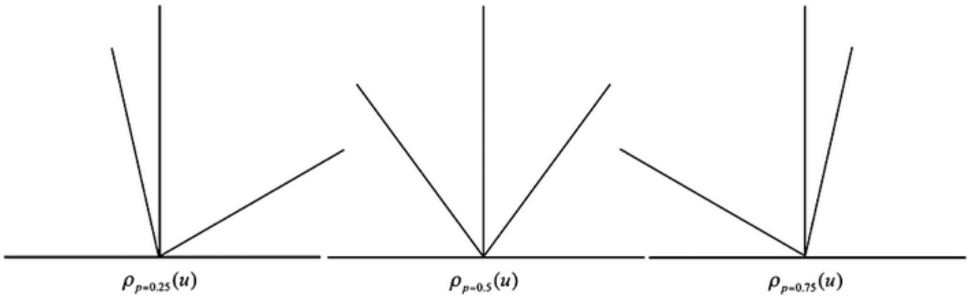
$$\rho_{(p)}(\varepsilon_i) = \begin{cases} p\varepsilon_i & eğer \;\; \varepsilon_i \geq 0, \\ (p-1)\varepsilon_i & eğer \;\; \varepsilon_i < 0 \end{cases} \qquad \varepsilon_i = Y_i - X_i{}'\beta \qquad (2.23)$$

biçiminde de ifade edilebilir. Fonksiyonun grafiği Şekil2.7'de ve kartil değerlerinde aldığı görünümler de Şekil 2.8'de verilmiştir:



Şekil 2.7. *Kontrol Fonksiyonu*

Kaynakça: *Koenker ve Hallock, 2001, s. 146*



Şekil 2.8. Kontrol Fonksiyonunun Kartillerdeki Görünümü

Kaynakça: *Park, Kim, ve Bae, 2012, s. 899*

Böylece *p*. kantil regresyon doğrusunun parametreleri

$$\hat{\beta}_{(p)} = \min_{\beta} \frac{1}{n}\left(\sum_{i:Y_i \geq X_i'\beta} p|Y_i - X_i'\beta| + \sum_{i:Y_i < X_i'\beta}(1-p)|Y_i - X_i'\beta|\right) \qquad (2.24)$$

denkleminin enküçüklenmesi ile elde edilmiş olur (Chiang ve Li, 2012, s. 38). (2.24) denkleminde p değeri 0.50 olduğunda denklem

$$\hat{\beta}_{(0.50)} = \min_{\beta} \frac{1}{n}\left(\sum_{i:Y_i \geq X_i'\beta} (0.50)|Y_i - X_i'\beta| + \sum_{i:Y_i < X_i'\beta}(1-0.50)|Y_i - X_i'\beta|\right)$$

$$\hat{\beta}_{(0.50)} = \min_{\beta}\left(\sum_i |Y_i - X_i'\beta|\right) \qquad (2.25)$$

halini alır. Bu denklem medyan regresyonun parametre kestirimini verir. Görüldüğü üzere kantil regresyon medyan regresyonun genişletilmiş halidir (Shulze, 2004, s. 24).

(2.24) denklemi kontrol fonksiyonu orijinde türevlenebilir olmadığı için EKK yöntemindeki gibi çözülemez. Bu yüzden denklem doğrusal programlama gösterimi ile ifade edilir. Y_i, sadece pozitif elemanların bir fonksiyonu olarak yazılırsa;

$$Y_i = \sum_{k=1}^{K} x_{ik}\beta_{k,p} + \varepsilon_{i,p} = \sum_{k=1}^{K} x_{ik}\left(\beta_{k,p}^1 - \beta_{k,p}^2\right) + \left(\varepsilon_{i,p} - v_{i,p}\right)$$

$$\beta_{k,p}^1 \geq 0,\ \beta_{k,p}^2 \geq 0,\ k = 1, \ldots, K \text{ ve } \varepsilon_{i,p} \geq 0,\ v_{i,p} \geq 0,\ i = 1, \ldots, n \qquad (2.26)$$

elde edilir. (2.26) denkleminin kısıt olarak yer aldığı doğrusal programlama modeli elde edilmiş olur.

$$\min_{\beta_{k,p}^1,\beta_{k,p}^2,\varepsilon_{i,p},v_{i,p}} \sum_{i=1}^{n} p\varepsilon_{i,p} + (1-p)v_{i,p}$$

Kısıtlayıcı;

$$Y_i = \sum_{k=1}^{K} x_{ik}\beta_{k,p} + \varepsilon_{i,p} = \sum_{k=1}^{K} x_{ik}\left(\beta_{k,p}^1 - \beta_{k,p}^2\right) + \left(\varepsilon_{i,p} - v_{i,p}\right)$$

$$\beta_{k,p}^1, \beta_{k,p}^2, \varepsilon_{i,p}, v_{i,p} \geq 0(\forall i, k) \tag{2.27}$$

Son olarak $A = (X, -X, I, -I)$, $z = (\beta_p^{1'}, \beta_p^{2'}, \varepsilon', v')'$, $c = (0', 0', pl', (1-p)l')'$, $Y = (Y_1, \dots, Y_n)$ dönüşümleri gerçekleştirildiğinde doğrusal programlama modelinin primal problemi

$$\min_{z} c'z$$

$$Az = Y \qquad (z \geq 0) \tag{2.28}$$

$0': Sıfırlardan\ oluşan\ Kx1\ vektörü$

$l: 1'ler\ nx1\ vektörü$

$I: n\ boyutlu\ birim\ matrisi$

olarak elde edilir (Akyol, 2013, s. 44). Buna ek olarak (2.28)'in dual problemi de

$$\max_{w} w'y$$

$$w'A \leq c' \tag{2.29}$$

olarak gösterilir (Buchinsky, 1998, s. 99).

2.4.2. Kantil regresyon modelinin özellikleri

1) Kantil regresyon, verilen bağımsız değişkenler aracılığıyla ve *p.* kantil değeri ile oynayarak, bağımlı değişkenin koşullu dağılımının tümünün modellenmesini sağlamaktadır (Shulze, 2004, s. 27).
2) EKMS yönteminde olduğu gibi, kantil regresyon modelinin amaç fonksiyonu da mutlak sapmaların ağırlıklı toplamına eşittir. Bu sayede katsayı vektörü, bağımlı değişkenin aşırı uç değerlerine karşı hassas olmamasını sağlar.
3) Model, tahmin etmeyi kolaylaştıran doğrusal programlama gösterimine sahiptir.

4) Hata terimlerinin normal dağılmadığı durumlarda EKK yöntemi ile elde edilen doğrusal regresyon analizinin sonuçlarından daha üstün sonuçlar üretir.
5) Farklı kantilleri modelleri tanımlayan regresyon doğrularından elde edilen sonuçlar sayesinde, bağımsız değişkenin bağımlı değişkenin çeşitli kantilleri üzerindeki farklı yön ve şiddetteki etkilerini inceleme imkanı sunar (Buchinsky, 1998, s. 89).
6) Kantil regresyonda p = 0.50 olması halinde medyan regresyon elde edilir (Akyol, 2013, s. 41).

ÜÇÜNCÜ BÖLÜM

TÜRKİYE İMALAT SEKTÖRÜ'NDE FİNANSAL KALDIRAÇ VE İŞLETME BÜYÜKLÜĞÜNÜN KARLILIĞA ETKİSİNİN İNCELENMESİ

Çalışmanın ilk iki bölümünde EKK, EKMS yöntemleri ile kantil regresyon yöntemi hakkında bilgiler verilmiştir. Üçüncü bölümde ise veri kümelerinin modellenmesinde farklı yaklaşımlarda bulunan bu yöntemlerin birbirlerinden eksiklik ve üstünlükleri ortaya koyulmaya çalışılmıştır.

3.1. Araştırmanın Amacı ve Önemi

Araştırmanın amacı, bağımlı değişkenin dağılımlarının modellenmesinde farklı yaklaşımları olan kantil regresyon ve EKK yönteminin birbirlerinden üstünlük ve eksikliklerini ortaya koymaktır. Bu amaçla, işletmelerin borçlanma düzeyi ve işletme büyüklüğünün kârlılığı nasıl etkilediği, söz konusu iki yöntem ile karşılaştırmalı olarak incelenecektir. Çalışmada Kamu Aydınlatma Platformu'nun sitesinden elde edilen ikincil verilerden yararlanılmıştır.

Araştırma sonucunda elde edilen bulgulara göre kantil regresyon ve EKK yöntemlerinin hangi durumlarda daha iyi sonuçlar üreteceği belirlenerek ileride yapılacak çalışmalarda bu yöntemlerden yararlanmak isteyenlere ışık tutması amaçlanmıştır. Buna ek olarak, güncelliğini kaybetmeyen bir tartışma konusu olan işletmelerin sermaye yapısının nasıl olması gerektiğine ilişkin bir katkıda bulunmak da araştırmanın bir diğer amacını oluşturmaktadır.

