

**HAYAT DIŐI SİGORTALARDA HASAR
MODELLEMESİNİ ETKİLEYEN FAKTÖRLERİN
KANTİL REGRESYON İLE İNCELENMESİ**

**ANALYSIS OF THE FACTORS AFFECTING LOSS
MODELING IN NON-LIFE INSURANCE USING
QUANTILE REGRESSION**

FATİH UYSAL

Dr. Öğr. Üyesi BAŐAK BULUT KARAGEYİK

Tez DanıŐmanı

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

Aktüerya Bilimleri Anabilim Dalı için Öngördüğü

YÜKSEK LİSANS TEZİ olarak hazırlanmıştır.

2024

ÖZET

HAYAT DIŐI SİGORTALARDA HASAR MODELLEMESİNİ ETKİLEYEN FAKTÖRLERİN KANTİL REGRESYON İLE İNCELENMESİ

Fatih UYSAL

Yüksek Lisans, Aktüerya Bilimleri Bölümü

Tez Danışmanı: Dr. Öğr. Üyesi BAŐAK BULUT KARAGEYİK

Ocak 2024, 76 sayfa

Hayat dıŐı branŐında yer alan kasko sigortaları ülkemizde üretilen sigorta primlerinin önemli bir bölümünü oluŐturmaktadır. Hayat dıŐı branŐta faaliyet gösteren sigorta Őirketleri ise, bu alanda poliçe üretebilmek için rakip Őirketlere kıyasla rekabetçi fiyatlar sunmak zorunda kalmaktadır. Bu durum özellikle kasko sigortalarında fiyatlandırmanın, sigortalı riskini dikkate alınarak yapılmasının önemini arttırmaktadır.

Geleneksel modeller ile yapılan prim hesaplamasında, sigortalanan risklerin birbirine bağımlılığının göz ardı edilmesi, seçilen modelin yeterliliğine ilişkin bilgi eksikliği ve uç hasarlara karşı duyarlı olunmaması gibi dezavantajlar bulunmaktadır. Fakat sigorta poliçesinin fiyatlandırmasında hasarı etkileyen uç değerler dahil bütün değışkenlerin etkisinin yer alması, rekabetçi fiyatlandırma açısından önemlidir. Aynı zamanda sigorta Őirketleri geleneksel yöntemler ile beklenen hasar üzerinden prim hesaplaması yaptıklarında risk yükleme faktörlerini bütün poliçe risk grupları için benzer oranda kullanarak poliçe grubunun riskliliğine göre prim yüklemesi ayrımını

gerçekleştirememektedir. Sigorta şirketleri risk grubuna ilişkin uç değerlerin poliçe fiyatlamasına etkisini de içerecek ve bu riski sigorta primlerine yansıtacak şekilde fiyatlandırma gerçekleştirmelidir.

Söz konusu fiyatlamanın risk grupları için öznel olarak belirlenebilmesi amacıyla, iki aşamalı kantil regresyon modeli önerilmiştir. Bu model, uç hasarlara ilişkin değerlerin riske yansıtılabildiği, sigortalanan risklerin bağımlılık durumunu ve seçilen modelin istatistiksel yeterliliğine ilişkin bilgi eksikliği gibi durumları göz ardı etmeden tahmin gerçekleştirilmesine olanak sağlamaktadır.

Bu çalışmada hayat dışı sigorta branşında hasar modellemesini etkileyen faktörlerin incelenmesinde iki aşamalı kantil regresyonun kullanılması önerilmiştir. Bu kapsamda, lojistik regresyon ve iki aşamalı kantil regresyon modelleri tüm detayları ile açıklanarak literatürdeki çalışmalara değinilmiştir. Çalışmanın uygulama aşamasında iki aşamalı kantil regresyon modeli ile Sigorta Bilgi ve Gözetim Merkezinden (SBM) elde edilen, bir hayat dışı sigorta şirketinin 2022 yılına ilişkin kasko sigortası hasar verileri üzerinden bir uygulama gerçekleştirilmiştir. Analizin ilk aşamasında, öncelikle lojistik regresyon ile hasarın oluşma ve oluşmama olasılıkları hesaplanmıştır. İkinci aşamada ise hasarın gerçekleştiğinin belirlendiği durumlar için kantil regresyon modeli uygulanmıştır. Pozitif bir sabit yükleme faktörü kullanılarak oluşturulan beklenen değer prim hesaplama yöntemi sonuçlarına göre elde edilen prim tutarları ile iki aşamalı kantil regresyon modeli sonucu kantil prim prensibine göre risk yüklemesi yapılarak elde edilen prim tutarları karşılaştırılmış ve sonuçlar yorumlanmıştır.

Elde edilen sonuçlara göre, iki aşamalı kantil regresyon modeli sonucu kantil prim prensibine göre yapılan prim hesabının, sigortalıların risklerini dikkate alması ve yükleme oranlarının risklilik düzeyleri göz önünde bulundurularak yapılması nedeniyle hem daha doğru hem de daha adil bir prim hesabına olanak sağladığı görülmüştür.

Anahtar Kelimeler: Toplam Hasar Modeli, Kantil Regresyon, Lojistik Regresyon, İki Aşamalı Kantil Regresyon, Risk Primi Yükleme, Kasko

ABSTRACT

ANALYSIS OF THE FACTORS AFFECTING LOSS MODELING IN NON-LIFE INSURANCE USING QUANTILE REGRESSION

Fatih UYSAL

Master of Science, Department of Actuarial Sciences

Supervisor: Assist. Prof. Dr. BAŞAK BULUT KARAGEYİK

January 2024, 76 pages

Motor insurance, which is included in the non-life branch, constitutes a significant part of the insurance premiums produced in Turkey. Insurance companies operating in the non-life branch have to offer competitive prices compared to rival companies in order to underwrite policies in this field. This situation increases the importance of pricing, especially in motor insurance, by taking the insured's risk group into account.

Traditional models have disadvantages such as ignoring the interdependence of insured risks, lack of information on the adequacy of the selected model and insensitivity to extreme losses. However, it is important for competitive pricing to include the effects of all variables, including outliers, that affect the loss in the pricing of the insurance policy. At the same time, when insurance companies calculate premiums based on expected

losses with traditional methods, they cannot differentiate premium loading according to the riskiness of the policy group by using risk loading factors at the same rate for all policy risk groups. Insurance companies should realize pricing in a way to include the impact of extreme values related to the risk group on policy pricing and reflect this risk in insurance premiums.

A two-stage quantile regression model is proposed to determine the pricing subjectively for risk groups. This model allows for estimation without ignoring situations where the values of extreme losses can be reflected in the risk, the dependency of the insured risks and the lack of information on the statistical adequacy of the selected model.

In this study, it is proposed to use two-stage quantile regression to examine the factors affecting loss modeling in the non-life insurance branch. In this context, logistic regression and two-stage quantile regression models are explained in detail and studies in the literature are mentioned. In the application phase of the study, an application of the two-stage quantile regression model is performed on the motor insurance claims data of a non-life insurance company for the year 2022 obtained from the Insurance Information and Surveillance Center (SBM). In the first stage of the analysis, the probabilities of occurrence and non-occurrence of the damage were calculated by logistic regression. In the second stage, a quantile regression model was applied for the cases where the loss was determined to have occurred. The premium amounts obtained according to the results of the expected value premium calculation method using a positive constant loading factor are compared with the premium amounts obtained by risk loading according to the quantile premium principle as a result of the two-stage quantile regression model and the results are interpreted.

According to the results obtained from the two-stage quantile regression model, it was observed that the premium calculation based on the quantile premium principle allows for a more accurate and fairer premium calculation since it takes into account the risks of the insured and the loading rates are calculated by taking into account the riskiness levels.

Keywords: Aggregate Claims Model, Quantile Regression, Logistic Regression, Two Stage Quantile Regression, Premium Risk Loading, Motor Insurance

TEŐEKKÜR

Tez alıőmam boyunca bilgi ve deneyimleriyle bana yol gsteren tm hoőgrsyle bu srete beni destekleyen tez danıőmanım sayın Dr. gr. yesi Baőak BULUT KARAGEYİK'e,

Tez savunma jrisinde yer alan; deęerli grő ve katkıları iin Prof. Dr. Sevtap SELUK KESTEL, Prof. Dr. Meral ETİN'e,

Fedakarlıklar gstererek beni ben yapan srekli beni destekleyen eőim Ayőe Zhal UYSAL ve oęlum Ali Efe UYSAL'a,

Sonsuz teőekkrlerimi sunarım.

İÇİNDEKİLER

ÖZET.....	i
ABSTRACT	iii
TEŞEKKÜR	v
İÇİNDEKİLER.....	vi
ŞEKİLLER DİZİNİ.....	viii
ÇİZELGELER DİZİNİ	ix
SİMGELER VE KISALTMALAR.....	x
1. GİRİŞ.....	1
2. LOJİSTİK REGRESYON.....	7
2.1. Lojistik Regresyon	7
2.1.1. Tek Değişkenli Lojistik Regresyon.....	8
2.1.2. İki Değişkenli Lojistik Regresyon.....	8
2.1.2.1. Lojit Model	9
2.1.2.2. Lojistik Model Varsayımları	11
2.1.3. Doğrusal Regresyon ve Lojistik Regresyon Karşılaştırması.....	12
2.1.4. Sıralı Lojistik Regresyon.....	13
2.1.5. İki Değişkenli LM’de Kestirim	14
2.1.5.1. En Çok Olabilirlik Yöntemi (EÇO).....	14
2.1.5.2. İteratif En Küçük Kareler	15
2.1.5.3. Minimum Lojit Ki-Kare Yöntemi	16
2.1.6. Uyum İyiliği Hesaplamaları	16
3. KANTİL REGRESYON VE RİSK HESAPLAMA YÖNTEMLERİ	18
3.1. Kantil.....	18
3.2. Dağılımın Modellenmesi.....	19
3.2.1. Kantil Fonksiyonu	19
3.2.2. Kantil Yoğunluk Fonksiyonu	20
3.2.3. Kantil Dağılıma İlişkin Parametreler	21

3.3. Kantil Regresyon	23
3.3.1. Kantil Regresyon Modeli.....	24
3.3.2. Kantil Regresyon Modelinin Faydaları.....	28
3.3.3. Kantil Regresyon ve Doğrusal Programlama	29
3.3.4. Kantil Regresyona İlişkin Parametre Tahmini.....	31
3.3.5. Belirlilik ve Düzeltilmiş Belirlilik Katsayıları.....	32
3.4. Risk Hesaplama Yöntemleri	33
3.4.1. Prim Tabanlı Risk Hesaplama Yöntemleri	33
3.4.1.1. Net Prim Hesaplama Prensibi	34
3.4.1.2. Beklenen Değer Prim Prensibi	34
3.4.1.3. Standart Sapma Prim Prensibi.....	34
3.4.1.4. Varyans Prim Prensibi.....	35
3.4.1.5. Üstel Prim Prensibi	35
3.4.1.6. Escher Prim Prensibi	35
3.4.2. Sermaye Yapısını Dayanak Alan Risk Hesaplama Yöntemleri.....	36
3.4.2.1. Riske Maruz Değer	36
3.4.2.2. Koşullu Riske Maruz Değer.....	36
3.4.3. Risk Hesaplamalarının Tutarlılık Kriterleri.....	37
3.5. Genelleştirilmiş Doğrusal Modeller.....	38
3.5.1. Gamma Dağılım.....	40
4. İKİ AŞAMALI KANTİL REGRESYON MODELİ TÜRKİYE UYGULAMASI... 41	
4.1. Veri Yapısı	41
4.2. Lojistik ve Kantil Regresyon Uygulaması.....	47
4.2.1. Lojistik Regresyon Uygulaması.....	47
4.2.2. Kantil Regresyon	51
4.3. Risk Yükleme.....	55
4.3.1. Uç Değersiz Analiz.....	61
5. SONUÇ VE ÖNERİ	62
KAYNAKLAR	64
EKLER.....	69
EK 1 – Kantil Regresyon Tahminleri	69
ÖZGEÇMİŞ	76

ŞEKİLLER DİZİNİ

Şekil 2.1. Lojistik Fonksiyon (Sigmoid Eğri) [34].....	10
Şekil 3.1. Kantil Fonksiyonu, $Q(p)$ [39].....	20
Şekil 3.2. Kümülatif Dağılım Fonksiyonu ve Kantil Fonksiyonu [41].....	21
Şekil 3.3. Normal Kantil Fonksiyonu [41].....	23
Şekil 3.4. Çarpık Kantil Fonksiyonu [41].....	23
Şekil 3.5. Doğrusal Regresyon ve Kantil Regresyon Çizgileri [41].....	26
Şekil 3.6. Kontrol Fonksiyonu [21].....	27
Şekil 3.7. RMD ve KRMD Grafikselleştirme [55].....	37
Şekil 4.1. Hasar Tutarları (İl Değişkeni Bazında).....	44
Şekil 4.2. Hasar Tutarlarına İlişkin Grafik (Sigortalı Basamak Grubu).....	46
Şekil 4.3. Sigortalı Risk Grubu ve İl Bazında Hasar Tutarı.....	47
Şekil 4.4. Kantil Regresyon Grafik Karşılaştırmaları.....	53
Şekil 4.5. KPP-BDPP-Risk Prim Yükleme.....	60
Şekil 4.6. Hasar Tutarları (İl Bazında Uç Değersiz).....	61

ÇİZELGELER DİZİNİ

Çizelge 4.1. Değişkenler ve Açıklamaları	41
Çizelge 4.2. Sigortalı Basamak Grubu ve İndirim Oranları	42
Çizelge 4.3. Hasar Frekansı (Toplam)	43
Çizelge 4.4. Hasar Tutarına İlişkin Özet İstatistikler (İl Ayrımında)	44
Çizelge 4.5. Hasar Tutarı Betimleyici İstatistikler (Sigortalı Basamak Grubu)	45
Çizelge 4.6. LR ile Poliçe Risk Grubuna Göre Hasar Oluşma Olasılığı Tahmini.....	48
Çizelge 4.7. LR ile Elde Edilen Olasılıklar (Grupların Sıfır Hasar Olasılıkları).....	49
Çizelge 4.8. $\theta=0,95$ Kantil Değeri ile Revize Hasar Oluşturmama Olasılıkları.....	50
Çizelge 4.9. $\theta=0,95$ kantil Değeri İçin Kantil Regresyon Parametre Tahminleri.....	52
Çizelge 4.10. $\theta=0,95$ Kantil Risk Grupları için Toplam Hasar Tahminleri.....	54
Çizelge 4.11. GDM (Gamma Hasar Tutarı)	56
Çizelge 4.12. Safi Prim, Beklenen Hasar, $\theta=0,95$ Kantil ve Safi Prim Arası Fark	57
Çizelge 4.13. Safi Prim, BDPP, KPP, BDPP ve KPP Kâr Oranları	58

SİMGELER VE KISALTMALAR

Kısaltmalar

BDPP	Beklenen Değer Prim Prensibi
DM	Doğrusal Model
DR	Doğrusal Regresyon
DRM	Doğrusal Regresyon Modeli
EÇO	En Çok Olabilirlik
EKK	En Küçük Kareler
GDM	Genelleştirilmiş Doğrusal Modeller
İAKR	İki Aşamalı Kantil Regresyon
KDF	Kümülatif Dağılım Fonksiyonu
KF	Kantil Fonksiyonu
KRMD	Koşullu Riske Maruz Değer
KPP	Kantil Prim Prensibi
KR	Kantil Regresyon
KRM	Kantil Regresyon Modeli
KYF	Kantil Yoğunluk Fonksiyonu
LR	Lojistik Regresyon
LRM	Lojistik Regresyon Modeli
ODSS	Göreceli Olasılıklar Oranı
OYF	Olasılık Yoğunluk Fonksiyonu
RMD	Riske Maruz Değer
SBM	Sigorta Bilgi ve Gözetim Merkezi
SLR	Sıralı Lojistik Regresyon
TSB	Türkiye Sigorta ve Reasürans Şirketleri Birliği

1. GİRİŞ

Sigortacılıkta verilen teminat karşılığı alınan ücret primdir. Anlaşmanın şartlarına bağlı olarak sözleşmede yer alan riske ilişkin prim, hasar gerçekleşmeden önce sigortalıdan alınır. Sözleşmede üstlenilen riskin maliyeti ise sözleşme süresi içerisinde ya da daha sonra ortaya çıkmaktadır. Sigorta poliçelerinin doğasından kaynaklı olarak önceden tam olarak hesaplanamayan maliyetler, istatistiki tahminler ve rekabet koşulları gibi değişkenlere bağlı olarak belirlenir. Sigorta genel hatlarıyla hayat ve hayat dışı branş şeklinde ikiye ayrılmaktadır. Hayat sigortaları, belirlenen bir prim üzerinden sigortalı ya da lehdara, sigortalının yaşam veya vefatı halinde, sigorta bedelinin ödenmesi prensibine dayanmaktadır. Hayat dışı sigortalar ise sigortalının sahip olduğu varlığın güvenceye alındığı, söz konusu varlığın hasarlanması ya da hiçbir şekilde kullanılmıyor olması durumunda ortaya çıkan kaybı maddi olarak karşılamaya dönük sözleşmelerdir. Ülkemizde hayat dışı sigortacılık alanında düzenlenen poliçe adeti ve prim tutarı, hayat sigortalarına kıyasla oldukça fazladır.

Türkiye Sigorta A.Ş. 2022 Yılı Ekonomik ve Sektörel Gelişmeler raporuna göre Türkiye’de 2022 yılında hayat dışı alanında 204,15 milyar TL, hayat alanında 30,85 milyar TL olmak üzere toplam 235 milyar TL’lik poliçe düzenlenmiştir [1]. Aynı rapora göre 2022 yılında hayat dışı sigorta prim üretiminin 102,1 milyar TL’si kara araçları ve bu araçların sorumluluk sigortaları branşlarında gerçekleşmiştir [1]. Türkiye Sigorta, Reasürans ve Emeklilik Şirketleri Birliği tarafından yayınlanan Kasko Prim Poliçe Adet Raporunda belirtildiği üzere 2023 yılı Ekim itibari ile hayat dışı sigorta branşında faaliyet gösteren 48 şirket bulunmakta olup, 39 şirketin Kara Taşıtları Sigortası (kasko) branşında prim üretimi bulunmaktadır [2]. Prim üretim tutarları ve sigorta şirket sayıları göz önünde bulundurulduğunda, sigorta poliçelerinin fiyatlandırılmasında oldukça rekabetçi bir ortamın oluşmasına neden olduğu gözlenmektedir. Özellikle hem sigortalıya, hem de kullanılan araca ilişkin birçok değişkenin dikkate alındığı kasko sigortalarında, fiyatlandırmanın önemi daha da artmaktadır.

Oldukça rekabetçi bir ortamda poliçe oluşturmaya çalışan sigorta şirketleri poliçeye ilişkin riskin fiyatlamasını en doğru şekilde yapmalıdır. Bunu yapan sigorta şirketi hem mevcut sigortalıların rakip şirketlere geçmesini engelleyebilir, hem de diğer şirketlerden

sigortalıları kendi portföyüne ekleyebilir. Doğru fiyatlanmış sigorta primi, şirketin etkin rekabet edebilmesi, sermaye yapısının korunması, denetleyici ve düzenleyici otorite ile ilişkilerin sorunsuz yürütülmesi gibi birçok konuda şirkete fayda sağlamaktadır. Ayrıca poliçe fiyatlamasının risk seviyelerini içerecek şekilde yapılması şirketin sigorta poliçelerine ilişkin riskin hassas bir şekilde tahmin edilmesine imkan sağlamakta ve tahmin edilenden daha yüksek hasar ile karşılaşılmasını engelleyebilmektedir.

Sigorta şirketlerinin en önem verdikleri konu aslında, üstlenilen riske ilişkin doğru prim hesabını yapabilmektedir. Doğru prim hesabı, hem sigorta şirketinin karşılaşılabileceği beklenen hasarları hem de düşük olmakla birlikte gerçekleşme olasılığı bulunan yüksek hasarları da karşılayabilmelidir.

Sigorta primlerinin doğru şekilde tahmin edilmesi için bilinen birçok aktüeryal yöntem bulunmaktadır. En küçük kareler (EKK) tahmin etme yöntemini temel alan doğrusal ve genelleştirilmiş doğrusal regresyon yöntemleri ise aktüeryal alanlarını da içeren birçok farklı alanda yaygın bir şekilde kullanılmaktadır.