Yapılan literatür taramasından anlaşıldığına göre, Türkiye'deki işletmelerin kârlılıkları ve borçlanma düzeyleri arasındaki ilişki, kantil regresyon yöntemi ile daha önce araştırılmamıştır. Ayrıca kantil regresyon yöntemi, faktörlerin işletmelerin mevcut kârlılık düzeylerine bağlı olarak, kârlılığı nasıl etkilediği hakkında yorumlar yapılmasına imkan tanımaktadır. Diğer bir ifadeyle, kantil regresyon yardımıyla, finansal kaldıracın düşük kârlı ve yüksek kârlı işletmeler üzerindeki etkileri veri kümesi bir bütün halinde ele alınarak ayrı ayrı yorumlanabilmektedir. Bu sebeplerden dolayı mevcut çalışma, finansal kaldıraç, kârlılık ve sermaye yapısı üzerine yapılan diğer çalışmalardan ayrılmaktadır.

3.2. Araştırmanın Kapsamı ve Sınırlılıkları

Araştırma Borsa İstanbul'da işlem gören imalat sektöründe yer alan 183 işletmenin 3 yıllık verileri ele alınarak incelenmiştir. Söz konusu imalat şirketleri incelenirken, işletmelerin gıda, dokuma, kimya gibi alt sektör ayrımları göz ardı edilmiştir. Buna ek olarak veriler, Kamu Aydınlatma Platformu'nun websitesinden toplandığı için ikincil kaynak veri sınıfındadır. Bu sebeple, işletmelerin finansal tablolarında yer alan verilerin doğru hesaplamalar sonucunda elde edildiği varsayılmaktadır.

3.3. İşletme Büyüklüğü ve Finansal Kaldıraç Kavramları

Bu kısımda işletme büyüklüğü ve finansal kaldıraç kavramları ve bu kavramların ölçülmesinde başvurulan ölçütler tanıtılmıştır.

3.3.1.İşletme büyüklüğü

İşletme büyüklüğünün nasıl tanımlanacağına dair net bir yaklaşım bulunmamaktadır. Literatürde işletme büyüklüğü kavramı farklı yaklaşımlarla, değişik biçimlerde tanımlanmıştır. Karalar (2010, s. 135), işletme büyüklüğünü girişim tarafından bir araya getirilen üretim araçlarının tümünün hacmi biçiminde tanımlamıştır. Öte yandan (Baştürk ve Ödül, 2008, s. 144), bu tanıma ek olarak işletmenin iktisadi faaliyet hacmi ve kapasitesi ifadesini eklemiştir. Bir işletmenin büyüklüğünü ifade edebilmek için niteliksel ve niceliksel ölçütlerden yararlanılır.

Söz konusu niteliksel ölçütler, işletmelerin politikaları, karar alma ve yönetim süreçleri ile ilgili matematiksel ve istatistiksel olarak ölçülemeyen ölçütleri kapsar. İşletmelerin uzman kişilere yönetimde yer verebilme düzeyi, yönetimin paydaşları ile iletişim ve ilişki kurabilme kabiliyeti, dış kaynaklardan sermaye sağlayabilme seviyesi, müşteri ve tedarikçilerle pazarlık yapabilme gücü ve rekabet gücü gibi ölçütler niteliksel ölçütler arasındadır.

Niceliksel ölçütler ise istatistiki olarak ifade edilebilen ve işletmelerin birbirleriyle karşılaştırılmasına olanak sağlayan ölçütlerdir. Çalışan sayısı, sermaye miktarı, maddi duran varlıklar, satış miktarı, üretim kapasitesi, araştırma-geliştirmeye ayırdığı bütçe, üretim merkezinin kapladığı alan gibi ölçütler niceliksel ölçütlerdir. İşletme büyüklüklerinin belirlenmesinde ve sınıflandırılmasında niceliksel ölçütlerden daha fazla yararlanılır.

Tüm bu ölçütler göz önüne alındığında bir işletmenin büyüklüğü konusunda sadece tek bir ölçüte bakılarak karar verilemez. Sağlıklı bir belirlemenin yapılabilmesi için işletmeyi içinde yer aldığı sektörün özelliklerine göre incelemek gerekir. Örneğin, çalışan sayısı ölçütünü ele alarak bir perakende işletmesi ile bir yazılım şirketini birbirleriyle kıyaslamak doğru olmayacaktır. Çünkü yazılım şirketinin ihtiyaç duyduğu çalışan sayısı ile perakende işletmesinin ihtiyaç duyduğu çalışan sayıları birbirlerinden çok farklıdır (Karalar, 2010, s. 137). Buna karşın çalışan sayısı, yıllık satış hasılatı ve bilanço büyüklüğü işletmelerin büyüklüklerine göre sınıflandırılmasında kullanılan başlıca ölçütlerdir. Avrupa Birliği Komisyonu'nun 6 Mayıs 2003 tarihinde yayınladığı, küçük ve orta büyüklükteki işletmelerin sınıflandırılmasında kullandığı tavsiye kararı Tablo 3.1'de verilmiştir (European Commission, 2014):

Tablo 3.1. *Avrupa Birliği Komisyonu KOBİ Sınıflandırması*

İşletme Türü	***Çalışan Sayısı***	***Yıllık Net Satış Hasılatı***	***Yıllık Mali Bilanço Büyüklüğü***
Mikro İşletme	< 10	< 2 milyon Avro	< 2 milyon Avro
Küçük İşletme	< 50	< 10 milyon Avro	< 10 milyon Avro
Orta İşletme	< 250	< 50 milyon Avro	< 43 milyon Avro

Buna karşın Bakanlar Kurulu'nın 18.11.2005 tarihinde Resmi Gazete'de yayınladığı "Küçük ve Orta Büyüklükteki İşletmelerin Tanımı, Nitelikleri ve Sınıflandırılması Hakkında Yönetmelik"e göre Türkiye'deki KOBİ'lerin sınıflandırılması Tablo 3.2.'de yer almaktadır:

Her iki sınıflandırmada da orta büyüklükteki işletmeler için belirlenmiş olandan daha fazla çalışana, satış hasılatına veya bilanço büyüklüğüne sahip işletmeler, büyük işletme olarak adlandırılmaktadır.

Tablo 3.2. Bakanlar Kurulu'nun KOBİ Sınıflandırması

İşletme Türü	***Çalışan Sayısı***	***Yıllık Net Satış Hasılatı***	***Yıllık Mali Bilanço Büyüklüğü***
Mikro İşletme	< 10	< 1 milyon TL	< 1 milyon TL
Küçük İşletme	< 50	< 8 milyon TL	< 8 milyon TL
Orta İşletme	< 250	< 40 milyon TL	< 40 milyon TL

3.3.2. Finansal kaldıraç

İşletmelerin gündelik faaliyetlerini sürdürebilmek için gerekli olan sermayeyi borçlanma veya özsermaye arttırma yoluyla elde etmesine sermaye yapısı politikası adı verilir. Sermaye yapısı politikasının amacı işletmenin sermaye yapısını en iyi biçimde belirleyen borç-özsermaye birleşimini belirlemektir. Bu politika belirlenirken risk ve getiri arasındaki dengenin doğru olarak belirlenmesi gerekir. Bunun sebebi borçlanmanın giderek artmasıyla, ileride alınacak kredilerin riskinin artacak olması ve bir noktadan sonra yüksek maliyetli borç alınmasına yol açacak olmasıdır. Öte yandan, işletme için gerekli olan fonu temin ederken hisse senedi ihraç etmek yerine borçlanmayı tercih etmek, faiz giderlerinden dolayı net kârı azaltsa da özsermaye getirisini arttıracaktır. Diğer bir ifadeyle riskin fazla olması hisse senetlerinin değerini azaltırken, beklenen getirinin artacak olması ise hisse senedinin değerini arttıracaktır. Bu sebeple, risk ve getiri arasında dengenin sağlanması önem arz etmektedir (Brigham ve Houston, 2007, s. 94).

Finansal kaldıraç kavramı, sermaye yapısı kavramının önemli bir bölümünü oluşturmaktadır. Finansal kaldıraç, işletmenin borç kullanarak finansman ihtiyaçlarını karşılamasını tanımlamak için yaygın olarak kullanılan bir kavramdır. Finansal kaldıracın artması ile işletmenin borçlarının özsermayeye göre daha fazla artması, finansal kaldıracın azalması ile de işletme borçlarının özsermayeye göre azalması ifade edilmektedir. Finansal kaldıracın artması, hissedarların beklenen getirilerini arttırırken öte yandan üstlendikleri riskleri de arttırmaktadır. Bu yüzden, finansal kaldıracın ne oranda ve hangi noktaya kadar arttırılması gerektiği önemli bir sorun olarak karşımıza çıkmaktadır. Hiç şüphesiz bu sorunun yanıtı, işletmeden işletmeye ve daha da önemlisi sektörden sektöre farklılık göstermektedir. Bu sebeple, yapılan çalışmaların belirli bir sektörü ele alarak yapılması gerekir (Brigham ve Houston, 2007, s. 597).

İşletmenin finansal yapısı hakkında bilgi edinmek için finansal oranlardan yararlanılması yaygın uygulamadır. Bu amaçla izleyen kesimde finansal oran türleri ve kullanılan başlıca finansal oranlar tanıtılacaktır.