Doğrusal regresyon (DR); bağımlı değişkenin bir ya da daha çok bağımsız değişken tarafından açıklanmaya çalışıldığı istatistiksel bir yöntemdir. Bu yöntem, değişkenler arasındaki ilişkiyi anlamak ve mevcut verilerden geleceğe ilişkin tahminlerde bulunmak için kullanılır. En küçük kareler yöntemi (EKK) ile gözlem ve tahmin değerleri arasındaki ilişkiyi doğrusal veri çizgisine en iyi şekilde uydurmayı amaçlar. Doğrusal regresyon gözlenen ve tahmin edilen veriler arasındaki hata olarak tanımlanan farkların normal dağılıma uyduğu varsayımı ile hareket eder [3].

DR'de yer alan hata terimlerinin normal dağılması varsayımı ya da sabit varyanslılık gibi durumlara ilişkin koşulları aşmak için genelleştirilmiş doğrusal modeller (GDM) önerilmiştir. GDM'ler DR'nin genişletilmiş bir formu olup, DR'nin bulundurduğu kısıtlayıcıları içermemektedir. Özellikle bağımlı değişkenin normal dağılmadığı, sürekli olmadığı veya bağımlı değişkenin olasılık dağılımının belirli bir yapıya uymadığı durumlarda tercih edilmektedir [4]. GDM sürekli, kategorik veya sayım verilerinin analizinde veya DR'nin temel varsayımlarının sağlanmadığı durumlar için daha uygundur. GDM'nin aktüerya alanında kullanıldığı durumlara örnek olarak, yaşam ya da ölüm modellemesi, sağlık sigortalarında çoklu modeller, hayat sigortası ayrılma oranları,

risk sınıflandırması (sigara kullananlarda yüksek ölüm oranının modellenmesi), hayat dışı sigortalarda prim belirlenmesi, hayat dışı branşta rezerv tahmini verilebilir [3]. GDM’de bağımlı değişkenin olasılık dağılımına bağlı olarak farklı bağ (link) fonksiyonları kullanır. Genellikle, binom dağılımı için logit dağılımı tercih edilirken, poisson dağılım için log-link fonksiyonu kullanılır. DR yalnızca doğrusal ilişkileri modellerken, GDM doğrusal olmayan ilişkileri de modelleyebilir [4]. Bahsedilen avantajlarından dolayı, aktüerya bilimleri içerisinde DR’ye göre GDM önemli kolaylıklar sağlamaktadır. DR ilk olarak 1975 yılında bezelyelerin genetik özelliklerine ilişkin çalışmada Galton ve Pearson tarafından kullanılmıştır [5]. 1894 yılında ise Galton tarafından doğrusal regresyon gösterimi ilk defa ifade edilmiştir [6]. 1896 yılında Pearson’ın korelasyon ve regresyon çalışması yayımlanmış ve Pearson teorisinde regresyon eğrisinin nasıl ortaya çıkartıldığı açıklanmıştır [5]. Daha sonra Pearson, çoklu regresyon teorisini geliştirmiş ve ki-kare gibi kavramlara değinerek istatistiğin farklı alanlarında önemli ilerlemeler sağlamıştır [7].

GDM’ye ilişkin ilk çalışmalar ise 1972’de Nelder ile Wedderburn tarafından gerçekleştirilmiştir [8]. McCullagh ve Nelder, GDM’ye ilişkin teorik çalışmaları temel seviyede gerçekleştirmişler [9]. 1991 yılında Renshaw ve 1994’te Renshaw ile Verall aktüerya alanında GDM’yi çalışmalarında uygulamışlardır [10, 11]. GDM’nin hayat ile hayat dışı sigortalar alanında kullanımı ise 1996 yılında Haberman ve Renshaw tarafından gerçekleştirilmiştir [3]. Anderson vd., GDM’nin aktüerya alanındaki kullanımına ilişkin uygulama rehberi isimli yayında GDM’nin teorisi, pratiği ve farklı uygulamalarına değinmiştir [12]. Sigorta veri analizlerinde kullanılan GDM’lere göre Jong ve Heller tarafından sigorta verileri üzerinden uygulamaları da içeren kapsamlı bir kitap yayımlanmıştır [4].

GDM grubunda yer alan ve kategorik değişkenler için kullanılabilen lojistik regresyon modelleri ise sigorta hasar olasılığı hesaplamalarında kullanılmaktadır. Uygulama kolaylığı ve birçok risk faktörünün dikkate alınmasının yanı sıra geleneksel regresyon modelleri (DR ve GDM) gerçek sigorta portföyleri için çelişkili olan birtakım varsayımlara da dayanabilmektedir. Bu modellerde öncelikle koşullu beklentilere ilişkin tahmin gerçekleştirilmektedir. Geleneksel regresyon modelleri katastrofik hasar olasılıkları, sigortalanan risklerin bağımlılık durumu ve seçilen modelin istatistiksel yeterliliğine ilişkin bilgi eksikliği gibi durumları göz ardı edebilmektedir. Ayrıca,

geleneksel modeller, verideki uç hasar değerlerine karşı duyarlı olmaması nedeniyle kısa kuyruklu dağılımları temel almaktadır. Bu da özellikle hasar tutarının kalın kuyruklu bir dağılım ile modellenememesi ve dağılımın kuyruk kısmındaki gerçekleşme olasılığı az olan katastrofik durumların göz ardı edilmesine neden olabilmektedir. Ayrıca, bu modeller, gözlenen değerlerin bağımsız olması varsayımına ve her bir risk değişkenine ilişkin oldukça fazla gözlem sayısına da ihtiyaç duymaktadır. Değinilen problemlere karşı çözüm olarak Heras [13] tarafından sigorta poliçesi fiyatlandırmasında uygulanmak üzere iki aşamalı kantil (quantile) regresyon (İAKR) yöntemi önerilmiştir. İAKR modeli bilinen ve uygulanan Lojistik Regresyon (LR) ve Kantil Regresyon (KR) üzerine kurulmuş belirli bir dağılıma uyma koşulu bulunmayan bir modeldir.

İAKR modeli toplam hasar tutarına ilişkin olasılık dağılımlarını KR kullanarak tahmin eden, ilk aşamada LR kullanarak hasar oluşma/oluşmama ihtimalini kullanan iki aşamalı bir modeldir. Lojistik regresyonun ikili kategorik veriler için kullanıldığı model ise probit model olarak ifade edilmektedir. İAKR yüksek toplam hasarlara ilişkin risk faktörlerinin hangi poliçe gruplarında yer aldığına belirlenmesine olanak sağlayarak bu risk gruplarına ilişkin poliçe fiyatlamasında kullanılmasına imkan sağlayan bir modeldir.

Lojistik regresyon nüfus artışının matematiksel formül ile açıklanmasına ilişkin çalışmalar ile çıkmıştır. İkili ya da çoklu hedef değişkenlere ilişkin açıklayıcı değişkenleri tahmin etmede kullanılmaktadır. Lojistik regresyona ilişkin ilk çalışmalar biyolojik deneylerin analizinde 1944 ve 1951 yıllarında Berkson [14, 15] tarafından yapılmıştır. Bu çalışmalardan sonra lojistik regresyon, hataların normal dağılıma uyduğu kabul edilen probit modele alternatif olarak Finney [16] tarafından geliştirilmiştir. 1970'li yıllardan itibaren probit ve logit modeller yaygın bir şekilde kullanılmaya başlanmıştır. Hosmer [17] lojistik regresyon uygulamalarına çalışmasında yer vermiştir. Hem probit hem logit modeller biyomedikal, ekonomi ve sosyal bilimlerde kullanılmıştır. Prentice [18] biometri alanında logit ve probit modeller üzerine çalışmıştır. Aldrich ve Nelson [19] ise logit ve probit modellerin doğrusal olasılıkları üzerinde çalışma gerçekleştirmiştir.

McCullagh ve Nelder [9] lojistik regresyonun varyans ve ortalamasını binom dağılım aracılığı ile GDM ile ilişkilendirmiştir. Menard [20] lojistik regresyon ile doğrusal regresyon karşılaştırmalarını gerçekleştirmiş ve lojistik regresyona ilişkin özet

istatistiklere çalışmasında yer vermiştir. Heras vd. [13] lojistik regresyonu sigorta poliçesi fiyatlamasına ilişkin çalışmalarında kullanmışlardır.

Kantil regresyon modeli ise Koenker ve Hallock [21], Koenker ve Machado [22] ve Koenker [23] tarafından gerçekleştirilen çalışmalarda tanıtılmıştır. Söz konusu yöntem geleneksel regresyon modelinin birtakım dezavantajlarını ortadan kaldırmaktadır. Kantil regresyon, başta piyasa, kredi ve operasyonel risk olmak üzere finansal alanda yer alan birçok risk tahmininde kullanılmaktadır. Crouhy vd. [24] ve Jorion [25] söz konusu risklerin bütünleşmiş çalışmalarını gerçekleştirmiştir. Kantil regresyon aynı zamanda finansal alanda riske maruz değer tahmin edilmesinde kullanılan popüler bir yöntemdir [26].

KR modeli bazı teknik problemleri çözmek amacıyla birçok sigorta uygulamasında kullanılmaktadır. Portnoy [27] mortalite tablolarına ilişkin oranların geliştirilmesinde KR modelini kullanmıştır. Avustralyada 1946 ve 1992 yılları arasında hazırlanan on adet hayat tablosundan, 1990 ve 1992 yılına ait olan iki tabloda yüzde 25 ve 75 kantil değerlerini kullanarak mortalite tablolarını hazırlamıştır. Pitt [28] kantil regresyon modelini, gelir koruma sigortalarına ilişkin talep sonlandırma oranlarının tahmininde kullanmıştır. Çalışmasında gelir sigortaları verisinde yer alan bağımlılığa ilişkin daha net açıklamalar sağlamak için farklı koşullu kantil eğrilerini kullanmıştır. Abduramanov ve Kudryavtsev [29] net prim tutarına ilişkin oranların tahmininde kantil regresyon modelini kullanmıştır. Daha sonra Kudryavtsev [30] sigorta fiyatlamasında EKK ve GLM gibi geleneksel tahmin modelleri ile kantil regresyon modelini karşılaştırmıştır. Yapılan simülasyon çalışmasında fiyat farklılaşması açısından kantil regresyon modelinin daha iyi sonuçlar verdiğini gözlemlemiştir. Heras vd. [13] çalışmalarında iki aşamalı kantil regresyon modelini sigorta primlerinin belirlenmesinde kullanmışlardır. Ayrıca kantil regresyon modeline dayalı prim prensibi ile hasar verileri üzerine risk yüklemesi yaparak alternatif bir risk yükleme yöntemi geliştirmiştir. Baione ve Biancalana [31] çalışmasında GDM ve KR modellerine ilişkin karşılaştırmayı çalışmalarında uygulamışlardır. Bu çalışma ile İAKR modelinin iki aşamalı GDM modeline göre daha etkin bir şekilde risk marjını belirlediğini göstermişlerdir.

İki aşamalı kantil regresyon modelinin GDM'lere göre önemli avantajları bulunmaktadır. İAKR uç hasarlara ilişkin risk faktörleri ile ilgili bilgi sağlamaktadır. Bu açıdan büyük

hasar oluřturma ihtimali olan polie sahipleri hakkında bilgi elde edilebilmektedir. Bu durum sigorta poliesine iliřkin hasarın ngrlmesi ve fiyatlandırılması aısından olduka nem tařımaktadır.

Bu tezde, sigorta verileri zerinde kantil regresyonu kullanarak, farklı hasar řiddeti seviyelerinde heterojen etkilere iliřkin igrleri ortaya ıkarmak ve risk profillerindeki potansiyel deėiřikliklere ıřık tutmak amalanmaktadır. Bu kapsamda, Trkiye’de faaliyet gsteren bir sigorta řirketine ait verilere iliřkin analiz yapılarak, kantil (quantile) regresyonun GDM modeli ile karřılařtırılması yapılmıřtır. Hasar tutarı zerinden hesaplanan prim tutarlarına beklenen deėer ve kantil prim prensibine gre risk ykleme yapılarak sonular deėerlendirilmiřtir. Teze iliřkin beř blmde ařaėıda yer alan alıřmalar gerekleřtirilecektir. Tez alıřmasının giriř blmn oluřturan ilk blmnde DR, GDM ve İAKR’yi oluřturan lojistik ve kantil regresyon ile bu modellerin geliřtirilmesinde kullanılan yntemlere iliřkin literatr taramasına yer verilmiřtir. Tezin ikinci kısmında, LR tekniėine iliřkin bilgi ve aıklamalar detaylı olarak aıklanmıřtır. Bu blmde, tek deėiřkenli ve ok deėiřkenli LR’ye deėinilmiř, lojistik model ve varsayımları incelenmiř, lojistik ve doėrusal regresyon karřılařtırması yapılarak uyum iyiliėi hesaplamaları hakkında bilgi verilmiřtir. Tezin nc kısmında ise, KR ve modele iliřkin matematiksel detaylar aıklanmıřtır. Ayrıca bu kısımda risk ykleme faktrlerinden olan prim tabanlı risk hesaplama yntemleri ile sermaye tabanlı risk hesaplama yntemlerinden bahsedilmiřtir. Tezin drdnc kısımda ise Trkiye’de faaliyette bulunan bir sigorta řirketinin kasko verileri zerinden sayısal uygulama gerekleřtirilmiřtir. Analizde ncelikle, LR modeli ile sigortalının bulunduėu iki kategorili risk grubu belirlenerek polienin hasar oluřturma ve oluřturmama olasılıėı hesaplanmıřtır. Sonraki ařamada ise, hasar oluřmama olasılıkları kullanılarak hasar tutarına iliřkin kantil deėerleri elde edilmiřtir. Pozitif risk ykleme ve beklenen deėer prim hesaplama yntemine gre safi prim ve kantil yntemine gre elde edilen prim deėerleri karřılařtırılmıřtır. Tezin beřinci ve son kısımda ise alıřmaya iliřkin bilgiler zetlenerek sonular yorumlanmıř ve alıřmanın geliřtirilebilecek alanları aıklanmıřtır.

2. LOJİSTİK REGRESYON

İki aşamalı kantil regresyon modeli ile poliçenin hasar meydana getirmeme olasılığı olan p_i oranı, seçilmiş olan poliçe grubu bazında hesaplanmaktadır. Bu aşamada öncelikle lojistik regresyon modeli ile poliçenin hasar oluşturmama olasılığı belirlenir. Hasar oluşmuş olan poliçe ve oluşmamış olan poliçe şeklinde iki kategorik veri bağımlı değişken olarak belirlenir. Bağımsız değişkenler ile ilişkilendirilen risk gruplarının hasar oluşturma ve oluşturmama olasılıkları hesaplanır. Bu bölümde lojistik regresyon kavramı detaylı olarak açıklanarak, dağılımın modellenmesi hakkında bilgi verilecektir.

2.1. Lojistik Regresyon

Lojistik regresyon modeli (LRM) çeşitli bağımsız değişken (x) ve y kategorik (0,1) bağımlı değişken arasındaki ilişkiyi belirlemek amaçlı kullanılan matematiksel modellerdir [32]. Lojistik regresyon modeli bağımlı değişkenin birden fazla kategoride bulunduğu durumda en iyi modelin kurularak risk faktörlerinin belirlenmesinde kullanılan bir metottür. Doğrusal regresyon (DR) ile kıyaslandığında aradaki en temel fark bağımlı değişkenin kategorik olmasıdır. DR ile bağımlı değişkenin gözlemlenen ve tahmin edilen değerleri arasındaki farkın normal dağılması varsayımı mevcut iken LRM’de bu koşul aranmamaktadır. Regresyon tarzı çalışma ile oldukça avantajlı olan bu metot kategorik verilerin analizinde etkin bir şekilde kullanılmaktadır. DR ile bağımlı değişken tahmin edilirken; LRM’de bu değişkenin alabileceği değerlerden bir tanesinin ortaya çıkma ihtimali tahmin edilmektedir. Lojistik regresyon modelinin tercih edilme sebepleri aşağıda özetlenmektedir [33].

- LR, bağımsız değişkenin sayısal veya kesikli olmasına mani değildir.
- LM’nin parametreleri kolay değerlendirilebilir olup matematiksel açıdan kullanımı rahat ve basit fonksiyonlar üretir.
- Lojistik modele bağlı analizlerin kullanımına yönelik çok sayıda istatistik programı bulunmaktadır.
- Bağımsız değişkenlerin fonksiyonlarının dağılımına ilişkin bir sınırlama bulunmamaktadır.
- LR’ye bağlı elde edilen tüm olasılıklar, 0 ile 1 aralığında değer alır. Bağımsız ile bağımlı değişken arasında olan ilişki doğrusal olmayabilir; doğrusallık yanında üstel ya da polinomsal ilişki problem yaratmaz. Değişkenler arasında

logit ilişki ön kabulü ile doğrusal olmayan modeller yaratılabilir. LR, doğrusal olmayan durumu değiştirmeden, ln dönüşüm ile doğrusal hale döndürebilir.

2.1.1. Tek Değişkenli Lojistik Regresyon

Tek değişkenli lojistik regresyon modeli, bağımlı değişkeni etkileyen yalnızca bir bağımsız değişkenin incelendiği model olup genel olarak doğrusal regresyon modeli;

$$Y = \beta X + \epsilon$$

eşitliği ile gösterilmektedir.

Bağımlı değişken 0 ve 1 değerleri ile ilişkilendirildiği durumlarda koşullu ortalama 0 ile 1 arasında değişim göstermektedir. Lojistik regresyonun gösterimini kolaylaştırmak için x değerinin bilindiği durumda Y 'ye ilişkin koşullu ortalaması $\pi(x) = E(Y|X)$ olup, ifadenin açık şekli ise;

$$\pi(x) = \frac{e^{\epsilon + \beta X}}{1 + e^{\epsilon + \beta X}} \quad (2.1)$$

olarak gösterilmektedir.

2.1.2. İki Değişkenli Lojistik Regresyon

Bağımlı değişkeninin 0 veya 1 ikili değerleri aldığı durumlarda $P(Y_i = 1)$, i değerinin 1 değerini alma olasılığını göstermektedir.

$$E\left(\frac{Y_i}{X_i}\right) = P(Y_i = 1) = \sum_{k=0}^p \beta_k X_{ik} \quad (2.2)$$

Eşitlik (2.2)'deki doğrusal olasılık modelinde eşitliğin solu sıfır ile bir arasında yer alan değerlerden oluşmakta, söz konusu değerler sonsuza giden açıklayıcı değişkenlerle ilişkilendirilmesi durumunda eşitlik durumların hepsinde sağlanabilmektedir. Böyle olduğunda tepki değişkeni farklı dönüşümlerle $(-\infty)$ ile $(+\infty)$ aralığında tanımlanabilir. Bu özellikler lojistik ve probit dönüşümler ile sağlanmaktadır.

2.1.2.1. Lojit Model

Doğrusal regresyon modelinde bağımlı değişkenin tanım aralığı 0 ile 1 arasında olması gerekirken doğrusal fonksiyon bu değerler arasında sınırlandırılmamıştır. Bu sorunun çözülmesi için bağımlı ve bağımsız değişkenin tanım aralığını örtüştürmek gerekmektedir. Bu durum ise DM'ye alternatif olan ve yaygın olarak kullanılan lojit model ile mümkün olabilmektedir.

Olasılık değerleri üzerinden $P/(1 - P)$ dönüşümü tepki değişkeninin sınırlarını (0) ile (∞) arasında tanımlı yapacaktır. Sınırlar bu oranın doğal logaritmasını almak suretiyle $(-\infty)$ ile $(+\infty)$ yapılabilir. Bu durumda $P_i = P(Y_i = 1)$ ve $i = 1, 2, \dots, n$ için lojistik model aşağıdaki eşitlikte tanımlanmaktadır.

$$E\left(\frac{Y_i}{X_i}\right) = P(Y_i = 1) = L_i = Ln\left(\frac{P_i}{1-P_i}\right) = \sum_{k=0}^p \beta_k X_{ik} . \quad (2.3)$$

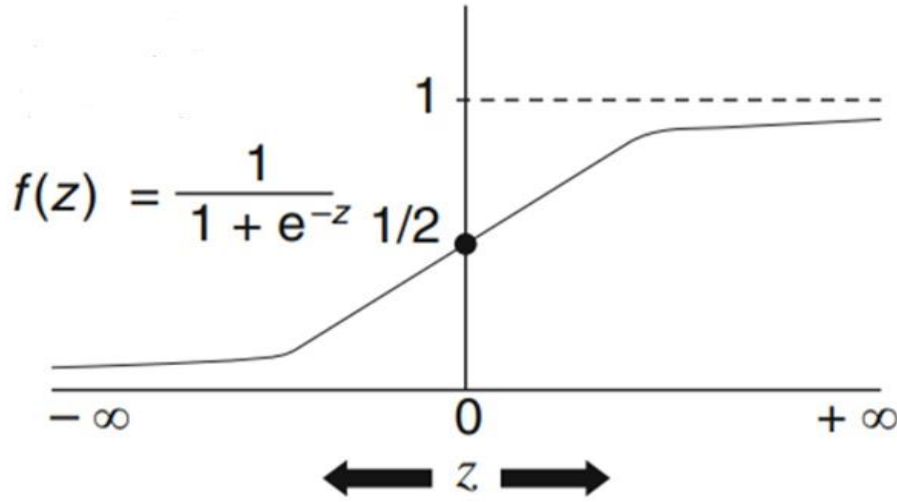
Ayrıca P_i olasılık değeri ise;

$$P_i = \frac{\exp(\sum_{k=0}^p \beta_k X_{ik})}{1 + \exp(\sum_{k=0}^p \beta_k X_{ik})}$$

olarak tanımlanır ve lojistik fonksiyon adını alır [32].

Tepki değişkeni kesikli olduğu için tepki değişkeni ile bu değişkeni açıklayan bağımsız değişkenler arasındaki ilişki açık bir şekilde gösterilemez. Bu sebepten tepki değişkeni yerine P_i olasılık değerine karşı grafik çizimi gerçekleştirilir. Bu şekil “lojistik fonksiyon” olarak adlandırılır ve s (Sigmoid) şeklinde bir eğridir. 0 ile 1 arasında olasılık değerleri alabilen bu eğri sürekli bir eğridir.

Lojistik regresyonun analizini önemli hale getiren iki ana sebep bulunmaktadır. Öncelikle lojistik fonksiyon, tahminlerin 0 ve 1 arasında olduğu durumlarda söz konusu durumu sağlamaktadır. Ayrıca, Sigmoid eğrisine (SE) bağlı olarak, bağımsız değişken $-\infty$ 'a giderken bağımlı değişken 0'a yaklaşır, bağımsız değişken $+\infty$ 'a giderken ise, bağımlı değişken 1'e yaklaşmaktadır [32]. Lojistik fonksiyon (Sigmoid Eğri) Şekil 2.1'de verilmektedir.



Şekil 2.1. Lojistik Fonksiyon (Sigmoid Eğri) [34]

Lojistik regresyon dağılımında tepki değişkenlerinin, $(0,1)$ şeklinde ikili değerler almasından dolayı hata terimlerinin ortalama değeri 0 ve varyansı $p(1 - p)$ olarak bulunur. Hata teriminin ise $E(Y|X)$ ortalama ve sabit bir varyans ile normal dağılıma uyduğu söylenebilir. X değişkeni bilindiğinde $Y = f(x) + \epsilon$ şeklinde ifade edilir. ϵ ise mümkün olan iki değerden birisini alabilir. Y 'nin 1 olduğu durumda $\epsilon = 1 - f(x)$, y 'nin 0 olduğu durumda ise $1 - f(x)$ olasılıkla $\epsilon = -f(x)$ değerini alır. Bu şekilde hata terimi olan ϵ , ortalaması 0 olan ve varyansı $[1 - f(x)] f(x)$ olan bir dağılımı gösterir.

Lojistik regresyonda, bir olayın gerçekleşme olasılığını tanımlamanın farklı yollarından biri olan Göreceli Olasılıklar Oranı (ODDS) kullanılır. Olayın ODDS'si olayın oluşma sayısının oluşmama sayısına oranı şeklinde elde edilir. ODDS, $O=P/(1-P)$ olarak elde edilmekte olup, bu durumun tersi $P=O/(1+O)$ olasılık durumunu ifade etmektedir.

ODDS değerinin 1 'den küçük olması durumunda, olasılık $0,5$ 'ten küçüktür. ODSS oranları hassas olup ikili kategorik değişken arasındaki ilişkinin ölçülmesinde ve lojistik regresyon analizinde kullanılırlar [34].

2.1.2.2. Lojistik Model Varsayımları

LM ile olayın oluşmasının, oluşmaması olasılığına oranının logaritması, bağımsız değişkenlere lineer şekilde bağlanır [34]. Rasgele değişken sürekli ya da kesikli değişkenlerden oluşursa diskriminantın bağımlı olmayan değişkenlerinin normal dağılımlı olması kabulü bozulabilir. Bu durumda herhangi bir kısıtı olmamakla birlikte bazı varsayımları içinde barındıran LM kullanılabilir. Lojistik model varsayımları şu şekilde özetlenmektedir:

- $Y \in (0,1)$,
- $P\left(y_i = \frac{1}{x_i}\right) = P_i$,
- $y_1 y_2 y_3 \dots y_n$ istatistiksel açıdan bağımsızdır,
- Açıklayıcı değişkenler x_k bağımsızdır [35].

LM'nin yanıt değişkenini oluşturan lojit dönüşümün $Lojit(P) = Ln\left(\frac{P}{1-P}\right)$ özellikleri ise şu şekilde belirtilmektedir [19].

- P arttıkça lojit(p) de artmaktadır.
- P, sıfır ile bir arasında ise Lojit(P), reel sayı şeklinde ifade edilebilir.
- $P < 0,50$ olduğunda $Lojit(P) < 0$ diğer durumda tersi olur [35].

Lojistik modelde, açıklayıcı değişkenlerin tamamının sürekli, tamamının kesikli ya da kesikli ve sürekli olduğu durumlar için modelin farklı tanımları mevcuttur.

Açıklayıcı değişkenlerin hepsinin kesikli olması durumunda, eşitlik,

$$E\left(\frac{Y_i}{X_i}\right) = P(Y_i = 1) = L_i = Ln\left(\frac{P_i}{1-P_i}\right) = \sum_{k=0}^p \beta_k X_{ik}$$

şeklinde tanımlanmaktadır.

Açıklayıcı değişkenlerin hepsinin sürekli olduğu durumda ise Eşitlik (2.4)'te gösterildiği şekilde tanımlama yapılmaktadır.

$$Ln\left(\frac{P(X_1, \dots, X_p)}{1-P(X_1, \dots, X_p)}\right) = \beta_0 + \sum_{k=1}^p \beta_k X_{ik}. \quad (2.4)$$

Açıklayıcı değişkenlerin sürekli ve kesikli olarak karışık olduğu durumda, LM modeli Eşitlik (2.5) ile tanımlanmaktadır.

$$\text{Ln} \left(\frac{P f_1(X_1, \dots, X_p)}{(1-P) f_0(X_1, \dots, X_p)} \right) = \beta_0 + \sum_{k=1}^p \beta_k X_{ik}. \quad (2.5)$$

Eşitlik (2.5)'de, β katsayıları diskriminant fonksiyonunun katsayılarını göstermekte olup, yapılan gözlemleri f_0 ve f_1 fonksiyonlarına karşılık gelerek ayırır.

$$P(x) = \int_0^{R(x)} Z^{n_1-1} \cdot (1 - Z)^{n_2-1} \cdot dZ / B(n_1, n_2) \quad (2.6)$$

$n_1, n_2 > 0$ olmak üzere, $B(n_1, n_2)$ β fonksiyonu ve $\beta = (\beta_0, \dots, \beta_p)$ olup $R(x)$ ise matris formunda;

$$R(x) = \exp(\beta' X) / (1 + \exp(\beta' X)) \quad (2.7)$$

şeklinde tanımlanır ve n_1 ve n_2 'nin bire eşit olduğu durum lojistik modele karşılık gelir.

Açıklayıcı değişkenler bağımlı değişkene doğrusal bir yapı ile bağlıdır. İlk olarak bağımlı değişkeni belirli sayıdaki denemelerde başarı sayısına ilişkin bilinen binom sayımlarından oluşmaktadır. İkinci olarak başarı sayısı doğrusal bir yapı ile açıklayıcı değişkenlere bağlıdır. Bağ fonksiyon ise, açıklayıcı değişkenlere ilişkin doğrusal yapıyı binom sayımlarının beklenen değerlerine bağlayan bir fonksiyondur. Fonksiyon LM için lojit dönüşümü şeklinde link fonksiyonu olarak ifade edilir. Gözlemlere ilişkin varyans, ortalama ile fonksiyonel olarak ilişkilidir [9].

2.1.3. Doğrusal Regresyon ve Lojistik Regresyon Karşılaştırması

Doğrusal regresyonda iki değişken arasındaki ilişki $Y = \alpha + \beta X$ olarak ifade edilmektedir. Bu denklemde α kesişim noktası ya da $X = 0$ iken Y 'nin değeri olarak ifade edilir. β ise eğim ya da X 'in bir birim artışında Y değerinde meydana gelen değişimi gösteren katsayıdır. α ve β 'ya ilişkin değerlerin tahmininde kullanılan metot ise EKK yöntemidir. Bu yöntem yukarıda yer alan tüm değişkenlerinin yanı sıra hata terimi olan ϵ_j 'yi de tahmin etmektedir. Hata terimi, Y 'ye ilişkin tahmin ile j durumunda Y 'nin değeri arasında gerçekleşen farktır. Bağımlı değişken olan Y 'ye ilişkin tahmin edilmiş değerler Y 'nin bazı olası değerleri için oldukça iyi tahminlerdir [20]. Ancak, $0 \leq E(Y|x) \leq 1$ eşitsizliğinin sağlanabilmesi için bahsedilen avantajlarından dolayı LR kullanılır ve bu dönüşüme lojit dönüşüm adı verilir.

Bağımlı değişkenin yalnızca iki sonuç gösterdiği durumda bu değerler 0 ya da 1'i gösterir. Değişken ikili değer aldığı için, tahmin edilen değerlerin ortalaması 0 ile 1 arasında yer almaktadır. Bağımlı değişkenin ikili sonuçlar aldığı durumda EKK kullanılması halinde bağımlı değişkene ilişkin tahminler 1'i aşabilir ya da 0'ın altında kalabilir. Bağımlı değişken sürekli değilse denklemin koşullu ortalaması 0 ile 1 değerleri arasında yer alır. Dolayısı ile DR'de yer alan hata terimlerinin normal dağıldığı varsayımı yerine LR'de hata terimlerinin binom dağıldığı varsayımı yer alır. Lojistik regresyon analizlerinde ise ilk amaç bağımlı değişken olan Y'nin değerini direkt olarak tahmin etmek yerine olayın oluşup oluşmayacağına ilişkin olasılığın tahmin edilmesidir. $Y = 1$, olayın gerçekleştiğini göstermekte olup, $P(Y = 1)$ olayın gerçekleşme ihtimalinin olasılığını göstermektedir. Tahmine ilişkin değerlerin 1'den fazla ya da 0'dan düşük olmasına ilişkin problem, yukarıda değinilen olasılık oranları ile aşılabilmektedir. $Y = 1$ in, $Y \neq 1$ 'e oranı olasılık değeri olarak hesaplanmaktadır. Eğer oran 1,85 ise, $Y = 1$ durumuna düşme ihtimali $Y = 0$ 'a düşme ihtimalinin 1,85 katıdır şeklinde yorumlanmaktadır. Bu oran 0'dan düşük değerler alamamakla birlikte, 1'den büyük değerler alabilir. Bu yüzden EKK olay ya da kazanın hangi olasılıkla gerçekleşeceğine ilişkin tahminlerde kullanılamaz.

2.1.4. Sıralı Lojistik Regresyon

Tepki değişkeni ikiden daha fazla sıralı seviyeye sahip olabilir. Amaç belki bu kategorilerden bir tanesinde olma ihtimalini belirlemek olabilir. Üç ya da daha fazla sıralı kategori bulunması durumunda Sıralı Lojistik Regresyon (SLR) kullanılmaktadır [36]. İkili lojistik regresyon modelinde 0 ve 1'den oluşan iki bağımlı değişken bulunmaktadır. Sıralı tepki değişkeninde ise üç ya da daha fazla kategori bulunabilir. Sıralı lojistik modele ilişkin formlar aşağıdaki eşitliklerde gösterilmektedir.

$$\log\left(\frac{p_1}{1-p_1}\right) = \alpha_k + \beta'X,$$

$$\log\left(\frac{p_1 + p_2}{1-p_1-p_2}\right) = \alpha_k + \beta'X,$$

ve

$$\log\left(\frac{p_1+p_2+\dots+p_k}{1-p_1-p_2-\dots-p_k}\right) = \alpha_k + \beta'X \quad (2.8)$$

SLR iki ya da daha fazla regresyon eğrisini eş zamanlı kullanan regresyon tekniğidir. Yukarıda gösterilen eşitlik $Y=1$ grubuna karşı $Y=2$ 'den k 'ya kadar olan gruba oranını göstermektedir. Seride yer alan denklem sayısı, sıralı kategorilerden 1 eksilti olarak bulunur. Eğer Y , 4 sıralı seviyeye sahipse modellenen eşitlik sayısı 3 olacaktır. Her eşitlik kendi lojistik modelini temsil etmektedir. 1'den k 'ya kadar olan olasılıkların toplamı 1'e eşittir. SLR modelin kümülatif olasılığını ifade etmektedir.

2.1.5. İki Değişkenli LM'de Kestirim

2.1.5.1. En Çok Olabilirlik Yöntemi (EÇO)

DR'ye ilişkin parametre kestirimi genellikle EKK yardımıyla tahmin edilir. EKK y 'nin kestirilen değerlerinin gözlenen değerlerine en uygun parametre tahminlerini en küçük hata kareler toplamı ile ölçerek bulmaya çalışır. EÇO birçok modelde yer alan genel bir yöntemdir. EKK ile kıyaslandığında EÇO lineer olmayan karışık modellere uygulandığında DM gibi doğru sonuçlar verebilir. Lojistik modelde doğrusal bir model olmadığı için LR'de EÇO tahmin yöntemi olarak kullanılabilir.

Lojistik modelde koşullu ve koşulsuz olmak üzere parametre tahmininde kullanılacak iki alternatif EÇO yöntemi bulunmaktadır. Kestirilecek parametre adeti gözlenenden az ise koşulsuz, fazla ise koşullu EÇO tercih edilir. Pratikte tahmin edilen parametre sayısı genellikle gözlem sayısından az olduğu için koşulsuz yöntem kullanılmaktadır. Bu sebeplerle koşulsuz EÇO en yaygın şekilde kullanılan yöntemdir. Olasılık modelinde denklemin sol tarafında yer alan olasılık P_i ise $P(y = 1, x)$ 'dir. Y , 1 ve 0 değerlerini alacağı için her bir gözlem P_i parametresi ile Bernoulli dağılıma uygunluk gösterir. Modelde yer alan n gözlem için olabilirlik fonksiyonu;

$$P(y/x) = \prod_{i=1}^n P_i^{Y_i} \cdot (1 - P_i)^{1-Y_i} \quad (2.9)$$

şeklinde tanımlıdır.

Y değerini gözleme olabilirliğini maksimum yapan bir parametre tahmini seçilmelidir. Elde edilen olabilirlik fonksiyonlarına ilişkin değeri en büyük yapan β değeri $\tilde{\beta}$ olarak belirlenir. Lojistik modele ilişkin olabilirlik fonksiyonu;

$$L\left[\left(\frac{Y}{x}, \beta\right)\right] = \prod_{i=1}^n \left[\frac{\exp(\sum_{k=0}^p \hat{\beta}_k x_{ik})}{1 + \exp(\sum_{k=0}^p \hat{\beta}_k x_{ik})} \right]^{Y_i} \left[\frac{1}{1 + \exp(\sum_{k=0}^p \hat{\beta}_k x_{ik})} \right]^{1-Y_i}$$

eşitliğinde gösterilmektedir. Eşitliğin logaritması alınarak aşağıdaki eşitlik elde edilir.

$$\text{Log } L\left[\left(\frac{Y}{x}, \beta\right)\right] = [Y_i \log P_i + (1 - Y_i) \log(1 - P_i)] \quad (2.10)$$

Olabilirlik fonksiyonunu maksimum yapan bu değer aynı zamanda bu fonksiyonun logaritmasını da maksimize edecektir. β' e göre alınan türevlerin 0'a eşitlenmesiyle $\hat{\beta}$ değeri bulunur.

Olabilirliğe ilişkin değerler $\hat{\beta}'$ de doğrusal olmadığı için bir seferde net sonuç elde edilemez ve iteratif çözümlenmeye ihtiyaç duyulur. İteratif çözümlenmede β değerlerine başlangıç değeri verilir. Ortaya çıkan ilk tahminlerde δ şeklinde ayarlama yapılarak türev alınır ve EÇÖ değeri tahmini gerçekleştirilir. δ terimi iterasyon değerini farklılaştırmadığı durumda yakınsamıştır. En az iterasyon ile uygun çözüme ulaşılmak hedeflenir. İtersayon adetini düşürmenin yöntemi ilk değeri uygun seçmektir [35].

Fazla sayıda örnekleme tahmin edicinin iyi özelliklere sahip olması EÇÖ tahmin yönteminin kullanımını yaygınlaştırmıştır. Bu yöntem asimtotik etkin, yansız ve normaldir. Tahminlerin örnekleme dağılımının normale yakın olması güven aralıkları oluşturulmasında normal ve Ki-Kare (χ^2) dağılımlarının kullanılmasına imkan sağlamaktadır [34].

2.1.5.2. İteratif En Küçük Kareler

En küçük kareler yöntemi, DR'de parametre tahmini için sıklıkla kullanılan bir yöntemdir. Kestirilen bağımlı değişkenlerin gözlem değerinden sapmalarını minimum yapan β değerlerine ulaşılır. EKK yöntemi çoğunlukla oldukça doğru tahminler üretirken; bağımlı değişkenin kesikli olduğu durumda varsayımları tam olarak sağlayamamaktadır [35].

2.1.5.3. Minimum Lojit Ki-Kare Yöntemi

Ağırlıklı en küçük kareler yönteminin özel bir biçimi olan ve Berkson [14] tarafından geliştirilen bu yöntemde çapraz tablolardaki beklenen ve gözlenen lojit değerleri arasındaki farktan yararlanılmaktadır. Tekrarlı veriler olması durumunda bu yöntem kullanılmaktadır. İteratif en küçük kareler yönteminde verilen P_j olasılığı üzerinden yapılan lojit dönüşümü bu yöntemde sonuç değişkenini oluşturmaktadır. Kestirimde bulunulan ağırlık değerleri $n_j \cdot P_j(1 - P_j)$ şeklinde elde edilmektedir. Lojit değeri olarak tanımlanan sonuç değişkeninin açıklayıcı değişkenler ile regresyonundan EKK kestirimlerini elde etmeye dayanmaktadır. Tek adımda bulunan EKK kestirimleri minimum lojit ki-kare kestirimi adını almaktadır [35].

2.1.6. Uyum İyiliği Hesaplamaları

Dahil edilmesi gerekli değişkenler dahil edildiğinde, modele tepki değişkeninin açıklanmasındaki etkinlik uyum iyiliği kontrolü ile mümkündür [37]. Modelde logaritmik dönüşüm yerine başka dönüşüm kullanılması gerektiğinde, gerekli değişkenler modele alınmadığı durumda, gerekli değişkenlerin modelde olmasına rağmen ölçeğin yanlış olması durumunda ve modelde aykırı değerlerin bulunması hallerinde varsayım bozulması mevcut olup bunlar için uyum iyiliği incelemesi gerçekleştirilir.

Lojistik modelde normallik varsayımı kısıt olmadığı için uyum iyiliği testlerinde kullanılan t ve F tablo değerleri karşılaştırma amacıyla kullanılamamakta; bunlar yerine ki-kare, G^2 gibi parametrik olmayan ölçütlerden yararlanılmaktadır [35].

O gözlenen, E beklenen değerleri OlogO ve OlogE sırasıyla gözlenen ve beklenen olabirlikleri göstermek üzere; parametrik olmayan ölçütlerden yararlanan uyum iyiliği testlerinde kullanılan G^2 aşağıda yer alan formüldeki şekilde tanımlanmıştır.

$$\chi^2 = \sum(O - E)^2/E$$

$$G^2 = -2 \sum O \text{LOG} \left(\frac{E}{O} \right) = 2 \sum O \text{LOG} \left(\frac{O}{E} \right) \quad (2.11)$$

$$G^2 = -2 \left[\sum O \text{LOG}(E) - \sum O \text{LOG}(O) \right] = -2 [\text{LOG}(E) - \text{LOG}(O)]$$

Bu ölçüm, basit bir hesaplama mantığına dayanması nedeniyle sıklıkla tercih edilmektedir [34].