3.4.Finansal Oranlar

İşletmelerin mali durumunu kavrayabilmek için finansal tablolar adı verilen bilanço, gelir tablosu ve nakit akış tablosu incelenir. Bu tablolarda yer alan kalemler işletme hakkında geniş bilgiler sağlamaktadır. İşletme hakkında daha fazla bilgi edinmek için ise kalemler arasındaki göreceli ilişkilerin incelenmesi gereklidir. Kalemlerin arasındaki ilişkileri anlamak için her bir kalemin aldığı değerleri matematiksel olarak ilişkilendiren finansal oranlardan yararlanılır.(Akgüç, 2010), finansal oranı, kalemler arasındaki ilişkilerin basit matematiksel ifadesi olarak tanımlamıştır. Oranlar hesaplanırken, aralarında anlamlı ilişkilerin olduğu düşünülen kalemler göz önüne alınır. Oranların hesaplanması kolay olsa da asıl güçlük hesaplanan oranların yorumlanabilmesindedir. Bunun için yorumlayıcının alanda yeterli bilgiye sahip olması ve belirli bir yargı gücüne ihtiyacı vardır.

Finansal oranların hepsi farklı ihtiyaçlara cevap vermektedir. Tek bir oranla işletmenin tüm analiz ihtiyaçları karşılanamaz. Bu sebeple işletme hakkında farklı konularda bilgiler veren oranlar aşağıdaki gibi sınıflandırılmıştırlar. (Aydın, Başar, ve Coşkun, 2007, s. 91):

- Likidite oranları
- Faaliyet oranları
- Finansal kaldıraç oranları
- Kârlılık oranları
- Piyasa performansını değerlendiren oranlar

3.4.1. Likidite oranları

İşletmenin kısa vadeli borçlarını ödeyebilme gücünü gösteren oranlara likidite oranları denir. Likidite oranları ile işletmenin günlük faaliyetlerine devam edebilmek için gerekli sermayeyi edinebilme gücüne sahip olup olmadığı belirlenir. Likidite oranları hesaplanırken daha çok işletmenin dönen varlıkları ve kısa vadeli yükümlülükler kısımlarındaki kalemler dikkate alınır. Bu kalemlerden başlıcaları dönen varlıklar, kısa vadeli yabancı kaynaklar, stoklar, hazır değerler, menkul kıymetlerdir (Çabuk, 2013, s. 64). Likidite oranlarının başlıcaları likidite oranı, cari oran ve nakit oranıdır.

3.4.2. Faaliyet oranları

İşletmelerin sahip olduğu varlıkların sağladığı gelirlerin nasıl kullanıldığı finansal analizin önemli bir bölümünü oluşturmaktadır. İşletmenin varlıklarını ne kadar etkin kullandığının belirlenmesinde etkinlik oranları olarak da adlandırılan faaliyet oranlarından yararlanılır. Faaliyet oranlarının yüksek olması işletme varlıklarının etkin bir biçimde kullanıldığı anlamına gelir. Alacakların devir hızı, stok devir hızı, sabit varlıkların devir hızı, ortalama alacak tahsil süresi oranı faaliyet oranlarının başlıcalarıdır (Aydın vd., 2007, s. 96).

3.4.3. Finansal kaldıraç oranları

İşletmelerin finansal kaldıraçlarının oran analizi ile ölçülmesi yaygın bir uygulamadır. Bu oranlar; finansal kaldıraç oranları, borçluluk oranları, finansal yapı oranları gibi çeşitli biçimlerde adlandırılmaktadır. Finansal kaldıraç oranlarının başlıcaları borç oranı, kısa vadeli borç oranı, uzun vadeli borç oranı, borç-özsermaye oranıdır.

3.4.4. Kârlılık oranları

Brigham ve Houston (2007, s. 99), kârlılığı işletmenin kararları ve politikalarının net sonucu olarak tanımlamıştır. Kârlılık oranları ile işletmenin varlık, borç ve likidite yönetimi hakkında bilgi elde edilir. Böylece işletmenin aldığı kararların ve faaliyetlerin ne kadar etkili olduğu sonucuna ulaşılır. Kârlılık oranları hesaplanırken gelir tablosundaki kalemlerden ve bilançodan yabancı kaynaklar, özsermaye ve varlıklar kısmındaki kalemlerden yararlanılır. Başlıca kârlılık oranları özsermaye kârlılık oranı, aktif kârlılık oranı, hisse başına kâr oranıdır.

3.5. Veri Kümesi ve Uygulama

Bu çalışmada Borsa İstanbul'da işlem gören ve imalat sanayide yer alan 183 işletmenin 2013, 2014 ve 2015 yıllarına ait finansal tablolarının incelenmesi ile toplam 536 gözlemden oluşan bir veri kümesi elde edilmiştir. Finansal tablolardaki kalemlerin değerleri hesaplanırken, sene başındaki ve sonundaki değerler toplanıp ikiye bölünerek ortalamaları elde edilmiştir.

3.5.1. Değişkenlerin tanımlanması

Çalışmanın amacı, işletmelerin finansal kaldıraç oranları ve işletme büyüklüklerinin, işletme kârlılıkları üzerindeki etkisini kantil regresyon ve EKK yöntemlerini karşılaştırarak incelemektir. Literatürdeki çeşitli finansal oranlar kullanılarak işletme kârlılıkları, büyüklükleri ve finansal kaldıraçları hesaplanabilmektedir. Uygulama veri kümesine dayanılarak, söz konusu işletmelerin her biri için kaldıraç, büyüklük ve kârlılıklarını gösteren finansal oranlar hesaplanacaktır. Bu amaçla literatürde işletmenin kârlılıklarını, büyüklüklerini ve finansal kaldıraçlarını gösteren finansal oranlardan yararlanılarak, örneklemdeki her bir işletmenin ilgili finansal oranları belirlenmiştir. Bu bağlamda regresyon modelinde kullanılacak kârlılık, kaldıraç ve büyüklüğü veren oran ve hesaplamalar Tablo (3.3)'te özetlenmiştir. Bağımsız değişkenler arasında yer alan "Varlıklar", "Satışlar" ve "Duran Varlıklar Oranı" işletme büyüklüğünü göstermektedir. Buna ek olarak "Borç Oranı", "Borç-Özkaynak Oranı", "Uzun Dönem Borç Oranı" ve "Uzun Dönem Borç/Devamlı Sermaye Oranı" işletmenin borçluluk seviyelerini gösteren finansal kaldıraç oranlarıdır.

Daha önce ifade edildiği gibi işletme büyüklüğünün nasıl ölçüleceği bir tartışma konusudur. Bu çalışmada işletme büyüklüğü finansal tablolardan elde edilen veriler aracılığıyla hesaplanmıştır. Büyüklüğün hesaplanmasında kullanılan "Varlıklar" değişkeni, işletmelerin bilançolarında yer alan toplam varlıklar kalemini göstermektedir. "Satışlar" değişkenine ait veri kümesi ise işletmelerin gelir tablosunda yer alan hasılat kaleminden elde edilmiştir. Bu iki değişken, diğer değişkenlerden farklı olarak birer oran değildir ve finansal tablolardan herhangi bir işlem yapılmadan elde edilmiştir. Söz konusu "varlıklar" ve "satışlar" değişkenlerinin veri kümelerinde yer alan değerleri çok yüksek olduğundan, modele baskı kurmamaları ve yalnızca gözlem değerleri arasındaki değişkenliğin göz önüne alınması için doğal logaritmaları alındıktan sonra modele eklenir hale getirilmiştir. Öte yandan maddi duran varlıklar kalemi, işletmelerin sahip olduğu bina, arsa, makine, teçhizat, taşıt vb. varlıklarını kapsamaktadır. İşletme büyüklüğünün hesaplanmasında kullanılan ölçütler arasında yer alan bu varlıkların, modelde yer almasını sağlamak için maddi duran varlıkların toplam varlıklara oranı (Duran varlık oranı) hesaplanmış ve modele eklenmiştir.

Tablo 3.3. *Değişkenlerin Tanımlanması*

Değişken Adı	**Açıklama**
Bağımlı Değişkenler	
Özsermaye Kârlılık Oranı (ROE)	Net Kâr/Özsermaye
Aktif Getiri Oranı (ROA)	Net Kâr/Toplam Varlıklar
Bağımsız Değişkenler	
Varlıklar (VARLIK)	Toplam Varlıkların Doğal Logaritması
Satışlar (SATIS)	Toplam Satışların Doğal Logaritması
Duran Varlıklar Oranı (MADDI)	Maddi Duran Varlıklar/Aktif Toplamı
Borç Oranı (BORC)	Toplam Borçlar/Pasif Toplamı
Borç-Özkaynak Oranı(BORCOZ)	Toplam Borç/Özkaynaklar
Uzun Dönem Borç Oranı (UZUN)	Uzun Vadeli Borçlar/Pasif Toplamı
Uzun Dönem Borç/Devamlı Sermaye Oranı (UOZ)	Uzun Vadeli Borçlar/(Uzun Vadeli Borçlar+Özkaynaklar)

"Borç oranı", işletmelerin kaynaklarının borç ile finanse edildiğini göstermektedir. Bir diğer ismi de kaldıraç oranıdır. "Borç-özsermaye oranı", işletmelerin özsermayeleri ile borçları arasındaki ilişkiyi göstermektedir. Oranın ne kadar olması gerektiği sektörden sektöre ve ekonomiden ekonomiye değişiklik göstermekte iken genel olarak, oranın 1'e eşit olması sağlıklı bir finansal yapıya işaret olarak yorumlanır. "Uzun dönem borç oranı" ise işletmelerin finansmanının ne kadarının uzun dönemli borçlarla sağlandığını gösterir. "Uzun Dönem Borç/Devamlı Sermaye Oranı" kalıcı sermayenin ne kadarının uzun vadeli dış kaynaklardan temin edildiğini gösterir.