Lojistik regresyon analizinde G^2 yaklaşımı, modelde kullanılması gereken açıklayıcı değişkenleri belirlemede ve kurulan modelin önemliliğini test etmede kullanılmaktadır. Burada önerilen sapma ölçütü, doymuş model; değişken sayısı kadar parametre içeren model ile kestirilmiş model; sadece önemli olduğu düşünülen değişkenleri içeren model olmak üzere;

$$D = -2\log(\text{kestirilen modelin olabilirliği})/(\text{doymuş modelin olabilirliği}) \quad (2.12)$$

biçiminde tanımlanmaktadır.

Sabit katsayı olan β_0 harici katsayıların 0'a eşitliğini test eden ve F testine karşılık gelen istatistiktir.

$$H_0 := \beta_1 = \beta_2 = \beta_3 = \dots = \beta_k = 0 \quad (2.13)$$

ve

$$H_1 = \text{En az biri farklıdır} \quad (2.14)$$

Temelde benzerlik oranı prensibine dayanır.

$$C = -2\log\left(\frac{L_0}{L_1}\right) \quad (2.15)$$

Eşitlik (2.15)'de L_0 , Sabit katsayı haricindeki tüm katsayıların sıfıra eşit olması halinde olabilirlik değerini; L_1 , verilere ilişkilendirilen bütün modele ilişkin olabilirlik fonksiyon değerini göstermektedir.

3. KANTİL REGRESYON VE RİSK HESAPLAMA YÖNTEMLERİ

Kantil regresyon, koşullu kantil fonksiyonların model tahmini için istatistiksel bir yöntem olarak ortaya çıkmaktadır. Klasik en küçük kareler (EKK) tahmin edicileri koşullu ortalama yöntemlerini temel alarak, bağımlı değişkenlerin tüm tepki dağılımı üzerindeki şekil, konum ve ölçek değerlerini belirlemek için sistematik bir strateji önermektedir. Fakat bağımlı değişkenin asimetric dağılıma sahip olduğu durumda veride değişen varyanslılık durumu ortaya çıkmaktadır. Bu durumun temel sebebi ortalamanın uç değerlere duyarlı olmasıdır. KR Koenker ile Bassett [38] tepki değişkeninin normal dağılımdan farklılık gösterdiğinde oluşan zorluğun üstesinden gelmek için medyan regresyon kullanmışlardır. MR, serbest değişkende meydana gelen farkın, dağılımın medyanında (0.5 kantil) ortaya çıkan etkisini açıklamaktadır. Kantil regresyon ise farklı kantiller için bağımlı değişkeninin dağılımına ilişkin değerleri tahmin etmektedir. Dağılımın kuyruğuna ilişkin her iki uçtaki kantil değerleri ve bunların koşullu değerleri incelenmekte olup, simetrik olmayan dağılıma sahip tepki değişkenine ilişkin değerler yansız ve doğru olarak tahmin edilebilir. Bu bölümde kantil kavramı açıklanacak, dağılımın modellenmesi ve kantil regresyon hakkında bilgi verilecektir. Ayrıca, bölümün devamında risk hesaplama yöntemlerinden bahsedilecektir.

3.1. Kantil

Kantiller genel olarak bir seriyi eşit büyüklükte parçaya bölerler. Bir seri eşit olarak iki parçaya bölünürse medyan, dört eş kısma bölünürse çeyreklik, on eş kısma bölünürse ondalık (desil) ve 100 eş kısma bölünürse ise santil olarak adlandırılmaktadır. Bölünme şekilleri ve sayılarına bağlı olarak kantil çeşitleri yüzdeler, desil ve kantil olarak sınıflanmaktadır.

Veriler azdan çoğa sıralandığında, seriyi 100 eş kısma ayıran tutarlara yüzdeler denilmektedir. En küçüğü P_1 'den başlayan ve P_{99} ' a kadar giden toplam 99 adet yüzdeler bulunmaktadır.

Veriler azdan çoğa sıralandığında, seriyi 10 eş kısma ayıran tutarlara desil denilmektedir. En küçüğü D_1 'den başlayan ve D_9 ' a kadar giden toplam 9 adet desil bulunmaktadır.

Veriler azdan çoğa sıralandığında, seriyi 4 eş kısma ayıran tutarlara kantil denilmektedir. En küçüğü Q_1 'den başlayan ve Q_3 ' a kadar giden toplam 3 adet kantil bulunmaktadır.

3.2. Dağılımın Modellenmesi

Dağılımların modellenmesinde ortalama, medyan, mod şeklinde merkezi eğilim ölçüleri ile varyans, standart sapma şeklinde yayılım ölçüleri kullanılmaktadır. Fakat çarpık ve basık dağılımlarda bu ölçütler bazen yeterli olmamaktadır. Bu kısımda anakütlenin şekil ile konumunu belirlemede uygulanan olasılık yoğunluk fonksiyonu (OYF), kümülatif dağılım fonksiyonu (KDF), kantil fonksiyonu (KF) ve kantil yoğunluk fonksiyonu (KYF) açıklanacaktır.

KDF herhangi X değişkeninin, yanıt rastlantı değişkeninden düşük çıkma ihtimalini gösteren düzenli yükselen ve kesikli olmayan fonksiyondur [39]. $F(x)$ fonksiyonu

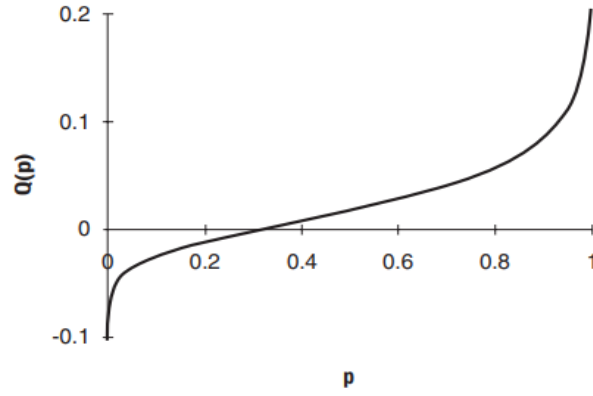
$$F(x) = P(X \leq x)$$

biçimde tanımlanmakta ve $0 \leq F(x) \leq 1$, $x \in R$ koşulunu sağlamaktadır.

X değeri $-\infty$ ' a yaklaşırken fonksiyon 0'a; x değeri $+\infty$ 'a yaklaşırken fonksiyon 1'e yaklaşmaktadır. Olasılık yoğunluk fonksiyonu değişkenin alabileceği değerler ile söz konusu değerlere ilişkin gerçekleşme olasılıkları arasındaki ilişkiyi gösteren fonksiyondur. Kümülatif dağılım fonksiyonunun türevi alındığı zaman olasılık yoğunluk fonksiyonu elde edilir.

3.2.1. Kantil Fonksiyonu

Kantil Fonksiyonu (KF), dağılımı tanımlamanın farklı bir yolu olarak ifade edilebilir ve $Q(p)$ ile gösterilmektedir. $x_p = P(X < x_p) = p$ olup $Q(p)$ ' ye ilişkin grafik Şekil 3.1'de gösterilmiştir.



Şekil 3.1. Kantil Fonksiyonu, $Q(p)$ [39]

$Q(p)$ F, KDF'nin kantil fonksiyonu şeklinde tanımlanır. KF, KDF'nin ters fonksiyonu olarak ifade edilebilir [40].

$$Q(p) = F^{-1}(p) = \inf\{x: F(x) \geq p\} \text{ olup } 0 < p < 1 \quad (3.1)$$

$Q(p)$, p 'ye ilişkin $0 < p < 1$ arası kantilleri gösterir.

3.2.2. Kantil Yoğunluk Fonksiyonu

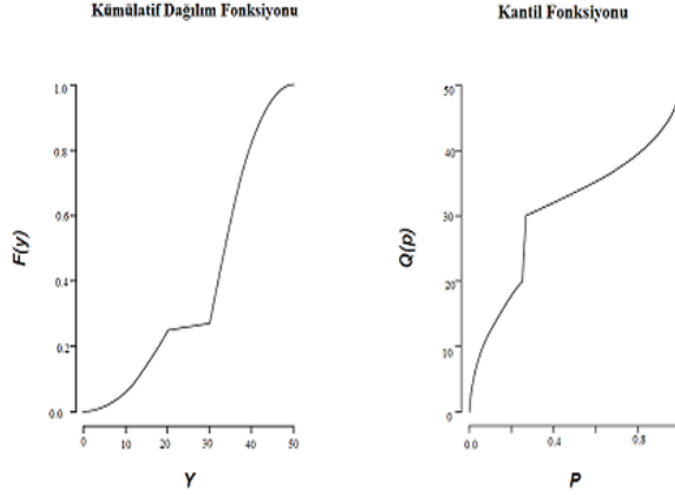
KDF'nin türevi olasılık yoğunluk fonksiyonuna eşit olup, KF türevlendiğinde Kantil Yoğunluk Fonksiyonuna eşittir.

$$q(p) = dQ(p)/dp$$

Kantil fonksiyonunun eğimi hiçbir zaman negatif olamaz çünkü azalmayan bir fonksiyona sahiptir. Bu sebeple $q(p)$, KYF sıfır ile bir arasında bir değer almaktadır. OYF sonsuz aralıkta tanımlıdır. KYF ve OYF arasındaki geçiş

$$q(p) = \frac{1}{f(F^{-1}(p))'} \text{ ve } f(x) = \frac{1}{q(Q^{-1}(x))'} \quad (3.2)$$

olarak ifade edilmiştir. KDF ve KF'ye ilişkin gösterimler Şekil 3.2'de gösterilmiştir.



Şekil 3.2. Kümülatif Dağılım Fonksiyonu ve Kantil Fonksiyonu [41]

y_1, y_2, \dots, y_n ile ifade edilen dağılımda, örneklemin dağılımına ilişkin KDF rastlantı y 'den düşük veya eşit olan örneklemin yüzdesini göstermektedir.

KDF

$$F^{\sim}(y) = P^{\sim}(Y \leq y) = \left(\frac{1}{n}\right) \sum_{i=1}^n I(Y_i \leq y) \quad (3.3)$$

veya

$$F^{\sim}(y) = P^{\sim}(Y_1, Y_2, \dots, Y_n \leq y) \quad (3.4)$$

şeklinde gösterilip, eşitlikte yer alan $I(Y_i \leq y)$, $Y \leq y$ olduğunda 1'e, $Y > y$ olduğunda ise sıfıra eşittir [40].

3.2.3. Kantil Dağılıma İlişkin Parametreler

$f = F^{\sim}$ olasılık yoğunluk fonksiyonlu ve $Q^{(p)}$ KF dağılım kullanan geniş bir y_1, \dots, y_n örnekleme örneği için dağılım $\hat{Q}^{(p)}, Q^{(p)}$ ortalamalı ve $\frac{p(1-p)}{n} \cdot \frac{1}{f(Q^{(p)})^2}$ varyanslı ve yaklaşık olarak normal dağılmaktadır. Örneklemin dağılım varyansı kantil için kullanılırken OYF'den faydalanılır. Örneklemin kantil değeri yoğunluk fazla ise çok değişkenlik göstermezken; kantil değeri yoğunluğun az olduğu durumda daha çok değişkenlik gösterir. Medyan ise ortalama için çarpık bir dağılımı ifade eder [41].

Çarpıklık (QSK) ile basıklık (QSC), değişkene ait dağılımı ifade eden ortalama ile standart sapmaya ilişkin terimler olup dağılım şeklinin iki ölçüsüdür. Çarpıklık, ortalama ve standart sapmadan sonra diğer bir momenttir. Dağılımla ilgili orantısızlığı ya da asimetrikliği gösterir. Simetrik olan dağılımda çarpıklık sıfır, pozitif ise sağa çarpıklığı ifade eder. Basıklık ise bir diğer momenttir. Normal dağılıma uygunluk gösterdiği durumda basıklık katsayısı üçtür. Üçten fazla bir değer normal dağılıma kıyasla sivri olduğunu, üçten az ise normal dağılımdan daha basık bir durumu göstermektedir [42].

$$QSK^{(P)} = \frac{(Q^{(1-p)} - Q^{(0.5)})}{(Q^{(0.5)} - Q^{(p)})} - 1 = \frac{Q^{(1-p)} - Q^{(p)} - 2Q^{(0.5)}}{Q^{(1-p)} - Q^{(p)}} \quad 0 < p < 0,5 \quad (3.5)$$

şeklinde ifade edilir.

Kantil merkezli basıklık ise;

$$QSC^{(P)} = Q^{(1-p)} - Q^p, \quad p < 0,5 \quad (3.6)$$

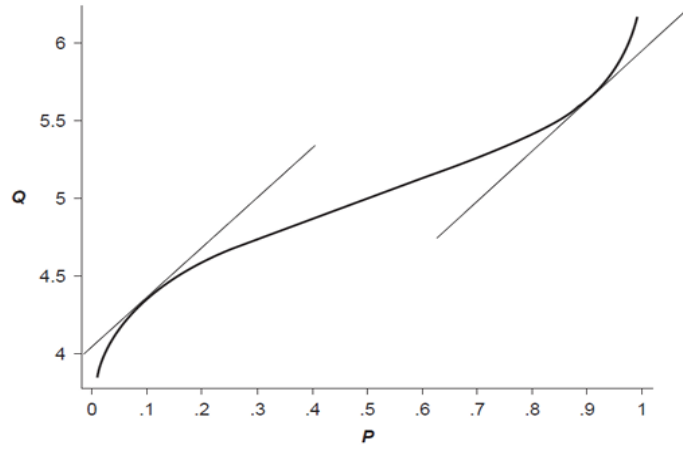
veya

$$QSC^{(P)} = \frac{Q^{(1-p)} - Q^{(p)}}{Q^{(1-n)} - Q^n} \quad 0 < p < 0,5 \quad 0 < p, n < 1 \quad (3.7)$$

olarak ifade edilir [43].

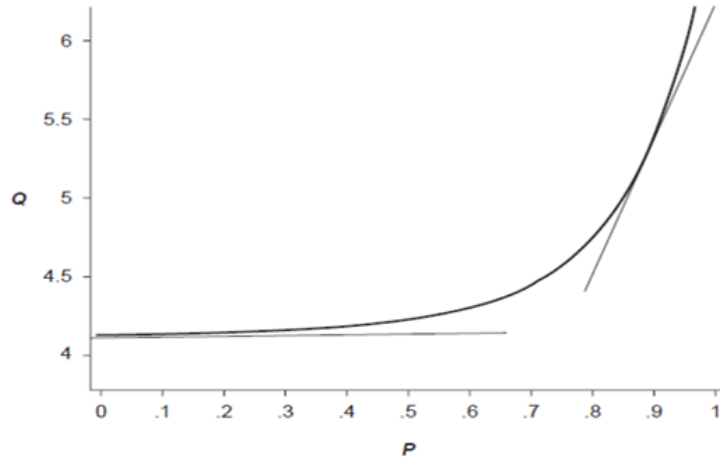
Basıklık ölçümünde genellikle standart sapma da kullanılır. Simetrik dağılımlarda standart sapma basıklığı tanımlamada etkili iken asimetrik dağılımlarda yanlış değerlendirmelere yol açabilir. Basıklığın ölçülmesi için birbirini tamamlayan iki kantil değeri birbirinden çıkartılır. Dağılımı ifade ederken çarpıklık değeri gözlenen değerlerin ortalama ile farklarının kübik fonksiyonu şeklinde gösterilir. Simetrik olan dağılıma ilişkin çarpıklık sıfırdır. Değer eksi ise çarpıklık sola, artı ise sağa doğrudur [41].

Birbirinin simetriği olan iki KF'ye ilişkin eğim Şekil 3.3'te görüldüğü gibi eşittir.



Şekil 3.3. Normal Kantil Fonksiyonu [41]

Şekil 3.4'te ise sağa çarpık bir dağılım gösterilmekte olup kantil fonksiyon değerleri asimimetrik olur.



Şekil 3.4. Çarpık Kantil Fonksiyonu [41]

3.3. Kantil Regresyon

KR, bağımlı değişkene ilişkin kantillerin koşulunu kestirmede yararlanılan lineer modeli ifade eder. İlk olarak 18.yüzyılda Boskovich medyan regresyon üzerine çalışma gerçekleştirmiştir. Daha sonra bu çalışmalar Koenker ve Basett [38] tarafından kantil regresyon modeli olarak geliştirilmiştir.

3.3.1. Kantil Regresyon Modeli

Kantil regresyon modeli oluşturulurken koşullu kantil fonksiyonlara ilişkin tahminlerin ağırlıklı mutlak sapmalar toplamı minimize edilir. Medyan (0,50'nci kantil) için ağırlıklar simetrik iken diğer kantillerde asimetrik ağırlık kullanılır. EKK ile koşullu ortalamaya ilişkin kestirim kullanılmakta iken, KR bağımlı değişkene ilişkin farklı bölgedeki belirleyiciler açıklanır. 0,60'ıncı kantil değerinde bir modelin kurulması durumunda, gözlenen değerlerin %60'ının tahminin aşağısında; %40'ının ise tahminin yukarısında yer alacağı belirlenir.

Doğrusal regresyon modelinde koşullu ortalamalar ile fonksiyon belirlenirken, kantil regresyon modelinde koşullu kantil fonksiyonu tarafından belirlenir. Dağılımın simetrik olduğu durumda medyan ve ortalama eşit olmaktadır. Fakat dağılım çarpık ise ortalama ve medyan farklı olmakta ve bu durumda dağılıma ilişkin merkezi eğilim ölçüsü olarak medyan değerinin kullanılması daha uygun olmaktadır [41].

Kantil regresyon tepki değişkeninin, bağımlı değişkenlerinin şekil ve ölçek konumuna nasıl etki ettiğine ilişkin sistematik metod sunmakla birlikte koşullu kantil fonksiyonların tahmin modelleri için oluşturulmuş istatistiksel bir metodolojidir [23]. KF'ler, hataların net tutarının ağırlıklandırılmış toplamının minimizasyonu ile kestirilir. Koşullu KF'nin bütünüyle kestirimi, tepki değişkeninin dağılımında katsayı etkilerine ilişkin bilgi sunar. Veride değişken varyans mevcut ise EKK sonuçları anlamsız olmaktadır. Böyle hallerde KR değişken varyansı dikkate almaktadır. Kantil uç hasara karşı güçlüdür. Veride uç değerler var ise ortalama bu değerlerden olumlu etkilenmeyebilir. Fakat medyana veride yer alan uç değerlerden etkilenmez. Diğer kantillerde, medyan gibi benzer özelliğe sahiptir.

KR, hataların normal dağılmadığı kabulü ile hareket eden metottur. Güçlü bir yöntemdir. KR bazı kantillere spesifik model sağlayan klasik DRM'lere ilişkin kısıtları aşma imkanı sağlamaktadır. KRM, aşağıdaki eşitlik ile gösterilmektedir.

$$Y_i = \beta_0^{(p)} + \beta_1^{(p)} X_i + \epsilon_i^{(p)} \quad 0 \leq p \leq 1 \quad (3.8)$$

Bu eşitlikte $\beta_0^{(p)}$ ve $\beta_1^{(p)}$ parametreleri, x üzerinde koşullu y'nin p'nci kantil değeri için bir model tanımlamaktadır. KR'da p'nci kantile ilişkin parametreleri tahmin etmek için aşağıda yer alan eşitlik minimize edilmektedir.

$$f(\beta_0^{(p)}, \beta_1^{(p)}) = \sum_{i \setminus y_i < \hat{y}_p(x_i)} (1-p) |y_i - \hat{y}_p(x_i)| + \sum_{i \setminus y_i > \hat{y}_p(x_i)} (p) |y_i - \hat{y}_p(x_i)|$$

Burada $\hat{y}_p(x_i) = \beta_0^{(p)} + \beta_1^{(p)}$ 'dir. y_i gözlemi ve p'nci kantil $\hat{y}_p(x_i)$ farkı kantil doğrusunun aşağısında olduğu durumda (1-p), gözleme ilişkin çizginin yukarısında olduğu durumda p olarak ifade edilir. Bu durumda değerlerin % p'si çizginin altında % (1-p)'si ise çizginin üzerinde kalmaktadır [44].