"Özsermaye kârlılığı oranı" ve "Aktif getiri oranı", modelde kârlılığı açıklayan oranlar olarak yer almaktadır. "Özsermaye kârlılığı oranı", hissedarların işletmeye yapmış olduğu yatırımların ne kadar verimli olduğunu göstermek için kullanılır.

Oranın yüksek olması, hissedarların yatırımlarının karşılığını aldıkları anlamına gelmektedir. "Aktif getiri oranı" ise işletmenin varlıklarının ne kadar etkin kullanıldığını göstermektedir.

Veri kümesinde yer alan değişkenlere ait tanımlayıcı istatistikler Tablo 3.4'te yer almaktadır:

Tablo 3.4. *Tanımlayıcı İstatistikler*

Değişken	*Gözlem*	*Ortalama*	*Std. Sapma*	*Minimum*	*Medyan*	*Maximum*	*Çarpıklık*	*Basıklık*
ROE	536	0.071	0.746	-2.274	0.067	15.848	17.658	375.866
ROA	536	0.045	0.174	-0.391	0.034	2.945	11.128	171.110
BORC	536	0.494	0.417	0.026	0.456	7.242	9.467	137.747
UZUN	536	0.146	0.136	0.004	0.107	1.443	2.803	17.844
UOZ	536	0.083	3.652	-83.858	0.162	6.063	-22.771	524.542
BORCOZ	536	0.884	7.486	-165.198	0.802	15.189	-20.480	454.952
MADDI	536	0.340	0.171	0.010	0.335	0.949	0.311	-0.269
SATIS	536	18.50	1.856	11.71	18.74	23.39	-0.441	0.238
VARLIK	536	18.77	1.720	13.40	18.97	22.78	-0.460	0.256

Değişkenler arasındaki ilişkilerin yönünü ve şiddetini gösteren korelasyon matrisi Tablo 3.5'te verilmiştir.

Tablo 3.5. *Korelasyon Matrisi*

Değişken	**ROE**	**ROA**	**BORC**	**UZUN**	**BORCOZ**	**MADDI**	**UOZ**	**SATIS**	**VARLIK**
ROE	1.00								
ROA	0.02	1.00							
BORC	-0.03	0.50***	1.00						
UZUN	-0.10**	0.19***	0.69***	1.00					
BORCOZ	0.95***	-0.01	-0.02	0.05	1.00				
MADDI	-0.04	-0.12***	0.02	0.19***	-0.02	1.00			
UOZ	0.00	-0.01	-0.07	-0.01	0.02	0.07	1.00		
SATIS	0.08	0.13***	0.05	0.07	-0.01	-0.01	0.00	1.00	
VARLIK	0.07	0.05	-0.01	0.11**	-0.02	0.09**	0.02	0.92***	1.00

Tablo 3.5 incelendiğinde SATIS ve VARLIK değişkenleri arasında yüksek korelasyonun var olduğu görülmektedir. Bu yüzden dolayı bağımsız değişkenler arasında herhangi bir çoklu bağıntının olup olmadığı varyans büyütme çarpanı ile incelenmiş ve sonuçlar Tablo 3.6.'da verilmiştir.

Tablo 3.6. *Varyans Büyütme Çarpanı (VBÇ) Tablosu*

ROE	**BORC**	**UZUN**	**BORCOZ**	**UOZ**	**MADDI**	**SATIS**	**VARLIK**
VBÇ	2.81	2.29	1.01	1.12	1.01	7.21	7.34
ROA	**BORC**	**UZUN**	**BORCOZ**	**UOZ**	**MADDI**	**SATIS**	**VARLIK**
VBÇ	2.16	2.18	3.00	1.14	1.02	7.49	7.44

Tablo 3.6.'da görüldüğü üzere her iki modelde de VBÇ değerleri 10'un altındadır. Bu yüzden bağımsız değişkenler arasında anlamlı bir çoklu bağıntının olduğu söylenemez.

3.5.2. Model

Araştırmada açıklanmak istenen bağımlı değişken işletme kârlılığıdır. İşletme kârlılığını inceleyebilmek için kârlılık oranlarından "özsermaye kârlılığı" ve "aktif getiri oranları" bağımlı değişkenler olarak belirlenmiştir. Buradan anlaşılacağı üzere uygulamada iki farklı model kurulacaktır.

3.5.2.1. Finansal kaldıraç ve işletme büyüklüğünün özsermaye kârlılığına etkisi

Finansal kaldıraç ve işletme büyüklüğünün kârlılığa etkisinin incelenmesinde özsermaye kârlılığının bağımlı değişken olarak yer aldığı doğrusal regresyon modeli (3.1.)'de verilmiştir:

$$(ROE) = \beta_0 + \beta_1(BORC) + \beta_2(UZUN) + \beta_3(UOZ) + \beta_4(BORCOZ) + \beta_5(MADDI) + \beta_6(SATIS) + \beta_7(VARLIK) \qquad (3.1.)$$

EKK yöntemi ile kestirilen katsayılar ve katsayılar hakkındaki diğer bilgiler Tablo 3.6'da sunulmuştur. Katsayıların p değerleri incelendiğinde UZUN değişkeni haricindeki diğer tüm değişkenlerin katsayılarının istatistiksel olarak anlamlı olduğu görülmektedir. EKK kestirimlerine göre yalnızca istatistiksel olarak anlamlı

[1]değişkenler göz önüne alınırsa; BORC, BORCOZ, MADDI ve VARLIK değişkenlerinin değerlerinin artması özsermaye kârlılığını olumsuz etkilerken, UOZ ve SATIS değişkenlerinin artmasının özsermaye kârlılığını olumlu etkilediği gözlemlenmektedir. Buna göre, işletmeler kendilerinin özsermaye yerine borçla finanse ettiklerinde özsermaye kârlılığı azalmaktadır. Bu durumdan işletmenin finansal kaldıracının artmasının kârlılığı düşürdüğü sonucuna varılır. Ancak UOZ değişkeni incelendiğinde, uzun dönemli borçların devamlı sermaye içindeki oranının artmasının kârlılığı arttırdığı çıkarımı yapılır.

Tablo 3.7. *EKK Kestirimleri (Özsermaye Kârlılığı)*

ROE	**Katsayı**	**Standart Hata**	**t değeri**	**p değeri**
Kesişim	-0.14836	0.1001	-1.473	0.1413
BORC	-0.0864***	0.0315	-2.748	0.0062
UZUN	-0.0420	0.0971	-0.432	0.6661
UOZ	0.0057**	0.0025	2.294	0.0223
BORCOZ	-0.0954***	0.0012	-79.029	0.0000
MADDI	-0.2047***	0.0554	-3.698	0.0002
SATIS	0.0625***	0.0130	4.827	0.0000
VARLIK	-0.0392**	0.0141	-2.773	0.0058
R^2	0.924			
F-Değeri	911.9			

Tablo 3.7. işletme büyüklüğünün kârlılığa etkisi açısından incelendiğinde işletmenin aktif toplamının ve maddi duran varlıkların aktifleri içindeki payının artması özsermaye kârlılığını düşürmektedir. Öte yandan işletmenin toplam satışlarının artması, özsermaye kârlılığını olumlu etkilemektedir.

Kantil regresyon için geliştirilmiş model ise aşağıdaki gibidir:

$$(ROE)^{(p)} = \beta_0 + \beta_1(BORC)^{(p)} + \beta_2(UZUN)^{(p)} + \beta_3(UOZ)^{(p)} + \beta_4(BORCOZ)^{(p)} + \beta_5(MADDI)^{(p)} + \beta_6(SATIS)^{(p)} + \beta_6(VARLIK)^{(p)} \quad (3.2)$$

[1] 1 ***: %99 anlamlılık düzeyi
**: %95 anlamlılık düzeyi
*: %90 anlamlılık düzeyi

Model (3.1) ve (3.2)'nin katsayı kestirimlerinin karşılaştırılmış hali Tablo 3.8'de yer almaktadır. Tablo 3.9'de EKK diye adlandırılan sütunda EKK kestirimleri gösterilmiştir. Diğer sütunlarda ise 5., 10., 25., 50., 75., 90., ve 95. kantiller için yapılan kantil regresyon kestirimleri yer almaktadır. Parantez içinde yer alan değerler standart hataları göstermektedir.

Tablo 3.8'de değişkenlerin değerlerinin düşük kantillerden yüksek kantillere doğru gittikçe farklılaştığı ve kantillerdeki katsayı değerlerinin de EKK kestirimlerinden farklı olduğu görülmektedir. Örneğin; BORC oranı düşük kantillerde negatif katsayıya sahip iken yüksek kantillerde pozitif katsayıya sahiptir. Bu durumda kârlılığı düşük olan işletmelerin borç oranları arttığında kârları daha düşmekte, yüksek kârlı işletmelerin borç oranları arttığında ise kârlılıkları da artmaktadır. UOZ ve UZUN oranlarının da katsayıları kantilden kantile değişse de katsayılar istatistiksel olarak anlamlı değildir. BORCOZ oranının katsayı değerler ise her kantil için birbirine yakın ve negatiftir. Böylece işletmelerin borçları, özsermayelerine göre arttıkça işletmelerin kârlılıkları düşmektedir.