DRM $\varepsilon_i = 0$ iken, verilen bir x_i için y_i 'nin koşullu ortalaması

$$E(y_i|x_i) = \beta_0 + \beta_1 x_i$$

olup, KR'de ise verilen bir x_i için p'nci koşullu kantil,

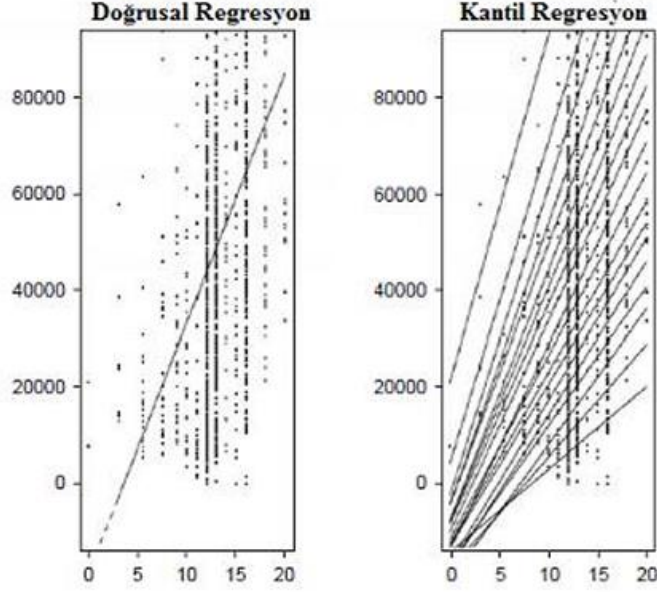
$$Q^{(p)}(y_i \setminus |) = \beta_0^{(p)} + \beta_1^{(p)} x_i \quad (3.9)$$

olarak ifade edilir. Bu nedenle koşullu p'nci kantil $\beta_0^{(p)}$ ve $\beta_1^{(p)}$ özel parametre ile x_i 'nin spesifik değeri ile ifade edilir. $\beta_0 + \beta_1 x_i$ sürekli olup

$$Q^{(p)}(y_i \setminus x_i) = \beta_0^{(p)} + \beta_1^{(p)} x_i + Q^{(p)}(\varepsilon_i) = \beta_0^{(p)} + \beta_1^{(p)} x_i \quad (3.10)$$

olduğu için hata teriminin p'nci kantili sıfır olmalıdır.

Yalnızca bir tek koşullu ortalamalı DRM'den değişik olan KRM birçok farklı kantil değerine sahiptir. Şekil 3.5'te DRM için çizilen soldaki şekil ortalama değişimi gösterirken, KRM için çizilen şekil farklı KR çizgilerini göstermektedir. Doğrusal Regresyon ve Kantil Regresyon Çizgileri Şekil 3.5'te gösterilmektedir.



Şekil 3.5. Doğrusal Regresyon ve Kantil Regresyon Çizgileri [41]

$Y_i = \beta + \varepsilon_i$ ile ifade edilen yerleştirme modeli ile kantil kullanılırsa, F fonksiyonu sürekli ise $Q^{(p)}$ şeklinde ifade edilen p'nci kantil ($0 < p < 1$), y'nin,

$$P(y < Q^{(p)}) = F(Q^{(p)}) = p \quad (3.11)$$

eşitliği sağlamaktadır. Burada Y_i , simetrik F dağılım fonksiyonuna sahip, bağımlı olmayan ve de özdeş dağılan, β medyana sahip bir değişkeni ifade eder. Modelde p'nci kantil,

$$\min_{\beta} \frac{1}{n} \left\{ \sum_{i/y_i \geq \beta} p |y_i - \beta| + \sum_{i/y_i < \beta} (1 - p) |y_i - \beta| \right\} \quad (3.12)$$

formülünden üretilir. Model genel olarak

$$Y_i = x_i' \beta + e_i \quad (3.13)$$

ile ifade edilir. Bu formülde e_i bağımsız, sıfırın etrafında simetrik ve de F dağılımına sahiptir. x_i' ise bağımsız değişken vektörüdür. Koşullu kantil tahmini ise ($0 < p < 1$) aralığında olmak üzere

$$P(e_i \leq 0 | x) = p \quad (3.14)$$

ifadesi ile çözülür [45].

Bu durumda ise p'nci kantil regresyon ($0 < p < 1$) minimizasyon problemine çözüm olarak

$$\min_{\beta} \frac{1}{n} \left\{ \sum_{i/y_i \geq x_i' \beta} p |y_i - x_i' \beta| + \sum_{i/y_i < x_i' \beta} (1 - p) |y_i - x_i' \beta| \right\} \quad (3.15)$$

ile ifade edilir. Bu durum

$$\min_{\beta} \frac{1}{n} \sum_{i=1}^n \rho_p(y_i - x_i' \beta) \quad (3.16)$$

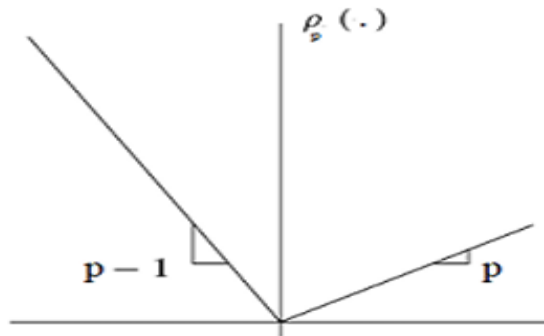
olarak ifade edilir. $\rho_p(\cdot)$, p'nci örnek kantiline çözüm olarak verilen ve şekilde gösterilen eğik mutlak değer fonksiyonudur [21]. Yani, ρ_p parçalı doğrusal fonksiyon olup

$$\rho_p(u) = u(p - I(u < 0)) \quad (3.17)$$

şeklinde tanımlanmaktadır. Bu durumda işaret fonksiyonu $I(\cdot)$ olup bu durumda $I(\cdot)$ hatalar artı ise sıfır, eksi ise bir olarak ifade edilir. Kantil regresyona ilişkin gösterim, p olasılığının herhangi bir değeri için simpleks doğrusal programlama gösterimidir [46]. Kontrol fonksiyonu ise

$$\rho_p(u) = \begin{cases} pu & u \geq 0 \\ (p - 1)u & u < 0 \end{cases} \quad (3.18)$$

biçiminde gösterilmektedir. Şekil (3.6)'da ise kontrol fonksiyonu gösterilmektedir.



Şekil 3.6. Kontrol Fonksiyonu [21]

$\rho_p(\cdot)$ kontrol fonksiyonu negatif ise $\rho_p(\cdot) = p - 1$ olur. $\rho_p(\cdot)$ kontrol fonksiyonu pozitif ise $\rho_p(\cdot) = p$ olur. Bu durumda doğrusal programlama yöntemleri kullanılarak kantil regresyondaki hesaplama zorluğu aşılabilir [47].

3.3.2. Kantil Regresyon Modelinin Faydaları

Regresyon teknikleri arasında en küçük kareler metodu en yaygın olarak kullanılan modeldir. Teknik hataların normal dağıldığı ve bağımsız olduğu varsayımı ile sonuçlar tahmin edilmektedir. Genelleştirilmiş doğrusal modeller gibi alternatif modeller ise verinin bazı karakteristik özellikleri ve dağılım çeşidinin güçlü varsayımlarını temel almaktadır. Fakat aktüerler pratik uygulamalarında modelleme aşamasında kabul edilen varsayımlardan daha karmaşık durumlarla karşılaşmaktadır. Bu durum geleneksel regresyon modellerinin yanlış ya da kullanışsız olmasına sebep olmaktadır. Bu durumda hasar dağılımının tahmininde yanılma, örneklemin spesifik özelliklerine dayanan tahmin hataları, homojenlik varsayımları ve/veya önceki veri setine ait olma varsayımlarını yok edebilmek mümkün olabilecektir. Bunun sonucunda hasar dağılımının tahmini, model hatası üreten gerçek dağılımdan çok farklı olabilir [30].

Kalın kuyruklu dağılımlarda, normal dağılım yaklaşımında kabul edilen değerlere göre, uç değerlere daha büyük ağırlık verilmesine ihtiyaç duyulmaktadır. Örnekleme yer alan uç değerler, sıklıkla dağılımı bozan ve önceki bilgilere dayanan düzgün veriden ayrılması imkânsız olan bozuk değerleri ifade eder. Bu durum örneklem dağılımını asimetric yapar ve geleneksel regresyon modelinin kullanılmasını kısıtlar. Bu sebeple örnekleme uygun olan ve daha hassas prosedürlerin kullanılması gerekli olabilir [30].

Ayrıca verinin bağımlılık yapısı karmaşık ise, bu yapıyı belirlemek zor hatta imkânsızdır. Bu nedenle örneklemin yapısını uygun modellerle analiz etmekte birtakım problemler ortaya çıkmaktadır. Bu durumda geleneksel regresyon metotları bazı pratik durumlar için uygun değildir. Sonuç olarak bu metotları temel alan çıkarımlar yanlış olabilir. Bu durum alternatif yaklaşımların kullanımını motive etmektedir. Söz konusu metotlardan bir tanesi, temel fikri ikinci dereceden sapmaları kesin olanlar ile değiştiren kantil regresyon modelidir. KR'nin avantajları aşağıda özetlenmektedir.

- Parametrik dağılım ailesinin varsayımlarına ihtiyaç duymayan herhangi bir varsayım içermeyen ve özelliklerini kullanmayan bir dağılım özelliği göstermez.
- Uç değerlere karşı oldukça güçlüdür.
- Bağımsızlık veya zayıf bir bağımlılık olmasına ihtiyaç duyulmaz [30].

Sonuç olarak KR metodu geleneksel regresyon modellerinin bazı dezavantajlarını ortadan kaldırmaktadır. Söz konusu varsayımlar aynı olduğu durumda geleneksel regresyon modelleri ile aynı sonuçları vermekle birlikte, varsayımların yanlış olduğu durumda KR metodunun daha iyi tahminler verdiği gözlenmektedir [30].

Kantil regresyonun diğer özelliklerine aşağıda değinilmektedir.

- Kantil tahmincilerin doğrusal kombinasyonuna dayanan L-tahmincileri, genel olarak EKK tahmincilerinden daha etkilidir [48].
- Farklı kantil değerleri için birbirinden farklı çözümler elde edilmesi, bağımlı değişkenin koşullu dağılımının değişik noktalardaki bağımsız değişkenlerdeki değişikliklere farklı cevaplar vermesi olarak ifade edilebilir [48].
- Hatalar normalden farklı dağılıyorsa KR tahmincileri, DR tahmin edicilerinden daha etkin olabilir [48].
- KR’de $p=0.5$ ise EKMS (medyan) regresyona ulaşılır [38].
- KR tahmincileri, lineer olarak programlanabilir [23].
- Veri setinde değişen varyans olduğunda, EKK yöntemi ile elde edilen tahminler güvenilir olmayabilir. Fakat böyle durumlarda KR değişen varyansı da ele almaktadır.
- KR sadece bir dağılımın merkezinin yanı sıra kuyruklarda yer alan değişkenlerin etkilerini ölçmek için de kullanılmaktadır.
- $\min_{\beta} \frac{1}{n} \sum_{i=1}^n \rho_{\theta}(y_i - x_i' b)$ ifadesinin minimizasyonu doğrusal programlamaya ilişkin gösterim olup bu sayede tahmin kolaylaşmaktadır.

3.3.3. Kantil Regresyon ve Doğrusal Programlama

Kantil regresyona ilişkin tahmin ediciler doğrusal programlama ile formüle edilebilir. 1950’lerden beri mutlak artıkların toplamlarını en aza indirmeye dayalı medyan regresyon yöntemlerinin doğrusal programlama problemleri olarak formüle edilmiş ve simplex algoritma yönteminin formları ile etkin bir şekilde çözülmüştür [21].

F dağılım fonksiyonuna sahip Y rastlantı değişkeni, β tahmin edilecek vektörün katsayısı, hata değeri $e_i = y_i - x_i\beta$ olan p' nci regresyon kantili ($0 < \theta < 1$);

$$\min_{\beta} \{ \sum_{i: y_i \geq x_i' \beta} p [y_i - \beta] + \sum_{i: y_i \leq x_i' \beta} (1 - p) [y_i - \beta] \} \quad (3.19)$$

eşitliğinin minimizasyonu ile elde edilir. Bağımsız değişken $\{y_i: i = 1, \dots, \dot{I}\}, \{x_i: i = 1, \dots, \dot{I}\}, i * k$ boyutlu matriste, β tahmin edilecek vektörün katsayısı ve hata değeri $e_i = y_i - x_i \beta$ olan p 'nci regresyon kantili ($0 < \theta < 1$)

$$\min_{\beta} \{ \sum_{i: y_i \geq x_i' \beta} p [y_i - x_i' \beta] + \sum_{i: y_i \leq x_i' \beta} (1 - p) [y_i - x_i' \beta] \} \quad (3.20)$$

eşitliğinin minimizasyonu ile elde edilir [38].

Sonlu sayıda simpleks iterasyonla tahmin yapılması, KR'nin doğrusal programlama gösterimini ifade etmektedir. Bu iterasyonun sayısı doğrusal programlama algoritmasından daha küçüktür [49]. Aşağıda ifade edildiği şekilde y_i bağımlı değişkeninin p 'nci kantili,

$$\hat{\beta}_p = \min_{\beta} \frac{1}{n} \{ \sum_{i: y_i \geq x_i' \beta} p [y_i - x_i' \beta] + \sum_{i: y_i \leq x_i' \beta} (1 - p) [y_i - x_i' \beta] \} \quad (3.21)$$

$$\hat{\beta}_p = \min_{\beta} \frac{1}{n} \sum_{i=1}^n \rho_p(y_i - x_i' \beta) \quad (3.22)$$

Eşitliğinin minimizasyonu yardımıyla gösterilmektedir. Bu durumu lineer program şeklinde göstermek için y_i yalnızca negatif olmayan elemanlara ilişkin fonksiyon olup, $\beta^1_{p_j} \geq 0, \beta^2_{p_j} \geq 0$ ve ($j = 1, \dots, K$) ve $e_{p_i} \geq 0, (i = 1, \dots, n)$ olduğu durumda aşağıdaki eşitlik elde edilir.

$$y_i = \sum_{j=1}^K x_{ij} \beta_{p_j} + e_{p_i} = \sum_{j=1}^K x_{ij} (\beta^1_{p_j} - \beta^2_{p_j}) + (e_{p_i} - v_{p_i}) \quad (3.23)$$

Bu durumda ilk eşitlik,

$$\min_{\beta} \frac{1}{n} \sum_{i=1}^n (p - \frac{1}{2} + 1/2 \operatorname{sgn}(y_i - x_i' \hat{\beta})) (y_i - x_i' \hat{\beta}) \quad (3.24)$$

şeklinde tekrar yazılabilir. $k * 1$ vektörlü birinci merteye koşulu

$$\frac{1}{n} \sum_{i=1}^n \left(p - \frac{1}{2} + \frac{1}{2 \operatorname{sgn}(y_i - x_i' \hat{\beta})} \right) x_i = 0 \quad (3.25)$$

şeklinde yazılır [48].

$$\min_{\beta} \frac{1}{n} \sum_{i=1}^n \rho_p(y_i - x_i' \beta) \quad (3.26)$$

Eşitlik (3.26) doğrusal programlama olarak tekrar formüle edilirse

$$\min_{(\beta, u, v) \in \mathbb{R}^p * \mathbb{R}_+^{2n}} \{pI_n^i u + (1-p)I_n^i v | X\beta + u - v = y\} \quad (3.27)$$

şeklinde gösterilir. X regresyon tasarım matrisine ilişkin gösterimdir. Hataya ilişkin vektör ise $y - x\beta$, hatalar negatif ve pozitif olarak ifade edilebilir. Dolayısıyla kısıta ilişkin kümede lineer fonksiyonun minimizasyonu sağlanabilir.

Gösterim;

$$\min_z c' z \quad \leftrightarrow \quad Az = y \quad z \geq 0$$

olarak ifade edilir. Burada,

$$A = (X, -X, I_n, -I_n), \quad y = (y_1, \dots, y_n)', \quad z = (\beta^1', \beta^2', e', v')', \quad c = (0', 0' p.l', (1-p) * l')', \quad X = (x_1, \dots, x_n)'$$

I_n : n boyutlu birim matrisini,

0^i : sıfırların k*1 vektörünü,

l : birlerin n*1 vektörünü ifade etmektedir. Doğrusal programlamanın dual problemi k*1 vektörlü birinci mertebe koşulu ile aynıdır ve

$$\max_w w' y \quad \leftrightarrow \quad w' A \leq c'$$

şeklindedir.

3.3.4. Kantil Regresyona İlişkin Parametre Tahmini

$\hat{\beta}_0$ ile $\hat{\beta}_1$ için çoklu doğrusal regresyon ile kalıntıların karelerini minimize eden parametre değerleri çözülür.

$$\min \sum_i (y_i - (\beta_0 + \beta_1 x_i))^2 \quad (3.28)$$

Yukarıda yer alan ifade $y = \hat{\beta}_0 + \hat{\beta}_1 x_i$ doğrusu ve $(x_i - y_i)$ arasında mesafelerin dikey karelerinin toplam değeridir. Denklem sıfır ile eşitlenir kısmi türev uygulanırsa çift bilinmeyenli olan çift denklem ortaya çıkar. Denklem ilişkisi eşitlik;

$$\hat{\beta}_1 = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_i^n (x_i - \bar{x})^2} \quad (3.29)$$

ve

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (3.30)$$

şeklindedir.

Kantil regresyonda, doğrusal regresyon parametre tahminlerinin kantil regresyon parametre tahminlerinden farklı olarak herhangi bir doğrudan noktaya olan uzaklığın dikey mesafenin toplam etkisinin kullanılarak ölçülmesidir. Etki ise doğru üzerindeki noktalar için p ve altındaki noktalar için (1-p)'dir. p= 0,10; 0,25; 0,50 gibi farklı seçimler farklı tahmini koşullu kantil fonksiyona sebep olur. Medyan regresyon için de benzer durum söz konusudur. Kalanların toplamını minimize ederek katsayılar bulunur.

$$\sum_i |y_i - \beta_0 - \beta_1 x_i|. \quad (3.31)$$

Yukarıda yer alan eşitlik minimize edilirse sonuç için, medyan regresyon doğrusuna bakılır. Kalan değerlerin yarısı pozitif yarısı negatiftir [41].

3.3.5. Belirlilik ve Düzeltilmiş Belirlilik Katsayıları

Modele ilişkin betimleyici fazla ise çoklu korelasyon olarak ifade edilir. Korelasyon katsayısının karesine çoklu korelasyon katsayısının karesi ya da çoklu belirlilik katsayısı adı verilir. R^2 ile gösterilir.

$$R_{y, X_1, X_2, \dots, X_K}^2 = \frac{\hat{\beta}_1 \sum yx_1 + \hat{\beta}_2 \sum yx_2 + \hat{\beta}_K \sum yx_K}{\sum y^2}$$

Fonksiyona ilave değişkenler eklendiği zaman azalan serbestlik derecesine dikkat ederek R^2 düzeltilir. Bu regresyon katsayısı ise;

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

şeklindedir.

Kantil regresyona modelleri için benzer ölçüm ise

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

olup R^2 ye benzer şekilde $R^1(\theta)$ değeri de (0,1) arasındadır [22].

3.4. Risk Hesaplama Yöntemleri

Risk hesaplama yöntemleri iki ana kategoriye ayrılmaktadır. Bu kategorilerden ilki, en optimal sigorta priminin hesaplanması için kullanılan primi esas alan risk hesaplama yöntemidir. İkinci kategori ise öz sermaye yeterliliğinin hesaplanması ve aynı zamanda en uygun sigorta risk priminin belirlenmesi için kullanılan sermaye yapısını dayanak alan risk hesaplama yöntemlerini içermektedir [50].

3.4.1. Prim Tabanlı Risk Hesaplama Yöntemleri

Aktüerya biliminde riskin nicel hale getirilmesi ve fiyatlandırılması için birtakım modeller kullanılmaktadır. Riski nicel hale getirmek, raslantısal gelecek kazanç ya da kayıpları kesin eşitlik haline getirmek ve bu eşitliği farklı riskler için karar verme amacı ile kullanmak için oldukça gereklidir. İlk ve geleneksel olan primi esas alan risk hesaplama yöntemidir.

Prim hesaplama yönteminde rastlantı değişkeni olan X hasar tutarı olarak ifade edilir ve olasılık uzayı olan (Ω, F, P) 'de negatif olmayan kümede tanımlanmaktadır. Eğer H risk hesaplama yöntemi olarak prim hesaplama fonksiyonunu ifade ederse, prim hesaplama yönteminin fonksiyonu $H : X \rightarrow R$ olup negatif olmayan değerleri hasar değişkeni olan X 'e tanımlar.