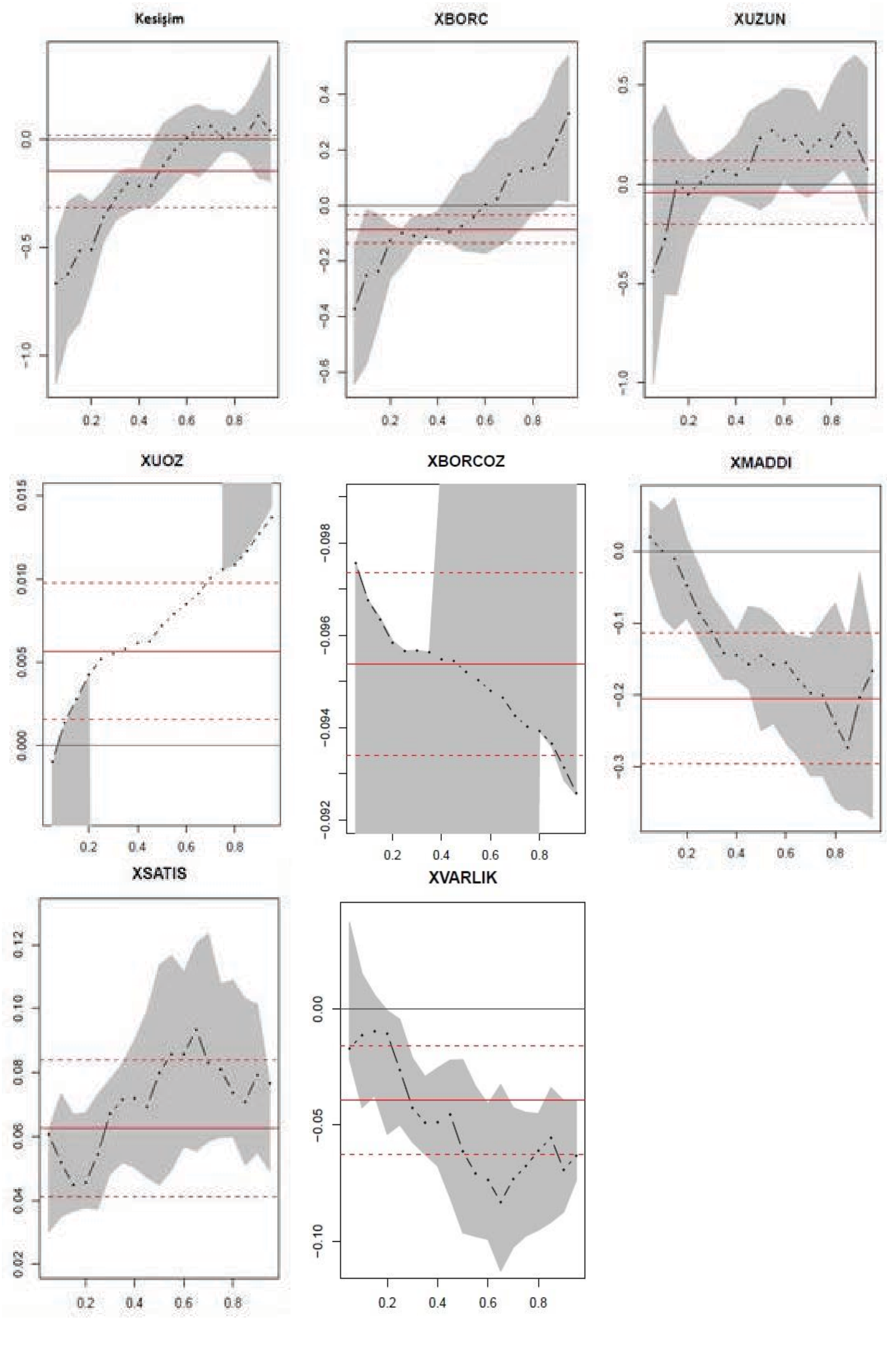
Öte yandan maddi duran varlıklar oranı da düşük kantillerde pozitif katsayıya, yüksek kantillerde ise negatif katsayıya sahiptir. Düşük kârlılığa sahip işletmelerin maddi duran varlıkları arttığında kârlılıkları artmakta, buna karşın yüksek kârlı işletmelerin kârlılıkları maddi duran varlıkları arttığında azalmaktadır. SATIS değişkeni incelendiğinde ise işletmelerin satışları arttığında, işletmenin kâr düzeyine bakılmaksızın kârlarının da arttığı sonucuna varılmaktadır. VARLIK değişkeninin artması ise tüm kantillerde negatif etki yaratmaktadır. Ancak işletmelerin toplam varlıklarının kârlılıkları üzerindeki etkisinin şiddeti yüksek kantillerde düşük kantillere göre daha fazladır.

Şekil 3.1. incelendiğinde kantiller arasındaki farklılıklar daha net olarak görülmektedir. Şekilde yer alan yatay düz kırmızı çizgi EKK katsayılarını, yatay kesikli çizgiler EKK katsayılarının güven aralıklarını göstermektedir. Kesik siyah çizgiler kantil regresyon doğrularını ve gri alanlar da kantil regresyon katsayılarının güven aralıklarını vermektedir. Bu grafikten kantil regresyonun veri kümesini tanımlamak için kullanılmasının faydalı olup olmayacağı belirlenebilir. Eğer kantil regresyon doğruları EKK güven aralıklarının içerisinde kalıyorsa, verinin ortalamaya yakın ve normal dağıldığı çıkarımı yapılır. Bu durumda EKK katsayılarının veriyi ve değişkenler arasındaki ilişkiyi tanımlamada yeterli olduğu sonucuna varılır.

Şekil 3.1 incelendiğinde kantil regresyon doğrularının BORCOZ değişkeni hariç hepsinde EKK güven aralıklarının dışında kaldığı görülmektedir. Buradan bağımsız değişkenlerin ROE'yi çeşitli kantillerde farklı düzeylerde ve yönlerde etkilediği, bu sebeple bağımlı değişken ROE'nin açıklanmasında EKK yönteminin yetersiz kaldığı sonucuna varılmaktadır.

Tablo 3.8. *Kantil Regresyon Kestirimlerinin EKK Kestirimleri ile Karşılaştırılması* (Özsermaye Karlılığı)

	EKK	*5. Kantil*	*10. Kantil*	*25. Kantil*	*50. Kantil*	*75. Kantil*	*90. Kantil*	*95.Kantil*
Kesişim	-0.1484 (0.1001)	-0.6671*** (0.1808)	-0.6242*** (0.1594)	-0.3614*** (0.0947)	-0.1230 (0.0996)	0.0037 (0.0844)	0.1080 (0.1323)	0.0401 (0.1516)
BORC	-0.0864*** (0.0315)	-0.3734*** (0.1041)	-0.2530** (0.1031)	-0.1002** (0.0478)	-0.0753 (0.0771)	0.1235 (0.1165)	0.2342 (0.1611)	0.3309 (0.2384)
UZUN	-0.0420 (0.0971)	-0.4388 (0.5182)	-0.2769 (0.3069)	0.0091 (0.1636)	0.2319 (0.2165)	0.2230 (0.2444)	0.2103 (0.3505)	0.0756 (0.4248)
UOZ	0.0057** (0.0025)	-0.0010 (0.2190)	0.0014 (0.1142)	0.0052 (0.0838)	0.0072 (0.1281)	0.0106 (0.1712)	0.0127 (0.2212)	0.0137 (0.2540)
BORCOZ	-0.0954*** (0.0012)	-0.0976*** (0.0150)	-0.0968*** (0.0157)	-0.0957*** (0.0096)	-0.0952*** (0.0132)	-0.0940*** (0.0139)	-0.0931*** (0.0166)	-0.0926*** (0.0280)
MADDI	-0.2047*** (0.0554)	0.0199 (0.0495)	0.0000 (0.0542)	-0.0862** (0.0408)	-0.1454*** (0.0542)	-0.2006*** (0.0671)	-0.2035** (0.0972)	-0.1662 (0.1073)
SATIS	0.0625*** (0.0130)	0.0606*** (0.0150)	0.0519*** (0.0133)	0.0544*** (0.0143)	0.0799*** (0.0176)	0.0810*** (0.0217)	0.0793*** (0.0166)	0.0766*** (0.0154)
VARLIK	-0.0392** (0.0141)	-0.0170 (0.0202)	-0.0113 (0.0173)	-0.0263 (0.0169)	-0.0613*** (0.0191)	-0.0676*** (0.0226)	-0.0693*** (0.0186)	-0.0632*** (0.0174)
R^2 **Düz. R^2**	0.924 0.923	0.315	0.280	0.247	0.186	0.130	0.101	0.261
F Testi	911.9							



Şekil 3.1. . *Finansal Kaldıraç Ve İşletme Büyüklüğünün Özsermaye Kârlılığına Çeşitli Kantillerdeki Etkisi*

Katsayıların değerlerinin kantiller arasında anlamlı bir farka sahip olup olmadığını anlayabilmek için her bir kantilin simetriği olan kantil ile arasında ANOVA testi yapılır. ANOVA testlerinin sonuçları Tablo 3.9'da verilmiştir. Tablo değerlerine göre katsayı değerlerinin kantiller arasında istatistiksel olarak anlamlı farklılıklar bulunmaktadır.

Tablo 3.9. *Kantiller Arası Fark Analizi (Özsermaye Kârlılığı)*

Kantil	F-değeri(p-değeri)
0.05 ve 0.95	6.869 (0.0000)***
0.10 ve 0.90	10.006 (0.0000)***
0.15 ve 0.85	6.675 (0.0000)***
0.20 ve 0.80	9.293 (0.0000)***
0.25 ve 0.75	6.797 (0.0000)***
0.30 ve 0.70	7.143 (0.0000)***
0.35 ve 0.65	5.685 (0.0000)***
0.40 ve 0.60	4.295 (0.0001)***
0.45 ve 0.55	6.601 (0.0000)***

3.5.2.2. Finansal kaldıraç ve işletme büyüklüğünün aktif getiriye etkisi

Aktif getiri oranının (ROA) bağımlı değişken olarak yer aldığı doğrusal regresyon modeli (3.3)'te verilmiştir:

$$(ROA) = \beta_0 + \beta_1(BORC) + \beta_2(UZUN) + \beta_3(UOZ) + \beta_4(BORCOZ) + \beta_5(MADDI) + \beta_6(SATIS) + \beta_7(VARLIK) \qquad (3.3)$$

Doğrusal regresyon modelin katsayılarının EKK yöntemi ile kestirimleri Tablo 3.10'da verilmiştir.

EKK kestirimleri incelendiğinde MADDI, UZUN ve BORC değişkenlerinin katsayılarının anlamlı olduğu sonucuna ulaşılmıştır. Diğer bağımsız değişkenler ile aktif getiri arasında istatistiksel olarak anlamlı bir ilişki bulunamamıştır. Tablo 3.10'daki katsayılar incelendiğinde borç oranının artması aktif getiriyi arttırmakta iken uzun dönem borç oranı ile maddi duran varlıklarının artması aktif getiriyi olumsuz etkilemektedir.

Tablo 3.10. *EKK Kestirimleri (Aktif Getiri)*

ROE	Katsayı	Standart Hata	t değeri	p değeri
Kesişim	-0.1949	0.0705	-2.765	0.0059
BORC	0.2842***	0.0220	12.907	0.0000
UZUN	-0.3515***	0.0680	-5.168	0.0000
UOZ	0.0016	0.0017	0.902	0.3676
BORCOZ	0.0003	0.0008	0.320	0.7492
MADDI	-0.0806**	0.0387	-2.080	0.0380
SATIS	0.0140	0.0091	1.539	0.1245
VARLIK	-0.0043	0.0099	-0.431	0.6668
R^2	0.3113			
F-Değeri	34.09			

Finansal kaldıraç ve işletme büyüklüğünün aktif getiriye etkisinin incelendiği kantil regresyon modeli ise (3.4.)'te verilmiştir:

$$(ROA)^{(p)} = \beta_0 + \beta_1(BORC)^{(p)} + \beta_2(UZUN)^{(p)} + \beta_3(UOZ)^{(p)} + \beta_4(BORCOZ)^{(p)} + \beta_5(MADDI)^{(p)} + \beta_6(SATIS)^{(p)} + \beta_6(VARLIK)^{(p)} \quad (3.4)$$

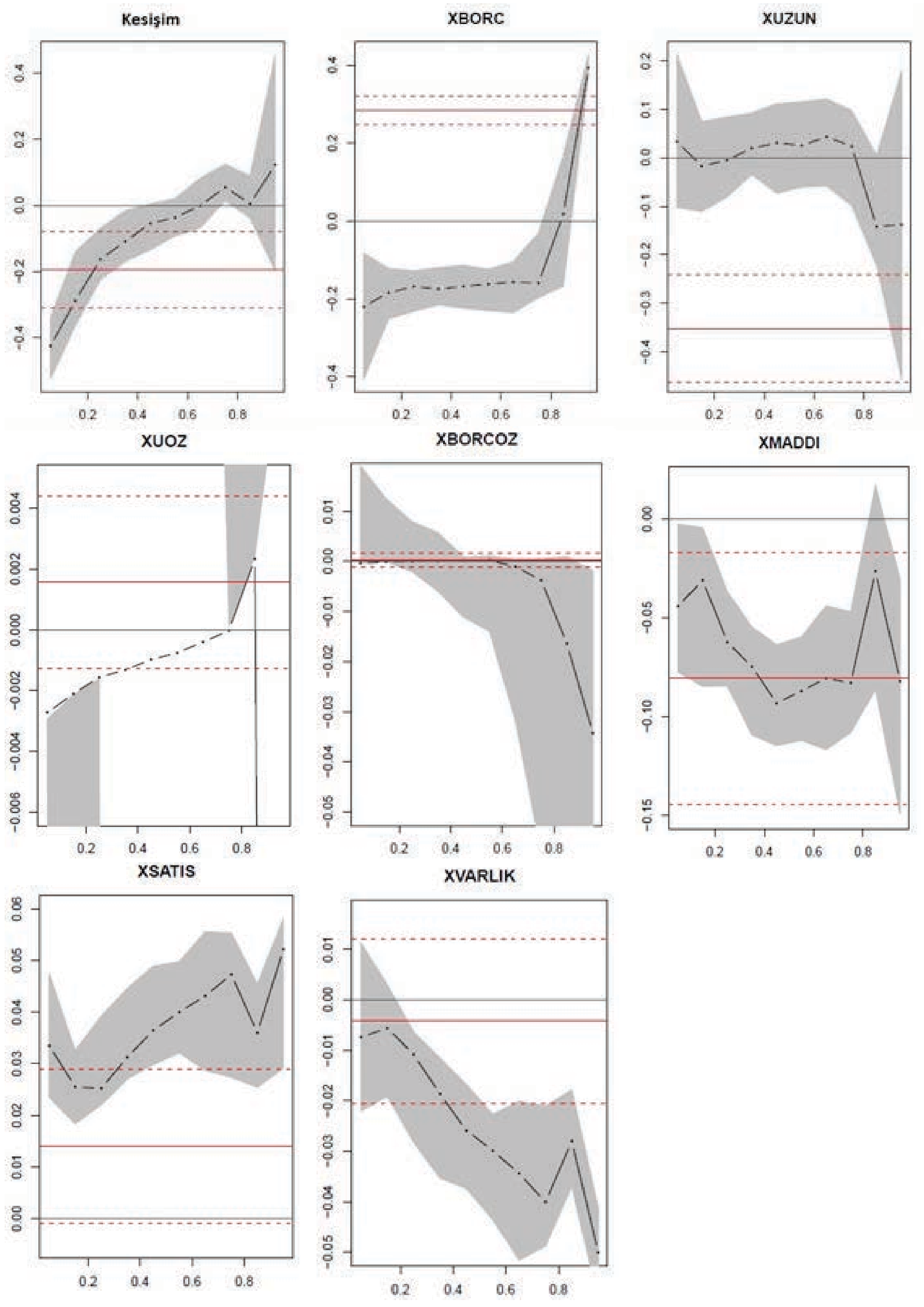
Model (3.3) ve (3.5)'in katsayı kestirimlerinin karşılaştırılmış hali Tablo 3.11'de yer almaktadır. EKK yöntemi ile yapılan analizde sadece UZUN, BORC ve UOZ değişkenlerinin katsayıları istatistiksel olarak anlamlı çıkmıştı. Kantil regresyon analizi incelendiğinde ise SATIS ve VARLIK değişkenlerinin katsayıları da çeşitli kantillerde anlamlı çıkmıştır. Buna karşın UOZ değişkeninin katsayıları hiçbir kantilde anlamlı değildir.

Tablo 3.11'deki değerlere göre UZUN ve UOZ ile aktif getiri arasında anlamlı bir ilişki bulunmamaktadır. Diğer finansal kaldıraç oranları incelendiğinde borç oranının düşük ve orta kantillerde negatif bir etkiye sahip olduğu gözlemlenmektedir. Böylece işletmelerin borç oranı arttığında, aktif getirisinin düşük kârlı işletmeleri olumsuz etkilemekte olduğu sonucuna varılır.