Raslantı deęişkenin marjinal daęılım fonksiyonu, hasar riskinin yeterli sigorta priminin hesaplanması suretiyle transfer edilmesi aısından oldukça önemlidir. Hasar tutarlarına iliřkin sınırların izilerek riskin belirlenmesi önemli olup, sınırları izilmeyen hasar tutarları sonsuz prim hesaplamasına neden olacağı için risklerin sigorta řirketleri tarafından sigortalanmasını engelleyecektir [51]. Prim esaslı risk hesaplama yöntemlerinden yaygın olarak kullanılan bazı modeller ařaęıda aıklanacaktır. TSE alışmasında prim tabanlı risk hesaplama yöntemlerini göstermiştir [54].

3.4.1.1. Net Prim Hesaplama Prensibi

Yalnızca gerekleşmesi beklenen hasara iliřkin beklenen deęeri hesaplayan en basit yöntem olarak kullanılmaktadır.

$$H(X) = E(X) \quad (3.32)$$

Risk için herhangi bir yükleme yapılmaz. Bu yöntem yalnızca risksiz (risk nötr) sigortalı için uygun bir hesaplama saęlar. Ama ortaya ıkabilecek farklı riskler için sigorta řirketinin sermaye aığı ile karřılařmasına sebep olabilir.

3.4.1.2. Beklenen Deęer Prim Prensibi

Beklenen deęer prensibi ile prim hesaplamada α yükleme faktörü $\alpha \geq 0$ olarak beklenen deęer tutarının üzerine eklenir. Kullanımı oldukça basit ve anlaşılır olduęu için sigorta řirketleri tarafından yaygın bir řekilde kullanılmaktadır. Formülü ise;

$$H(X) = (1 + \alpha) E(X) \quad (3.33)$$

řeklinde gösterilir. Fakat bu yöntem büyük hasar tutarlarına karřı hassas bir hesaplama gerekleřtiremez.

3.4.1.3. Standart Sapma Prim Prensibi

Standart sapma prim prensibi ile prim hesaplamada α yükleme faktörü $\alpha \geq 0$ olarak beklenen deęer tutarının üzerine baęımsız deęişkenin standart sapma deęeri arpılarak eklenir. Formülü ise;

$$H(X) = E(X) + \alpha\sqrt{Var(X)} \quad (3.34)$$

şeklinde gösterilir. Bu yöntemin en önemli özelliği hasar tutarlarında meydana gelen değişime karşı duyarlı bir hesaplama gerçekleştirilmesine imkân sağlamasıdır.

3.4.1.4. Varyans Prim Prensibi

Varyans prim prensibi ile prim hesaplamada α yükleme faktörü $\alpha \geq 0$ olarak beklenen değer tutarının üzerine bağımsız değişkenin varyans tutarı çarpılarak eklenir. Formülü ise;

$$H(X) = E(X) + \alpha Var(X) \quad (3.35)$$

şeklinde gösterilir. Bu yöntemin en önemli özelliği büyük hasar tutarlarında meydana gelen değişime karşı duyarlı bir hesaplama gerçekleştirilmesine bir imkân sağlar.

3.4.1.5. Üstel Prim Prensibi

Üstel prim prensibi ile prim hesaplamada α yükleme faktörü $\alpha \geq 0$ olarak;

$$H(X) = \frac{1}{\alpha} \ln E(e^{\alpha X}) \quad (3.36)$$

şeklinde gösterilir. Fakat bu yöntem ile prim hesaplaması yapılırken birçok varsayım kullanılması gerekmekte olup sigorta prim hesaplaması için pratikte kullanılan bir yöntem değildir.

3.4.1.6. Escher Prim Prensibi

Escher prim prensibi ile prim hesaplamada α yükleme faktörü $\alpha \geq 0$ olarak aşağıdaki eşitlik ile gösterilir.

$$H(X) = \frac{E(Xe^{\alpha X})}{E(e^{\alpha X})} \quad (3.37)$$

Bu yöntem ile uç hasar tutarlarına hassas sigorta prim tutarı hesaplamaları gerçekleştirilebilir.

3.4.2. Sermaye Yapısını Dayanak Alan Risk Hesaplama Yöntemleri

Yüzdesel prensiplerin kullanılmasına bağlı risk hesaplamasına, sermaye yapısını dayanak alan risk ölçüm yöntemi denir ve sigorta şirketleri ekonomik sermayenin oluşturulmasında bu yöntemi kullanır. Ekonomik sermaye belirli bir zamanda ekonomik açıdan faaliyetlerini sürdürebilecek düzeyde tutulması gereken sermaye olarak ifade edilir [52]. Bu bölümde yaygın şekilde kullanılan Riske Maruz Değer (RMD) ve Koşullu Riske Maruz Değer (KRMD) yöntemleri açıklanacaktır.

3.4.2.1. Riske Maruz Değer

Standart bir risk ölçüsü olarak RMD'nin kullanımı finansal ve aktüeryal alanda oldukça yaygındır. Belirli bir zaman diliminde normal piyasa koşullarının devam ettiği varsayımı ile belirli bir güven seviyesinde beklenen hasara ilişkin tutardır [25]. Belirli bir dilimde dağılımın kantil değerini göstermekte olup, riskin volatilitelerini ve tutarını göstermektedir. Sabit bir değer ile riskin tutarını verebilmektedir. Bu sebeple şirketler kendine özel sermaye tutarını RMD ile belirler. Hasarlar $1 - \alpha$ olasılık değeri ile oluşursa α , hasar dağılımının kantil değeri olarak ifade edilir.

$$RMD_{\alpha}(X) = F_X^{-1}\alpha = x_{\alpha} \quad (3.38)$$

Fonksiyonunu göstermektedir. X hasarı ifade eden rastlantı değişkeni olmak üzere F^{-1} ters kümülatif dağılım olarak ifade edilir. Bazı durumlarda X sürekli raslantı değerlere sahip olmayabilir bu durumda RMD'nin genel formülü;

$$RMD_{\alpha}(X) = \inf\{x \in [0, \infty): F_X(x) \geq \delta\} \quad (3.39)$$

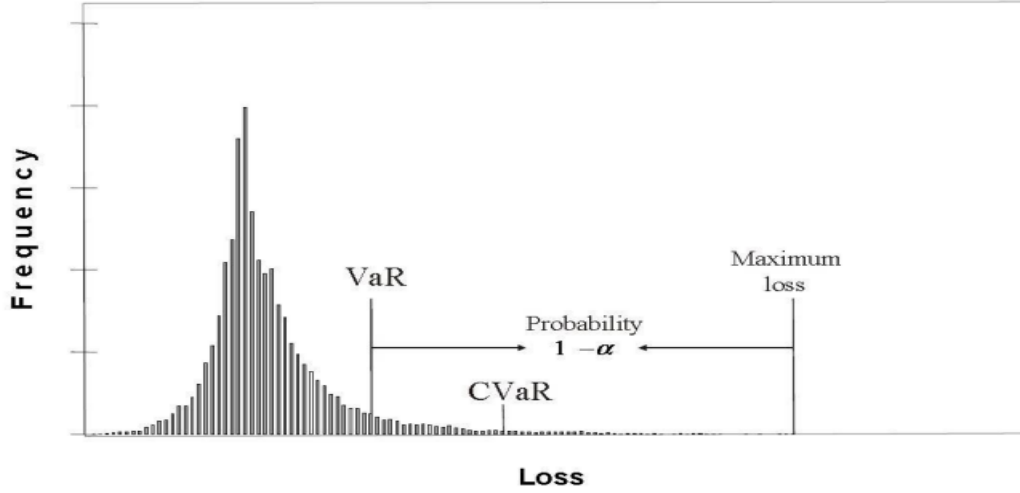
şeklinde gösterilir.

3.4.2.2. Koşullu Riske Maruz Değer

RMD yönteminin kuyruk bölgesinde yer alan risk ölçümünde yetersiz olmasından kaynaklı alternatif bir yöntem olarak KRMD geliştirilmiştir. KRMD $(1 - \alpha)$ değerinden sonra yer alan kuyruk kısmındaki riskin ölçümünü göstermektedir. KRMD aşağıdaki eşitlik ile gösterilmektedir.

$$KRMD_{\alpha}(X) = E[X|X > RMD_{\alpha}(X)] \quad (3.40)$$

Şekil 3.7 RMD ve KRMD'ye ilişkin grafiksel gösterimim açıklamaktadır. Şekil 3.7'de görüldüğü üzere KRMD değeri her durumda RMD değerinden büyüktür. Ayrıca KRMD ile kuyruk bölgeleri için daha iyi tahminler elde edilir.



Şekil 3.7. RMD ve KRMD Grafiksel Gösterim [55]

3.4.3. Risk Hesaplamalarının Tutarlılık Kriterleri

Optimal risk hesaplaması için tutarlılık kriterleri dönüştürülebilir değişmezlik, pozitif homojenlik, alt toplamsallık ve monotonluk olarak ifade edilmektedir [53]. Tutarlılık, riske ekonomik olarak rasyonel katkılar sağlayan ve primin etkin dağıtılmasını sağlayan kriterleri ifade etmektedir. Ayrıca düzenleyici ve denetleyici otoritelerin kabul noktasını da belirler. Bu sebeple tutarlılık prensiplerine sahip herhangi bir risk ölçümü uygun ve optimal risk yönetim aracı olarak kabul edilir. Risk yönetim mekanizmasına ilave olarak tutarlı ve rekabetçi primler tüm kriterler sağlandıktan sonra elde edilebilir. Böyle sigorta şirketleri hasara ilişkin risk faktörlerini efektif şekilde yönetebilir.

Dönüştürülebilir Değişmezlik: Eğer $H(X + c) = H(X) + c$ ise ve bu durum tüm $c \geq 0$ ve $X \in \chi$ ise dönüştürülebilir değişmezlik özelliği sağlanmış olur. Diğer bir ifade ile dönüştürülebilir değişmezlik hasarın artış/azalışında riski aynı oranda artırır/düşürür. Bu durum hasar olasılığını verir.

Pozitif Homojenlik: Eđer $H(cX) = cH(X)$ ise ve bu durum tüm $c \geq 0$ ve $X \in \chi$ ise ölçek deęişmezlięi özellięi saęlanmış olur. Artan/azalan hasar oranları riskin de artıp/azaldıęını gösterir. Başka bir ifade ile risk seviyesini ifade etmektedir.

Alt Toplamsallık: Eđer $H(X + Y) \leq H(X) + H(Y)$ ise ve bu durum tüm $X, Y \in \chi$ ise alt toplamsallık özellięi saęlanmış olur. Alt toplamsallık özellięi portföyün çeşitlendirme etkisini yani risk azaltmayı açıklamaktadır. Sigorta portföyü alt parçalara bölünerek risk azaltılamaz ya da poliçeler birleştirildięi zaman risk artmaz.

Monotonluk: Eđer $X \leq Y$ olduęu durumda $X, Y \in \chi$ iken $H(X) \leq H(Y)$ ise bu durum monotonluk özellięi saęlanmış olur. Monotonluk özellięi portföyün çeşitlendirme etkisini yani risk azaltmayı açıklamaktadır. Hasar arttırıldıęı zaman riskinde büyüdüęünü ifade eder. Başka ifade ile monotonluk riskin hasar tutarının büyüklüęünü göstermektedir.

3.5. Genelleştirilmiş Doğrusal Modeller

Genelleştirilmiş doğrusal modeller deęişkenler arasındaki ilişkiyi modelleyen bir metottur. GDM klasik ve normal doğrusal modeli, bazı kısıtlayıcı varsayımları rahatlatarak genelleştirir. GDM'ye ilişkin ilk çalışmalar Nelder ve Weddeburn [8] tarafından yayımlanmıştır. Yayımlandıktan sonra birçok farklı alanda uygulanmakta ve temel istatistikler arasında yer almaktadır [4].

GDM açıklayıcı deęişkenler ve baęımlı deęişken arasındaki ilişkiyi deęerlendirmek ve ölçmek için kullanılmaktadır. GDM, sıradan regresyon modellemesinden iki önemli özellikte ayrılmaktadır. İlk olarak baęımlı deęişkenin daęılımı üstel daęılım ailesinden seçilmektedir. Bu nedenle, baęımlı deęişkenin daęılımı normal daęılıma uymak ya da yakın olmak zorunda deęildir ve hatta açık bir şekilde normal daęılmayabilir. İkinci olarak baęımlı deęişkenin ortalamasının dönüştürülmesi açıklayıcı deęişkenler ile doğrusal olarak saęlanmaktadır. Baęımlı deęişkenin üstel daęılım ailesinin bir üyesi olmasını saęlamanın sonucu olarak, baęımlı deęişken genellikle ve çoęu zaman heteroskedastiktir. Bu sebeple varyans, ortalama ile birlikte deęişim göstermekte ve aynı zamanda açıklayıcı deęişkenlerin deęişimi ile farklılaşmaktadır [4].

GD'ler sigorta verilerinin analizinde oldukça önemli etkiye sahiptir. Sigorta verilerinde normal modelin varsayımları genellikle sağlanamamaktadır. Hasar frekansı, hasar şiddeti ve bir poliçenin hasar getirmesi olasılığına ilişkin değişkenler normal dağılım göstermemektedir [4].

Üstel dağılım ailesinde binom, poisson, negatif binom gibi kesikli dağılım modelleri yer alır. Ayrıca normal, gamma, ters gamma gibi sürekli dağılım modelleri de yer almaktadır. GDM ile normal dağılım birim bağ fonksiyonu, gamma dağılım ters bağ fonksiyonu, poisson dağılım logaritmik bağ fonksiyonu ve binom dağılım ise lojit bağ fonksiyonu ile genelleştirilir ve model oluşturulur.

GDM'nin lineer tahmin edicisi $\eta_i = X_i\beta$ olup bağ fonksiyonu $g(\mu_i) = \eta_i$ şeklinde ifade edilir. Üstel dağılım ailesi için OYF;

$$f(y) = c(y, \theta) \exp\left\{\frac{y\theta - a(\theta)}{\phi}\right\}$$

şeklindedir. θ doğal parameter, ϕ yayılım parametresidir. θ y'nin değerlerine bağlı olup, ϕ ise gözlemler için sabittir. Varyans fonksiyonu ise $E(y) = \mu = \mu(\theta)$ için;

$$f(y; \theta, \phi) = \exp\left\{\int \frac{[y - \mu(\theta)]}{\phi V(\mu)} d\mu(\theta) + c(y, \phi)\right\}$$

olarak ifade edilir. Doğrusal bileşen ise $\eta_i = \sum_{j=1}^p x_{ij}\beta_{ij} = X_i'\beta_i$ olup $i=1, \dots, n$ şeklindedir. Bağ fonksiyonu aracılığı ile ortalama ile doğrusallık ilişkilendirilir. $g(\mu_i) = \mu_i^p$ olan bağ fonksiyonunda $p = -1$ ise ters bağ fonksiyonu gamma dağılımında $g(\mu_i) = \frac{1}{\mu_i^2}$ şeklindedir.

Sigorta verilerinde hasar sayısının modellenmesinde poisson dağılım yaygın şekilde kullanılmaktadır. Sektörde log-normal, gamma ve ters gamma gibi dağılımlardan hasar tutarının modellenmesinde en yaygın kullanılan dağılımlardan bir tanesi gammadır.

Hasara ilişkin tutarın ortalamasının hesaplanmasında genel olarak gamma GDM kullanılmaktadır. Gamma GDM’de varyans ortalamasının bir fonksiyonu olup ortalamaya göre değişkenlik göstermektedir. Üstel dağılım ailesinin üyesi olan Gamma GDM’de ters bağ fonksiyonu kullanılmaktadır [9]. Sürekli veriler için kullanılan GDM’lerden olan Gamma GDM sigorta tutarına ilişkin hesaplamalarda ayrıca pozitif ve sürekli değerlerin mevcut olduğu diğer aktüeryal hesaplamalarda kullanılmaktadır.

Hasar frekansı ve hasar tutarına ilişkin gamma regresyon ile hesaplanan hasar tutarı çarpılarak safi prim hesaplanacaktır. Burada safi prim, üzerine risk yükleme faktörü eklenmemiş prim tutarını göstermektedir. Çalışmanın temel amacı LR ve KR ile hesaplanan prim yüklemesinin, klasik regresyon modeli ile kıyaslanması ve kantil regresyon değerlerinin poliçe fiyatlamasında kullanılmasıdır. Hasar tutarının modellenmesine sürekli dağılımlar kullanılmaktadır. Fakat çalışmada özellikle sürekli ve sağa çarpık dağılımlarda kullanılan ve KR ile hesaplanan modele ilişkin karşılaştırma sunan gamma regresyon ile hesaplanan tutarlar üzerinden poliçe risk grubu çerçevesinde risk prim yüklemesi gerçekleştirilecektir. Gamma dağılıma ilişkin açıklama ise bir sonraki bölümde açıklanmıştır.

3.5.1. Gamma Dağılım

Gamma dağılım iki parametresi bulunan şekil parametresi k ve ölçek parametresi θ olan bir dağılımdır. Sağa çarpık ve pozitif değerli değişkenler için yaygın şekilde kullanılmakta olup özellikle sigorta hasar tutarının modellenmesinde sık kullanılan bir dağılımdır.

$$f(x) = f(x; k, \theta) = \frac{\theta^k}{\Gamma(k)} x^{k-1} e^{-\theta x} \quad x, \theta, k > 0$$

Beklenen değer;

$$E(x) = k/\theta$$

Varyans ise;

$$\text{Var}(x) = k/\theta^2$$

şeklinde ifade edilir.

4. İKİ AŞAMALI KANTİL REGRESYON MODELİ TÜRKİYE UYGULAMASI

Tezin bu kısmında, Bölüm 2 ve 3’de detaylı olarak açıklanan lojistik regresyon ve iki aşamalı kantil regresyon metodolojileri ile risk faktörlerinin analizi, Türkiye’de faaliyet gösteren bir hayat dışı sigorta şirketi verileri üzerinde uygulanmış ve elde edilen tüm durumlar detaylı olarak açıklanmıştır. Bu bölümde sayısal uygulaması yapılan veri yapısı detaylı olarak açıklanarak kantil regresyonun kullanılma nedenleri veri üzerinden incelenmiştir. Devamında, lojistik regresyon ile hasarın gerçekleşme ve gerçekleşmeme olasılıkları tahmin edilmiş ve bu tahminler üzerinden kantil regresyon değerleri bulunmuştur. Hasar tutarları üzerine risk primi yüklemesi yapılarak, kantil regresyon ile elde edilen prim tutarları, kantil regresyonun kullanılmadığı durum için elde edilen prim tutarları ile karşılaştırılmıştır.

4.1. Veri Yapısı

Tezin uygulama kısmında kullanılan veri, TSB bünyesindeki Sigorta Bilgi ve Gözetim Merkezinden (SBM) tez çalışmasında kullanılmak üzere talep edilerek alınmıştır. Veri, Türkiye’de hayat dışı sigorta branşında aktif olarak faaliyet gösteren bir sigorta şirketinin 2022 yılına ilişkin bilgilerini içermektedir. 225.940 adet bağımsız sigorta poliçesinden oluşan portföyde, en az bir hasar gerçekleşen 36.829 poliçe üzerinde analiz gerçekleştirilmiştir. Analizde R bilgisayar programı ve program dahilindeki quantreg paketi kullanılmıştır [56]. Uygulamaya esas alınan değişkenler ve açıklamaları Çizelge 4.1’de detaylı olarak gösterilmektedir.

Çizelge 4.1. Değişkenler ve Açıklamaları

DEĞİŞKEN	AÇIKLAMA
Hasar	Hasar Gerçekleşme Durumu (0=hasar yok, 1=hasar var)
Hasar Sayısı	Sigortalı Hasar Frekansı
Toplam Hasar	Sigortalı Toplam Hasar Tutarı (Hasar Yok =0)
Sigortalı Basamak Grubu	Sigortalı Basamak Grubu (0-5 Basamak)
Araç İl Grubu	Araçların Bulunduğu İl Bazında Grup (İstanbul-Ankara-İzmir-Diğer)

Veride öncelikle ilk deęişken olarak hasarın gerçekleşip gerçekleşmedięi bilgisini gösteren *Hasar Deęişkeni* incelenmiştir. Bu deęişken hasarın gerçekleşme durumuna göre, toplam hasarın mevcut olması durumunda 1; hasarın gerçekleşmemesi durumunda ise 0 deęerini almaktadır. Hasarın gerçekleştięinin bilindięi durumlar için sigortalının yıl içerisinde gerçekleştirdięi hasar sayıları bilgileri de bulunmaktadır. Hasar tutarı tüm kazalar için toplu bir şekilde yer alan *Toplam Hasar Tutarı* deęişkeni altında gösterilmiş olup hasar mevcut deęil ise hasar tutarı 0 olarak ifade edilmiştir. Veride ayrıca, Kasko poliçelerine ilişkin 0 ile 5 arasında sınıflanan 6 basamak grubunda yer alan *Sigortalı Basamak Grubu* da tanımlanmıştır. Geçmiş yıllarda gerçekleşmiş hasar durumlarına baęlı olarak risklilik durumları göz önünde bulundurularak, prim indirim oranlarına baęlı olarak düzenlenmiş sigortalı basamak grupları ve prim indirim oranları Çizelge 4.2’de gösterilmektedir.

Çizelge 4.2. Sigortalı Basamak Grubu ve İndirim Oranları

Prim İndirim Oranı (%)	Sigortalı Basamak Grubu
0	0
30	1
40	2
50	3
55	4
60	5

Çizelge 4.2’ye göre, sigortalı basamak grubu 0 olan sigortalının şirkete ilk kez kasko sigortası yaptırdığını ya da bir önceki yıldaki gerçekleşen kaza durumuna baęlı olarak bu grupta bulunduęu varsayılmaktadır. Bu risk grubu için prim tutarı en yüksek deęere sahip olup herhangi bir indirim yapılmamaktadır. Sigortalı risk grubunun 1 olduęu durum ise; sigortalının bir önceki yıl kaza yapmadığı ve bir üst gruba yükselerek %30 indirim kazandığını ifade eder. Sigortalı risk grubunun 2 olması ise sigortalının iki yıl boyunca kaza yapmadığı ve bu gruba yükselerek %40 indirim kazandığını ifade etmektedir. Sigortalı risk grubunun 3 olması ise sigortalının üç yıl boyunca kaza yapmadığı ve bu gruba yükselerek %50 indirim kazandığını ifade eder. Risk grubunun 4 olması ise sigortalının dört yıl boyunca kaza yapmadığı ve bu gruba yükselerek %55 indirim kazandığını ifade eder. Son olarak risk grubunun 5 olması ise sigortalının beş yıl boyunca

kaza yapmadığını ve bu gruba yükselerek %60 indirim kazandığını ifade etmektedir. Gerçekleşen kazalarda, sigortalının kusur durumu dikkate alınmakta ve sigortalının kusurunun bulunmadığı durumlarda, sigortalının kaza yapmadığı varsayılmaktadır.

Plaka İl Kodu verilen poliçe üretim ve hasar verileri *Ankara, İstanbul, İzmir* ve bu iller dışındaki şehirler için *Diğer* şeklinde sınıflandırılmış ve yeni sütun olarak veriye eklenmiştir. Çizelge 4.4'te yer alan veriler incelendiğinde üç büyük şehirde gerçekleşen prim üretiminde önemli bir yoğunlaşma olduğu görülmüştür. Bu üç şehir için gerçekleşen prim üretiminin, üç il dışındaki diğer bütün illerin toplamına yakın olması nedeniyle bu şekilde bir gruplama gerçekleştirilmiştir.

Yazılan poliçelere ilişkin hasar frekans oranları Çizelge 4.3'te gösterilmektedir. Bu portföyde hasar frekansı en fazla 6 olacak şekilde düzenleme yapılmıştır. Veriye ilişkin incelemede %83,70'lik oran ile poliçelerin büyük çoğunluğunun 0 hasar frekansına sahip olduğu gözlenmiştir. Hasar frekansı arttıkça, gözlenen poliçe sayısı ve hasar oranında azalma olduğu da gözlenmiştir.

Çizelge 4.3. Hasar Frekansı (Toplam)

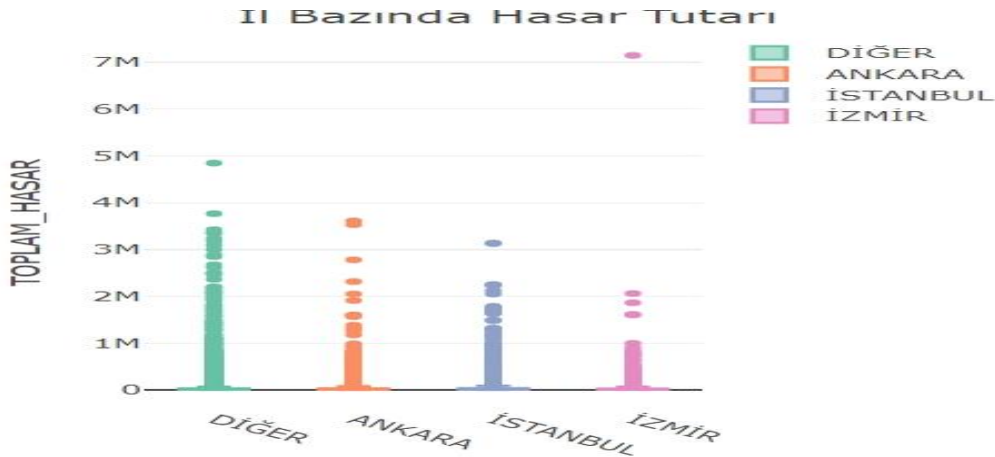
Hasar Sayısı	Hasar Frekansı						
	0	1	2	3	4	5	6
Poliçe Sayısı	189.111	24.532	7.601	2.598	1.147	603	348
Yüzde (%) (0 Hasar Dahil)	83,70	10,86	3,36	1,15	0,51	0,27	0,15
Yüzde (%) (0 Hasar Hariç)	-	66,61	20,64	7,05	3,11	1,64	0,94

Hasar frekansına ilişkin Çizelge 4.3 ayrıca en az bir hasar frekansı bulunan poliçeler dikkate alınarak incelendiğinde hasar frekansı 1 olan poliçelerin %66,61 ile en yüksek poliçe sayısı ve yüzdesine sahip olduğu da gözlenmektedir. Çizelge 4.4'te hasar tutarına ilişkin il ayrımında özet istatistikler gösterilmektedir. Çizelge 4.4'e göre hasar getiren sigortalı sayısı en yüksek *Diğer* olarak adlandırılan sınıfta gözlenmiştir. Hasar tutarına ilişkin standart sapmanın (σ) en fazla 171.531 ile İzmir'de olduğu gözlenmektedir. Bu durum için, İzmir ilinde gözlenen maksimum hasarın (Max) 7.149.000 TL olmasının etkisi ile olabileceği düşünülmektedir.

Çizelge 4.4. Hasar Tutarına İlişkin Özet İstatistikler (İl Ayrımında)

	İSTANBUL	ANKARA	İZMİR	DİĞER	TOPLAM
n	8.126	4.499	2.596	21.608	36.829
μ	48.675	45.644	39.763	50.721	48.877
σ	133.133	140.146	171.531	159.450	152.669,9
γ	2,74	3,07	4,31	3,14	3,12
Σ	395.531.079	205.352.611	103.226.109	1.095.987.407	1.800.097.206
Min	236	204	294	204	204
Max	3.139.056	3.611.590	7.149.000	4.847.734	7.149.000
Q_1	6.049	5.648	5.455	5.040	5.316
Q_2	15.425	13.241	12.046	10.968	12.160
Q_3	39.633	36.399	31.725	32.584	34.874
m_3	10	13	30	10	13
m_4	153	262	1.157	167	288

Verilere ilişkin üçüncü çeyrek (Q_3) değerleri incelendiğinde İstanbul için 39.633 TL olan riskin diğer şehirlere kıyasla daha yüksek olduğu gözlenmektedir. Çarpıklık katsayısının (skewness, m_3) ise bütün şehir kırılımlarında pozitif olduğu, bu nedenle %95 güven düzeyinde normal dağılıma göre daha sağa çarpık bir dağılım beklenmektedir. Ayrıca basıklık (kurtosis, m_4) değerine bağlı olarak bütün şehir kırılımlarında hasara ilişkin dağılımların normal dağılıma göre daha sivri bir dağılıma uygunluk gösterdiği %95 güven düzeyinde söylenebilmektedir. Hasar tutarlarına ilişkin şehir ayrımında yer alan grafikler ise Şekil 4.1’de gösterilmiştir.



Şekil 4.1. Hasar Tutarları (İl Değişkeni Bazında)

Şekil 4.1 incelendiğinde il kırılımında en yüksek hasarın 7.149.000 TL ile İzmir’de ve 3.139.056 TL ile en düşük İstanbul’da olduğu gösterilmiştir. Grafik ve betimleyici istatistiklerden hasara ilişkin dağılımın hem il bazında ve hem de toplam durum incelemelerinde, kalın kuyruklu ve sağa çarpık bir dağılıma uygunluk gösterdiği gözlemlenmektedir. Çalışmada poliçelere ilişkin hasar getiren veriler üzerinden kutu grafiği gösterilmiştir. 0 hasar getiren poliçeler ise veriye ilişkin İAKR modelinde LR ile modellendiği için kutu grafiğine dahil edilmemiştir.

Çizelge 4.5’te ise sigortalı basamak grubu sınıflamasında 0’dan 5’e kadar gruplanmış basamaklarda yer alan toplam hasar tutarlarına ilişkin betimleyici istatistikler gösterilmektedir.

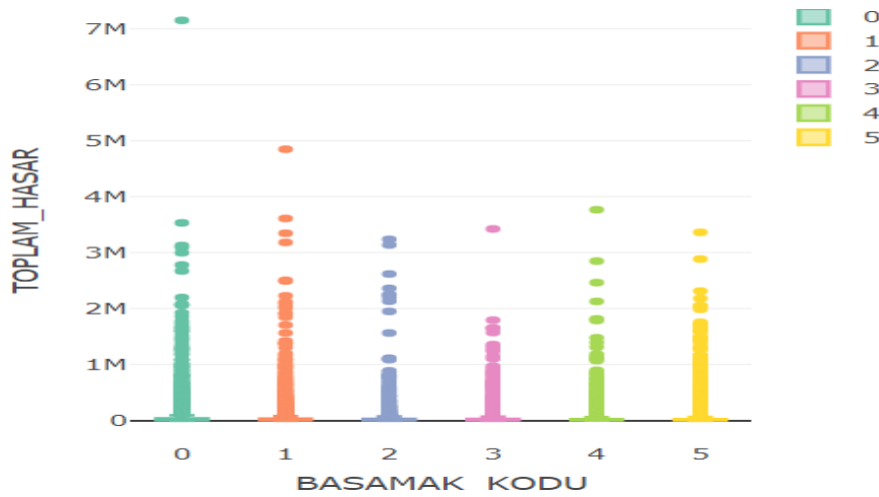
Çizelge 4.5. Hasar Tutarı Betimleyici İstatistikler (Sigortalı Basamak Grubu)

	Basamak 0	Basamak 1	Basamak 2	Basamak 3	Basamak 4	Basamak 5
n	9.056	6.259	4.120	3.018	2.941	11.435
μ	57.453	53.643	46.800	44.947	44.639	42.353
σ	176.601	170.904	148.681	133.312	154.435	124.418
γ	3,07	3,19	3,18	2,97	3,46	2,94
Σ	520.291.547	335.752.918	192.813.957	135.650.379	131.284.654	484.303.750
Min	295	236	204	247	348	204
Max	7.149.000	4.847.734	3.238.420	3.424.751	3.767.677	3.364.124
Q_1	5.224	5.409	5.602	5.548	5.564	5.183
Q_2	12.968	13.169	12.057	13.159	11.446	11.253
Q_3	42.398	36.915	34.825	33.810	29.862	30.577
m_3	14	12	12	11	12	10
m_4	371	216	191	188	209	151

Çizelge 4.5’e göre en yüksek sigortalı sayısı, 5’inci sigortalı basamak grubunda bulunmaktadır. Bu durum 5 yıl boyunca kazası bulunmayan araçları kapsamaktadır. Hasar ortalaması (μ), 0’ıncı basamak için 57.453 TL iken riski en düşük olan 5’inci basamakta 42.353 TL olarak gözlenmektedir. Standart sapma (σ) değeri ise risk seviyesi

yüksek olan 0'ıncı basamakta 176.601 iken diđer basamaklarda azalarak 5'inci basamakta 124.418 ile en düşük değere inmektedir. Ayrıca basamak gruplamasına göre en yüksek hasar, 7.149.000 TL hasar tutarı ile 0'ıncı basamakta gözlenmiştir. Üçüncü çeyrek değerinde de benzer şekilde 0'ıncı basamakta 42.398 olarak gözlenen hasar tutarı beşinci basamağa giderken azalmakta ve 30.577 TL ile beşinci basamakta en düşük değere ulaşmaktadır.

Çarpıklık ve basıklık katsayıları ise bütün basamak gruplarında benzer şekilde pozitif değerler almaktadır. Bu da tüm basamak gruplarına ilişkin hasar tutarlarının normal dağılıma göre %95 güven düzeyi ile daha sağa çarpık ve daha sivri olduğunu göstermektedir. Basamaklara ilişkin kırılımdaki hasar tutarlarına ilişkin grafikler ise Şekil 4.2'de sunulmaktadır.

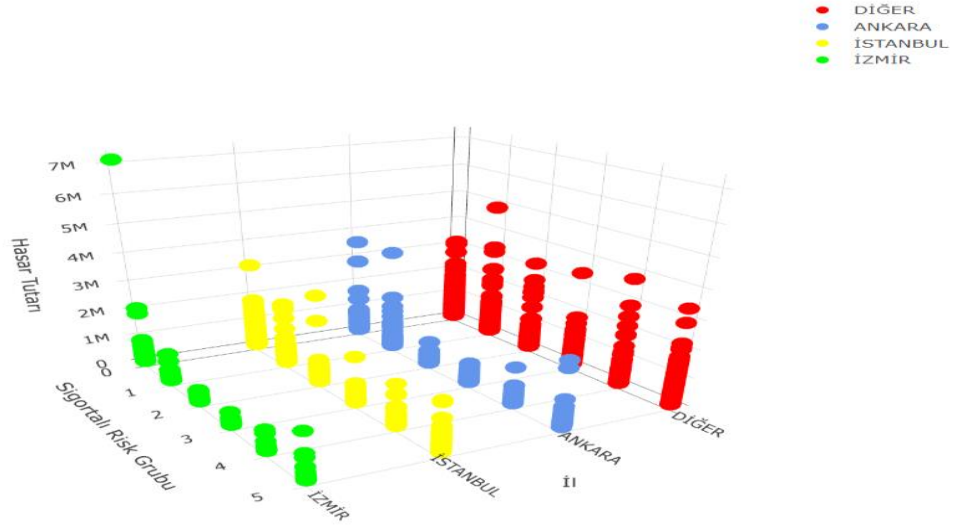


Şekil 4.2. Hasar Tutarlarına İlişkin Grafik (Sigortalı Basamak Grubu)

Şekil 4.2 incelendiğinde maksimum hasarın 7.149.000 TL ile 0'ıncı basamakta ve 3.364.024 TL ile en düşük 5'inci basamakta olduğu görülmektedir.

Şekil 4.3'te ise Sigortalı Basamak Grubu ve Poliçe İl grubu ayrımında poliçelere ilişkin hasar tutarları gösterilmiştir. Bu grafik ile poliçe bazında toplam hasar büyüklüğünün il ve sigortalı basamak grubu kesişiminde gösterimi özetlenmeye çalışılmıştır. Analizde kullanılacak risk grupları il ve sigortalı basamak grubu bilgisi kullanılarak toplam 24 risk grubu oluşturulacaktır. Örneğin İstanbul ve 0 Sigortalı Basamak Grubunda bulunan

poliçe risk grubu, Ankara ve 2 Sigortalı Basamak Grubunda bulunan poliçe risk grubu vb.



Şekil 4.3. Sigortalı Risk Grubu ve İl Bazında Hasar Tutarı

Şekil 4.3'te görüldüğü üzere en yüksek hasar 7.149.000 TL ile İzmir ili ve 0 Sigortalı Basamak Grubunda yer almaktadır. İkinci en yüksek hasar ise 4.847.734 TL ile diğer 1 Sigortalı Basamak Grubunda bulunmaktadır. İzmir ve diğer şehirlerde ortalama hasarın Ankara ve İstanbul'a göre daha yüksek olduğu tespit edilmiştir. 0 Sigortalı Basamak Grubunda yer alan poliçelerin diğer risk gruplarına kıyasla daha yüksek ortalama hasara sahip olduğu da belirlenen diğer bir husustur.

Veriler üzerinden çalışılan basamak grubu ve il değişkeni ise ayrı şekilde gruplar bazında incelenmiştir. Lojistik regresyon için bu iki risk grubu birleştirilerek İAKR modeli uygulanmıştır.

4.2. Lojistik ve Kantil Regresyon Uygulaması

4.2.1. Lojistik Regresyon Uygulaması

Lojistik regresyonun bağımlı değişkeni poliçe üzerinde hasarın gerçekleşme ve gerçekleşmeme durumlarına ilişkin (Hasar var ise 1, yok ise 0) sonucu vermektedir. Burada sigortalı basamak grubu 0 ile 5 arasında değerlerden oluşan 6 gruba ayrılmıştır. Poliçe il grubu bazında ise İstanbul, Ankara, İzmir ve diğer iller için diğer olarak 4 gruba ayrılmıştır. Lojistik regresyona ilişkin eşitlik aşağıda gösterilmektedir:

$$\ln\left(\frac{1-p_i}{p_i}\right) = \beta_0 + \beta_1(\text{Basamak Kodu}) + \beta_2(\text{il grubu}) \quad (4.1)$$

Eşitlik (4.1)'de $(1 - p_i)$, i'nci poliçe grubunun hasar getirmeme olasılığını göstermektedir. Basamak grubu 3 ve 4 dışında yer alan risk gruplarında parametrelerin önemli derecede etkiye sahip olduğu gözlenmiştir. Lojistik regresyon ile 0 hasarın gerçekleşme durumuna ilişkin hesaplanan sonuçlar Çizelge 4.6'da gösterilmektedir.

Çizelge 4.6. LR ile Poliçe Risk Grubuna Göre Hasar Oluşma Olasılığı Tahmini

	Tahmin	Std. Hata	Z Değeri	Pr (> Z)	
Sabit Terim	-1,9249	0,0123	-157,1390	<2 e-16	***
Basamak 1	0,1838	0,0181	10,1300	<2 e-16	***
Basamak 2	0,0860	0,0207	4,1550	3.25 e-05	***
Basamak 3	-0,0097	0,0231	-0,4210	0,6730	
Basamak 4	-0,0306	0,0233	-1,3150	0,1880	
Basamak 5	0,1330	0,0154	8,6130	<2 e-16	***
ANKARA	0,5653	0,0184	30,7320	<2 e-16	***
İSTANBUL	0,7557	0,0148	51,0550	<2 e-16	***
İZMİR	0,3626	0,0230	15,8010	<2 e-16	***

Anlamlılık. kodları: 0 '***' 0.001 '**' 0.01 '*'

Çizelge 4.7'de ise il ve basamak gruplarının birlikte değerlendirilmesi ile oluşturulan 24 tarife grubu için lojistik regresyon aracılığı ile belirlenen 0 hasarın gerçekleşmesi olasılığını veren model tahminleri gösterilmektedir. Çizelge 4.7'de yer alan tahminler Çizelge 4.6'da yer alan LR tahminleri ile elde edilmiştir.

Çizelge 4.7. LR ile Elde Edilen Olasılıklar (Grupların Sıfır Hasar Olasılıkları)

Grup Numarası	Risk Grup (Basamak& İl)	Police Sayısı	Hasar Sayısı	0 Hasar Olasılığı
1	(1&İs)	5.091	1.440	0,6267
2	(5&İs)	8.417	2.272	0,6452
3	(2&İs)	4.014	1.007	0,6615
4	(0&İs)	8.052	1.886	0,6894
5	(1&A)	3.254	749	0,6914
6	(3&İs)	3.514	778	0,6924
7	(4&İs)	3.403	743	0,6987
8	(5&A)	6.008	1.444	0,7067
9	(2&A)	2.441	474	0,7202
10	(0&A)	4.906	1.035	0,7432
11	(3&A)	2.041	411	0,7457
12	(1&İz)	2.186	392	0,7480
13	(4&A)	2.104	386	0,7510
14	(5&İz)	4.409	975	0,7605
15	(2&İz)	1.508	267	0,7715
16	(0&İz)	3.434	572	0,7903
17	(4&İz)	1.265	184	0,7967
18	(3&İz)	1.249	206	0,8097
19	(1&D)	24.588	3.678	0,8247
20	(5&D)	49.121	6.744	0,8334
21	(2&D)	16.699	2.372	0,8410
22	(0&D)	43.594	5.563	0,8541
23	(3&D)	12.380	1.623	0,8555
24	(4&D)	12.262	1.628	0,8585

Çizelge 4.8’de ise θ_i^* için olasılık değerlerini her bir grup için tahmin edilen lojistik regresyon aracılığı ile 0 hasar getirme olasılıkları üzerinden modifiye edilen KR parametreleri tahmin edilmiştir.