Tablo 3.11. *Kantil Regresyon Kestirimlerinin EKK Kestirimleri İle Karşılaştırılması (Aktif Getiri)*

	EKK	*5. Kantil*	*10. Kantil*	*25. Kantil*	*50. Kantil*	*75. Kantil*	*90. Kantil*	*95. Kantil*
Kesişim	-0.1949*** (0.0705)	-0.4247*** (0.0900)	-0.3445*** (0.0810)	-0.1632*** (0.0499)	-0.0607 (0.0439)	0.0546 (0.0586)	-0.0080 (0.1018)	0.1222 (0.1342)
BORC	0.2842*** (0.0220)	-0.2197*** (0.0776)	-0.1937*** (0.0656)	-0.1679*** (0.0361)	-0.1616*** (0.0595)	-0.1589 (0.1530)	0.1707 (0.2551)	0.3934 (0.2437)
UZUN	-0.3515*** (0.0680)	0.0335 (0.1183)	-0.0061 (0.1131)	-0.0050 (0.0743)	0.0096 (0.1374)	0.0239 (0.1606)	-0.2752 (0.2878)	-0.1371 (0.3784)
UOZ	0.0016 (0.0017)	-0.0027 (0.0485)	-0.0023 (0.0787)	-0.0016 (0.0533)	-0.0009 (0.1018)	-0.0000 (0.1203)	0.0045 (0.1644)	-0.1241 (0.1411)
BORCOZ	0.0003 (0.0008)	-0.0003 (0.0029)	-0.0002 (0.0026)	0.0001 (0.0024)	0.0002 (0.0035)	0.0038 (0.0104)	-0.0242 (0.0159)	-0.0343** (0.0171)
MADDI	-0.0806** (0.0387)	-0.0441 (0.0300)	-0.0526 (0.0339)	-0.0623*** (0.0170)	-0.0939*** (0.0199)	-0.0830** (0.0341)	-0.0285 (0.0523)	-0.0821 (0.0800)
SATIS	0.0140 (0.0091)	0.0335*** (0.0080)	0.0295*** (0.0064)	0.0252*** (0.0062)	0.0381*** (0.0072)	0.0473*** (0.0129)	0.0379*** (0.0141)	0.0523*** (0.0148)
VARLIK	-0.0043 (0.0099)	-0.0074 (0.0103)	-0.0068 (0.0075)	-0.0109 (0.0071)	-0.0268*** (0.0077)	-0.0400*** (0.0130)	-0.0299** (0.0147)	-0.0500*** (0.0171)
R^2	0.311	0.315	0.280	0.247	0.186	0.130	0.101	0.261
Düz. R^2	0.302							
F Testi	34.09 < 2.2e-16							

Öte yandan işletme büyüklüğünü gösteren SATIS ve VARLIK değişkenleri işletmenin aktif getirisini çeşitli kantillerde farklı düzeylerde etkilemektedir. Bu durum Şekil 3.2’den daha net olarak anlaşılabilir. Buna göre SATIS ve VARLIK değişkenleri yüksek kantillerde aktif getiri üzerindeki etkilerini arttırmaktadırlar. SATIS değişkeninin katsayıları tüm kantiller boyunca pozitif iken VARLIK’ın katsayı değerleri tüm kantiller boyunca negatiftir. Ancak etkileri kantilden kantile farklılık göstermektedir.



Şekil 3.2. *Finansal Kaldıraç Ve İşletme Büyüklüğünün Aktif Getiriye Çeşitli Kantillerdeki Etkisi*

Katsayıların değerlerinin kantiller arasında anlamlı bir farka sahip olup olmadığını anlayabilmek için her bir kantilin simetriği olan kantil ile arasında ANOVA testi yapılır. ANOVA testlerinin sonuçları Tablo 3.12'de verilmiştir. Tablo değerlerine göre 0.10-0.90, 0.40-0.60 ve 0.45-0.55 kantilleri haricinde, diğer kantiller arasında katsayı değerlerinin istatistiksel olarak anlamlı farklı değerlere sahip olduğu görülmektedir. Buradan sadece koşullu ortalamanın göz önüne alarak verinin modellenmesinin doğru olmadığı sonucuna varılır.

Tablo 3.12. *Kantiller Arası Fark Analizi (Aktif Getiri)*

Kantil	**F-değeri(p-değeri)**
0.05 ve 0.95	7.256 (0.0000)***
0.10 ve 0.90	1.945 (0.0600)*
0.15 ve 0.85	3.255 (0.0020)***
0.20 ve 0.80	5.952 (0.0000)***
0.25 ve 0.75	2.967 (0.0044)***
0.30 ve 0.70	6.996 (0.0000)***
0.35 ve 0.65	2.428 (0.0185)**
0.40 ve 0.60	1.365 (0.2167)
0.45 ve 0.55	1.308 (0.2431)

SONUÇ

Bu tezde kantil regresyon yöntemini açıklamak ve kantil regresyon yöntemini EKK yöntemi ile karşılaştırarak hangisinin daha doğru sonuçlar ürettiğini belirlemek amaçlanmıştır. Bu doğrultuda kantil regresyon ve EKK yöntemleri açıklanmış, her iki yöntemin asimetrik dağılımların modellenmesindeki başarıları karşılaştırılarak incelenmiştir. Karşılaştırmanın yapıldığı uygulamada finansal kaldıraç ve işletme büyüklüğünün kârlılık üzerindeki etkileri ele alınmıştır.

Doğrusal regresyon analizi, değişkenlerin gerçek hayatta karmaşık olan ilişkilerini, doğrusallık varsayımına dayanarak basit ve anlaşılır bir biçimde modellemeyi hedeflemektedir. İlişkilerin matematiksel olarak modellenmesi sayesinde bağımlı değişkenin değeri tahmin edilebilir, ilişkinin yönü ve şiddeti hakkında da bilgiler elde edilebilmektedir. Doğrusal regresyon analizi değişken sayısına bağlı olarak basit doğrusal regresyon ve çoklu doğrusal regresyon olarak ikiye ayrılmaktadır. Doğrusal regresyon analizinde en sık olarak kullanılan parametre tahmin yöntemleri EKK ve EKMS yöntemleridir.

EKK yöntemi hata terimlerinin karelerini enküçükleyerek parametre tahminlerini gerçekleştirmektedir. Geleneksel anlamdaki regresyon analizinin bir parçası olan EKK yönteminin en büyük dezavantajı, hata terimlerinin normal dağıldığını varsaymasıdır. Bu varsayım altında EKK yöntemi, bağımlı değişkenin koşullu ortalamasını modelleyerek değişkenler arasındaki ilişkiyi modellemeyi hedeflemektedir. Ancak gerçek hayattaki uygulamaların çoğunda hata terimleri normal dağılmamaktadır. Böyle durumlarda EKK yöntemi ilişkiyi doğru olarak tanımlayacak regresyon modelini oluşturamamaktadır.

EKMS yönteminde ise hata terimlerinin mutlak değerleri enküçüklenerek parametre tahmini gerçekleştirilir. EKMS yönteminde diğer bir merkezi eğilim ölçüsü olan medyan dikkate alınır. Medyan, ortalamanın aksine uç değerlere karşı daha az hassas bir merkezi eğilim ölçüsü olduğundan dolayı, dağılımın normal olmadığı durumlarda EKMS yöntemi EKK yöntemine göre daha doğru sonuçlar üretir. Buna karşın EKMS yöntemi de dağılımın tek bir noktasını (medyanı) modellemektedir.

EKMS yöntemini temel alan ve bu yöntemin geliştirilmiş hali olan kantil regresyon yönteminde ise dağılımın çeşitli kantilleri dikkate alınarak, her bir kantil için regresyon doğruları tahmin edilir. Farklı parametrelere sahip olan bu regresyon doğruları sayesinde, dağılımın sadece orta noktası değil tamamı modellenmiş olur. Böylece bağımsız değişkenin, bağımlı değişken üzerindeki etkilerinin çeşitli kantillerde farklı olup olmadığı da incelenebilmektedir. Özellikle asimetrik dağılımlarda, koşullu kantillerin modellenmesi daha doğru sonuçlar üretmektedir.

Çalışmanın uygulama kısmında Borsa İstanbul'da yer alan 183 işletmenin 2013, 2014 ve 2015 yıllarına ait finansal tabloları incelenerek 586 birimden oluşan veri kümesi derlenmiştir. Kârlılığı etkilediği düşünülen işletme büyüklüğünü temsilen üç ve finansal kaldıracı temsilen ise dört tane değişken, modele bağımsız değişkenler olarak eklenmiştir. Bunun yanında, kârlılığı temsil eden özsermaye kârlılığı oranı ile aktif getiri oranı bağımlı değişkenler olarak belirlenmiş ve böylece iki farklı regresyon modeli elde edilmiştir.

Elde edilen modeller öncelikle tanımlayıcı istatistiklerle tanımlanmış ve iki modelde yer alan bağımlı değişkenlerin çarpık ve basık olduğu gözlemlenmiştir. Buradan koşullu ortalamanın uygulama veri kümesini doğru ve yansız olarak tanımlamakta yetersiz kalabileceği sonucuna varılmıştır.

Uygulamada öncelikle özsermaye kârlılığının bağımlı değişken olarak yer aldığı model EKK yöntemi ile incelenmiştir. Özsermaye kârlılığının bağımlı değişken olarak yer aldığı model, anlamlı ve açıklayıcılığının yüksek olduğu sonucuna ulaşılmıştır. EKK yönteminden elde edilen bulgulara göre işletmelerin daha fazla borçlanmasının özsermaye kârlılığını olumsuz etkilediği görülmektedir. Ayrıca işletme büyüklüğünün artması da kârlılığı olumsuz etkilemektedir. Ancak işletmenin satışlarının artması kârlılığı artıran bir faktör olarak yer almaktadır.

Aynı model kantil regresyon analizi ile incelendiğinde ise farklı sonuçlar elde edilmiştir. Buna göre işletmelerin borç oranının artmasında düşük kârlı işletmeler olumsuz etkilenirken, yüksek kârlı işletmeler olumlu etkilenmektedir. Buna ek olarak, uzun dönemli borcun özsermaye kârlılığının herhangi bir kantilinde anlamlı bir etkiye sahip olmadığı sonucuna ulaşılmıştır. Ayrıca maddi duran varlıklarda gerçekleşecek bir artışın yüksek kârlı işletmeleri, diğer işletmelere göre daha olumsuz etkilediği gözlemlenmiştir. Toplam satışlar ve toplam varlıkların çeşitli kantiller üzerindeki etkilerinin çok farklı olmadığı sonucuna ulaşılmıştır. Son olarak özsermaye kârlılığının incelendiği modelde, regresyon parametrelerinin kantiller arasında anlamlı bir farklılık gösterip göstermediği varyans analizi ile incelenmiş ve kantiller arasında istatistiksel olarak anlamlı farklılıkların olduğu sonucuna ulaşılmıştır.

Aktif getirinin bağımlı değişken olarak yer aldığı model EKK ile incelendiğinde ise yalnızca duran varlık oranının, borç oranının ve uzun dönemli borç oranının anlamlı etkilere sahip olduğu sonucuna ulaşılmıştır. Buna göre borç oranı ile aktif getiri arasında aynı yönlü ilişki olduğu sonucuna ulaşılmışken; uzun dönemli borç ve duran varlıklar ile aktif getiri arasında ise ters yönlü bir ilişki mevcuttur.

Aktif getiri modeli kantil regresyon ile analiz edildiğinde ise uzun dönemli borcun kârlılık üzerinde anlamlı bir etkisinin olmadığı sonucuna ulaşılmıştır. EKK yönteminde borç oranı ile aktif getiri arasında aynı yönlü ilişki tespit edilmişken, kantil regresyon analizinde ise borç oranı düşük ve orta kantiller üzer inde negatif bir etkiye sahiptir. Buna ek olarak satışların artması kârlılığı olumlu etkilemekte iken toplam varlıkların artması ise yüksek kârlılığa sahip işletmeleri olumsuz etkilemektedir. Kantiller arasında farklılıkların olup olmadığı varyans analizi ile incelendiğinde ise çoğu kantil arasında anlamlı farklılıkların olduğu sonucuna varılmıştır.

Sonuç olarak finansal kaldıraç ve işletme büyüklüğü ile kârlılık arasında anlamlı bir ilişki olduğu bulgusuna ulaşılmıştır. Analizde başvurulan EKK ve kantil regresyon yöntemleri aynı veri kümesi için farklı sonuçlar üretmiştir. EKK yönteminin koşullu ortalamayı modellemesine ek olarak kantil regresyon yöntemi sayesinde dağılımın farklı kesimleri de modellenebilmiştir. Düşük, orta ve yüksek karlılığa sahip işletmeler üzerinde finansal kaldıraç ve işletme büyüklüğünün etkilerinin aynı olmadığı; EKK ve kantil regresyon yöntemlerinin farklı sonuçlar ürettiği sonucuna ulaşılmıştır. Veri kümesi aykırı değerlere sahip olduğu için kantil regresyon ve EKK yöntemi farklı sonuçlar elde etmiştir. Farklılıklar göz önüne alındığında, kantil regresyonun istatistiksel olarak daha anlamlı ve doğru sonuçlar elde ettiği sonucuna varılmıştır.

KAYNAKÇA

Akgüç, Ö. (2010). *Finansal yönetim* (9. baskı). İstanbul: Avcıol.

Akyol, K. (2013). *Kantil regresyon modeli yardımıyla ülkelerin insani gelişmişlik indeksi üzerinde etkili olan faktörlerin incelenmesi*. Yayımlanmamış Yüksek Lisans Tezi. Erzurum: Atatürk Üniversitesi.

Alpar, R. (2013). *Uygulamalı çok değişkenli istatistiksel yöntemler* (4. baskı). Ankara: Detay.

Altındağ, İ. (2010). *Quantile regresyon ve bir uygulama*. Yayımlanmamış Yüksek Lisans Tezi. Konya: Selçuk Üniversitesi.

Aydın, N., Başar, M., ve Coşkun, M. (2007). *Finansal yönetim* (2. baskı). Eskişehir: Genç Copy Center.

Baştürk, F. H., ve Ödül, Y. (2008). Firma büyüklüğü ile firma büyümesi arasındaki ilişkinin Gibrat yasası çerçevesinde ele alınması. *Mufad Journal*, *39*, 142–154.

Berry, W. D. (1993). *Understanding regression assumptions*. Londra: Sage

Birkes, D., ve Dodge, Y. (1993). *Alternative methods of regression* (1. baskı). New York: Wiley.

Brigham, E. F., ve Houston, J. F. (2007). *Fundamentals of financial management*. *engineering and process economics.* (3. baskı). ABD:South Western.

Buchinsky, M. (1998). Recent advances in quantile regression models: a practical guideline for empirical research. *The Journal of Human Resources*, 33 (1), 88-126

Cade, B. S., ve Noon, B. R. (2003). A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment*, *1*(8), 412–420.

Chatterjee, S., ve Simonoff, J. S. (2013). *Handbook of Regression Analysis*. *Handbook of Regression Analysis*. New Jersey: Wiley

Chiang, T. C., ve Li, J. (2012). Stock returns and risk: evidence from quantile regression analysis. *Journal of Risk and Financial Management*, *5*, 20–28.

Çabuk, A. (2013). Finansal analiz teknikleri. S. Önce (Ed.), *Finansal tablolar analizi* içinde (1. baskı). Eskişehir: Anadolu Üniversitesi Açıköğretim Yayınları

Draper, N. R., ve Smith, H. (1998). Applied Regression Analysis. *Technometrics*, *47*(3), 706.

European Commission. (2014). What is an SME? http://ec.europa.eu/enterprise/policies/sme/facts-figures-analysis/sme-definition/

Freund, R. J., Wilson, W. J., ve Sa, P. (2006). *Regression analysis* (2. baskı). Burlington: Elsevier.

Friendly, M., ve Denis, D. (2005). The early origins and development of the scatterplot. *Journal of the History of the Behavioral Sciences*. 41 (2), 103-130.

Gilchrist, W. G. (2000). *Statistical modeling with quantile functions*. Boca Raton: Chapman & Hall/CRC.

Gujarati, D. N. (2004). *Basic econometrics*. (4. baskı). New York: Mcgraw Hill.

Gürsakal, N. (2013). *Çıkarımsal istatistik* (5. baskı). Bursa: Dora.

Hao, L; Naiman, D. (2007). Quantile regresssion. Sage

Huang, J. C. (2009). *Cumulative distribution networks: Inference, estimation and applications of graphical models for cumulative distribution functions*. Yayımlanmamış Doktora Tezi. Kanada: University of Toronto.

Jones, M. C. (1992). Estimating densities, quantiles, quantile densities and density quantiles. *Annals of the Institute of Statistical Mathematics*, *44*(4), 721–727.

Karalar, R. (2010). *Genel işletme*. İzmir: Meta Basım.

Koenker, R. (2005). *Quantile regression* (1. baskı). New York: Cambridge

University Press.

Koenker, R., ve Bassett, G. (1978). Regression quantile. *Econometrica*, *46*(1), 33–50.

Koenker, R., ve Hallock, K. (2001). Quantile regression. *Journal of Economic Perspectives*, *15*(4), 143–156.

Marill, K. A. (2004). Advanced statistics: linear regression, part ı: simple linear regression. *Academic Emergency Medicine*, 11(1), 87–93.

Montgomery, D. C., Peck, El. A., ve Vining, G. G. (2013). *Doğrusal regresyon analizine giriş*. (M. A. Erar, Ed.) (5. baskı). Ankara: Nobel.

Pardoe, I. (2012). *Applied regression modeling* (2. baskı). Hoboken: Wiley.

Park, J. I., Kim, N., ve Bae, S. J. (2012). A genetic-based iterative quantile regression algorithm for analyzing fatigue curves. *Quality and Reliability Engineering International*, 28(8), 897–909.

Parzen, E. (1979). Nonparametric statistical data modeling. *Journal of the American Statistical Association*, *74*(365), 105–121.

Poole, M. A., ve O'Farrell, P. N. (1971). The assumptions of the linear regression model. *Transactions and Papers, The Institute of British Geographers*, 52, 145–158.

Rawlings, J. O., Pantula, S. G., ve Dickey, D. A. (1998). *Applied regression analysis* (2. baskı). Ann Arbor: Springer.

Serper, Ö. (2014). *Uygulamalı istatistik* (7. baskı). Bursa: Ezgi Kitabevi.

Shulze, N. (2004). *Applied quantile regression: microeconometric, financial and environmental analyses*. Yayımlanmamış Doktora Tezi. Almanya: Universtat Tübingen.

Şıklar, E. (2000). *Regresyon analizine giriş*. Eskişehir: Anadolu Üniversitesi.

Uyar, U., Kangallı Uyar, S., ve Gökçe, A. (2016). Gösterge faiz oranı ve dalgalanmaları ile BİST endeksleri arasındaki ilişkinin eşanlı kantil regresyon ile analizi. *Ege Akademik Bakış*, *16*(4), 587–598.

Vural, A. (2007). *Aykırı değerlerin regresyon modellerine etkileri ve sağlam kestiriciler*. Yayımlanmamış Yüksek Lisans Tezi. İstanbul: Marmara Üniversitesi.

Weisberg, S. (2005). *Applied linear regression*. *Statistics* (Vol. 528).