Çizelge 4.8. $\theta=0,95$ Kantil Değeri ile Revize Hasar Oluşturmama Olasılıkları

Grup	Risk Grup(Basamak & İl)	Police Sayısı	Hasar Sayısı	$\theta^*_{0,95}$
1	(1&İs)	5.091	1.440	0,8661
2	(5&İs)	8.417	2.272	0,8591
3	(2&İs)	4.014	1.007	0,8523
4	(0&İs)	8.052	1.886	0,8390
5	(1&A)	3.254	749	0,8380
6	(3&İs)	3.514	778	0,8375
7	(4&İs)	3.403	743	0,8340
8	(5&A)	6.008	1.444	0,8295
9	(2&A)	2.441	474	0,8213
10	(0&A)	4.906	1.035	0,8053
11	(3&A)	2.041	411	0,8034
12	(1&İz)	2.186	392	0,8016
13	(4&A)	2.104	386	0,7992
14	(5&İz)	4.409	975	0,7912
15	(0&İz)	3.434	572	0,7812
16	(2&İz)	1.508	267	0,7615
17	(4&İz)	1.265	184	0,7541
18	(3&İz)	1.249	206	0,7372
19	(1&D)	24.588	3.678	0,7148
20	(5&D)	49.121	6.744	0,7000
21	(2&D)	16.699	2.372	0,6855
22	(0&D)	43.594	5.563	0,6573
23	(3&D)	12.380	1.623	0,6539
24	(4&D)	12.262	1.628	0,6466

Birinci aşamada hesaplanan en az bir hasar oluşma olasılığı üzerinden her bir poliçe grubu için hasar oluşma/oluşmama olasılıkları tahmin edilmiştir. Eşitlik 4.3'te yer alan formül ile %95 kantil değeri ile modifiye edilmiş ve hasar oluşmama olasılıkları üzerinden KR

tutarlarının tahmininde kullanılan kantil değerleri her bir poliçe risk grubu için Çizelge 4.8’de gösterilmektedir. Hasar getirme olasılığı yüksek olan değerler için KR’de kullanılan ve riski belirleyen kantil değerlerinin daha yüksek olduğu ve bu değerlerin risk azaldıkça 0,6466’ya düştüğü Çizelge 4.8’de gözlenmektedir.

4.2.2. Kantil Regresyon

Kantil regresyonda bağımlı değişken, poliçe bazında toplam hasar tutarını göstermektedir. Bu yüzden poliçelere ilişkin veriler hasar getiren poliçeler bazında seçilmiştir. Bu sınıflamaya göre hasar getirme olasılığı dikkate alınarak ayrılan 24 grup için kantil değerleri hesaplanmıştır. Kantil regresyona ilişkin model ise Eşitlik (4.2)’de gösterilmektedir.

$$Q_{S_i|N_i>0}(\theta_i^* I x_i) = \exp(\gamma_{\theta_i^*,0} + \gamma_{\theta_i^*,1}(\text{Basamak Kodu}) + \gamma_{\theta_i^*,2}(\text{il grubu})) \quad (4.2)$$

$$S_i = \left(1 - \frac{\exp(\beta_0 + \beta_1(\text{Bas.}) + \beta_2(\text{il}))}{1 + \exp(\beta_0 + \beta_1(\text{Bas.}) + \beta_2(\text{il}))}\right) * \exp(\gamma_{\theta_i^*,0} + \gamma_{\theta_i^*,1}(\text{Bas.}) + \gamma_{\theta_i^*,2}(\text{il}))$$

Eşitlik (4.2)’de S_i tepki değişkeni olup, i ’nci poliçe sahibinin toplam hasar tutarını göstermektedir. Analizde, $\theta = 0,95$ olasılık değeri için KR tahmin değerleri elde edilmiştir. Kantil regresyon parametreleri aşağıda yer alan formül ile tahmin edilmiştir.

$$\theta_i^* = (\theta - p_i)/(1 - p_i) \quad (4.3)$$

Çizelge 4.9 ise kantil değeri $\theta = 0,95$ olan kantil regresyona ilişkin parametre tahminlerini göstermektedir. Her sütun risk grubunda yer alan (1’ den 24’e) poliçeler için model tahminlerini ifade etmektedir. Burada her bir risk grubu lojistik regresyon ile 0 hasar olasılığına göre belirlenen grupları ifade etmektedir. Lojistik regresyon modelinde iki grup dışında diğer gruplar önemli iken; bu gruplardan bazılarının ikinci aşama olan kantil regresyon modelinde önemli etkiye sahip olmadıkları belirlenmiştir. Bu açıdan sigorta şirketi en tehlikeli poliçe grubunu belirlemede iki aşamada da önemli etkiye sahip olan poliçe grubunu kullanmalıdır.

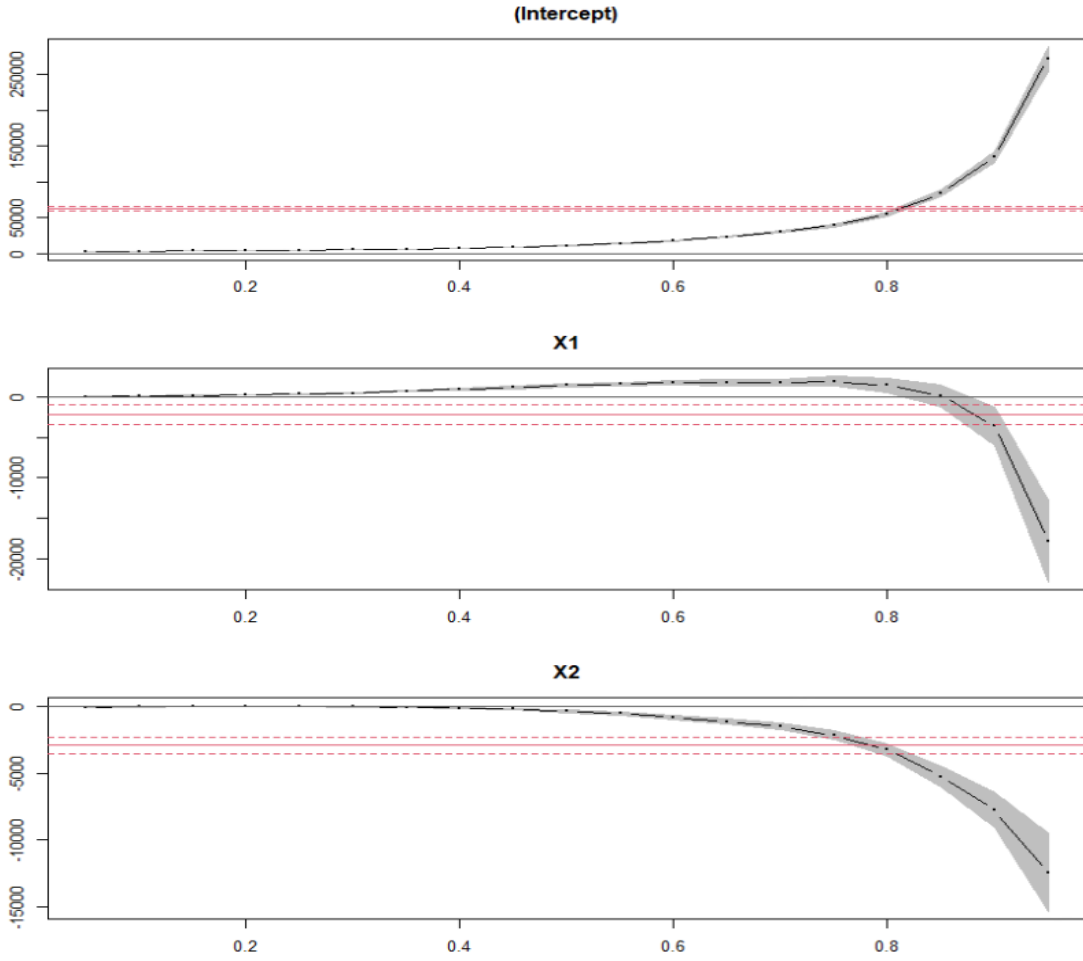
LR sonucunda hasar getirme olasılığı önemli olan risk grubu eğer ikinci aşama kantil regresyon sonucunda önemsiz ise bu grup daha az riskli olabileceğini göstermektedir. Hem LR hem de KR sonucunda önemli etkiye sahip olan gruplar ise diğer gruplara kıyasla daha riskli poliçe gruplarını içerebilecektir. Basamak 1, basamak 2 ve İstanbul

grubunda yer alan poliçeler ise LR ve KR açısından önemli etkiye sahip olup diğer gruplara oranla daha riskli poliçe gruplarını gösterebilecektir.

Çizelge 4.9. $\theta=0,95$ kantil Değeri İçin Kantil Regresyon Parametre Tahminleri

Model	1	2	3	4	5	6	7	8	9	10	11	12
Sabit Terim	11,438 ***	11,378 ***	11,315 ***	11,215 ***	11,315 ***	11,206 ***	11,183 ***	11,143 ***	11,068 ***	10,956 ***	10,940 ***	10,927 ***
Basamak1	-0,118 **	-0,130 **	-0,116 (.)	-0,130 *	-0,135 *	-0,139 *	-0,145 *	-0,154 *	-0,145 *	-0,146 **	-0,148 **	-0,143 **
Basamak2	-0,320 ***	-0,324 ***	-0,316 ***	-0,323 ***	-0,320 ***	-0,317 ***	-0,341 ***	-0,323 ***	-0,303 ***	-0,280 ***	-0,271 ***	-0,263 ***
Basamak3	-0,336 ***	-0,352 ***	-0,366 ***	-0,350 ***	-0,345 ***	-0,347 ***	-0,339 ***	-0,323 ***	-0,304 ***	-0,284 ***	-0,288 ***	-0,281 ***
Basamak4	-0,464 ***	-0,488 ***	-0,480 ***	-0,478 ***	-0,474 ***	-0,473 ***	-0,474 ***	-0,456 ***	-0,424 ***	-0,424 ***	-0,426 ***	-0,423 ***
Basamak5	-0,387 ***	-0,403 ***	-0,388 ***	-0,391 ***	-0,392 ***	-0,394 ***	-0,395 ***	-0,391 ***	-0,370 ***	-0,363 ***	-0,362 ***	-0,362 ***
Ankara	0,050	0,080	0,082	0,082 (.)	0,084	0,085 (.)	0,084	0,081 (.)	0,089 (.)	0,093 (.)	0,100 *	0,107 *
İstanbul	0,055 (.)	0,078 (.)	0,092 *	0,097 *	0,096 *	0,098 *	0,103 *	0,110 **	0,122 **	0,139 ***	0,144 ***	0,148 ***
İzmir	-0,137 **	-0,109 *	-0,087	-0,102 (.)	-0,104 (.)	-0,105 (.)	-0,100 (.)	-0,101 (.)	-0,080	-0,049	-0,051	-0,051
Model	13	14	15	16	17	18	19	20	21	22	23	24
Sabit Terim	10,915 ***	10,867 ***	10,803 ***	10,674 ***	10,626 ***	10,517 ***	10,394 ***	10,316 ***	10,245 ***	10,105 ***	10,088 ***	10,057 ***
Basamak1	-0,159 **	-0,174 **	-0,163 **	-0,156 **	-0,151 **	-0,144 **	-0,124 **	-0,114 **	-0,110 **	-0,085 *	-0,084 (.)	-0,078 (.)
Basamak2	-0,264 ***	-0,261 ***	-0,250 ***	-0,225 ***	-0,218 ***	-0,196 ***	-0,177 **	-0,172 **	-0,169 **	-0,182 **	-0,183 **	-0,177 **
Basamak3	-0,289 ***	-0,292 ***	-0,286 ***	-0,269 ***	-0,270 ***	-0,217 ***	-0,192 ***	-0,177 **	-0,174 **	-0,140 **	-0,135 **	-0,131 *
Basamak4	-0,427 ***	-0,438 ***	-0,431 ***	-0,391 ***	-0,383 ***	-0,337 ***	-0,305 ***	-0,291 ***	-0,273 ***	-0,258 ***	-0,257 ***	-0,269 ***
Basamak5	-0,367 ***	-0,364 ***	-0,361 ***	-0,350 ***	-0,338 ***	-0,321 ***	-0,293 ***	-0,280 ***	-0,275 ***	-0,268 ***	-0,265 ***	-0,260 ***
Ankara	0,106 *	0,107 *	0,117 *	0,138 **	0,138 **	0,139 **	0,145 ***	0,155 ***	0,171 ***	0,172 ***	0,175 ***	0,168 ***
İstanbul	0,149 ***	0,161 ***	0,169 ***	0,187 ***	0,196 ***	0,216 ***	0,219 ***	0,221 ***	0,235 ***	0,262 ***	0,265 ***	0,269 ***
İzmir	-0,048	-0,059	-0,042	-0,048	-0,038	-0,009	0,017	-0,005	0,003	0,034	0,035	0,038

Not: Anlamlılık. kodları: 0 '****' 0.001 '***' 0.01 '**' 0.05 '.' 0.1 ' ' 1.



Şekil 4.4. Kantil Regresyon Grafik Karşılaştırmaları

Şekil 4.4'te kesişim bölümü bağımlı değişken olan toplam hasarı göstermektedir. Bağımsız değişkenlerden X1 il grubu bazında sınıflamayı gösterirken, X2 ise sigortalı basamak grubu bazında sınıflamayı göstermektedir. Kırmızı çizgi EKK ile hesaplanan ortalama hasar tutarını göstermektedir. Kesikli kırmızı çizgiler ise EKK'nın güven aralığını göstermektedir. Eğer bağımsız değişkenler kırmızı kesikli çizgilerin arasında yer almış olsaydı bu durumda KR'ye ilişkin değerlerin önemli bir etkisinin olmadığını gösterebilirdi. Şekil 4.4'te yer alan en alt grafikte görüldüğü üzere 0.8 kantil değerine kadar veriler anlamlı sonuçlar ortaya çıkarmaktadır. 0,8 kantil değerinde EKK ile aynı değer ya da güven aralığında yer almakta olup önemli bir fark ortaya çıkarmamaktadır. 0,9 ve üzeri kantil değerlerde ise EKK'ya göre önemli bir etki bulunmaktadır. Bu durum ise uç değerlerin bulunduğunu ve kantil hesaplamaların riskin belirlenmesi açısından önemli olduğunu gösterebilir. Çizelge 4.10'da ise her bir risk grubu için kantil değerlerinin tahmini, lojistik regresyon aracılığı ile tahmin edilen hasar oluşmama

olasılıklarının %95 kantil değeri ile modifiye edilmesi sonucu oluşan değerler ile tahmin edilmiştir (LR+KR).

Çizelge 4.10. $\theta=0,95$ Kantil Risk Grupları için Toplam Hasar Tahminleri

	Grup	Poliçe Sayısı	Hasar Sayısı	0 Hasar Olasılığı	LR+KR Hasar Tutarı ($\theta = 0.95$)
1	(1&İs)	5.091	1.440	0,6267	87.141,63
2	(5&İs)	8.417	2.272	0,6452	63.133,70
3	(2&İs)	4.014	1.007	0,6615	59.849,60
4	(0&İs)	8.052	1.886	0,6894	81.819,43
5	(1&A)	3.254	749	0,6914	77.945,10
6	(3&İs)	3.514	778	0,6924	57.338,07
7	(4&İs)	3.403	743	0,6987	49.609,12
8	(5&A)	6.008	1.444	0,7067	50.641,16
9	(2&A)	2.441	474	0,7202	51.701,91
10	(0&A)	4.906	1.035	0,7432	62.899,91
11	(3&A)	2.041	411	0,7457	46.749,55
12	(1&İz)	2.186	392	0,7480	45.847,69
13	(4&A)	2.104	386	0,7510	39.912,31
14	(5&İz)	4.409	975	0,7605	34.360,74
15	(2&İz)	1.508	267	0,7715	36.718,28
16	(0&İz)	3.434	572	0,7903	41.226,64
17	(4&İz)	1.265	184	0,7967	27.025,33
18	(3&İz)	1.249	206	0,8097	29.467,40
19	(1&D)	24.588	3.678	0,8247	28.843,50
20	(5&D)	49.121	6.744	0,8334	22.833,64
21	(2&D)	16.699	2.372	0,8410	23.755,99
22	(0&D)	43.594	5.563	0,8541	24.468,45
23	(3&D)	12.380	1.623	0,8555	21.008,03
24	(4&D)	12.262	1.628	0,8585	17.814,00

Çizelge 4.10’da yer alan son sütunda lojistik regresyon ile kantil regresyonun birleşimine ilişkin tahminler yer almaktadır. Burada yer alan sonuçlara göre model hasar getirme olasılığının ne kadar iyi tahmin edildiğine bağlı olarak sonucu etkilemektedir.

4.3. Risk Yüklemesi

Aktüerler genellikle sigortalanan riske ilişkin prim hesabı yaparken, toplam hasar beklentisine belirli bir risk veya güvenlik yüklemesi eklemektedir. Beklenen değer prim hesaplama yöntemine göre, risk yükleme tutarı net hasar tutarının belirli bir oranı olmak üzere, P_i ;

$$P_i = E(S_i) + \alpha E(S_i) = (1 + \alpha)E(S_i) \quad (4.4)$$

eşitliği ile elde edilmektedir. Eşitlik (4.4)’de kullanılan α yükleme faktörünü göstermektedir ve $0 < \alpha < 1$ koşulunu sağlamaktadır. Bu risk yükleme faktörü oldukça basit bir yöntem olup yaygın bir şekilde kullanılmaktadır [54].

Beklenen değer prim hesaplama yöntemi, hasar tutarının varyans ya da standart sapmasını içermediğinden özellikle portföydeki poliçelerin heterojen yapıda olması ve beklenen hasar tutarlarında yüksek farklılıkların beklenmesi durumunda doğru ve yeterli bir prim hesabına imkân veremeyecektir. Bu durumda daha risk odaklı olan hasarın standart sapma ve varyansını dikkate alan varyans ve standart sapma prim hesaplama prensipleri kullanılmaktadır. Varyans prim prensibi aşağıdaki eşitlikte gösterilmektedir.

$$P_i = E(S_i) + \alpha Var(S_i) \quad (4.5)$$

Ayrıca standart sapma prim prensibi ise

$$P_i = E(S_i) + \alpha \sqrt{Var(S_i)} \quad (4.6)$$

şeklindedir. Bu iki prim prensibi daha iyi teorik özelliklere sahiptir. Fakat bu yöntemlerden varyans prim prensibi yalnızca dönüşümsel değişmezlik özelliğini sağlar ve standart sapma prim prensibi monotonluk özelliğini sağlamaz [54]. Bu yöntemlere bir alternatif olarak önceden belirlenmiş toplam hasarın θ' ncu kantiline ilişkin $Q_{S_i}(\theta)$ ve beklenen değere $E(S_i)$ arasındaki fark kullanılabilir. Sonuç olarak nihai prim $Q_{S_i}(\theta)$ ve $E(S_i)$ konvex birleşimi olarak aşağıdaki eşitlik ile hesaplanabilir.

$$P_i = E(S_i) + \alpha(Q_{S_i}(\theta) - E(S_i)) = \alpha Q_{S_i}(\theta) + (1 - \alpha)E(S_i) \quad (4.7)$$

Eşitlik (4.7)'de basit bir şekilde gösterildiği üzere kantil değerlere ilişkin düzensiz risk yüklemesi olmaması, dönüştürülebilir değişmezlik ve ölçek değişmezliği gibi bütün özellikleri içermektedir [13]. Bu yöntem Kantil Prim Prensibi olarak ifade edilebilir.

Çizelge 4.11'de ise hasar tutarı için GDM ile elde edilen prim tutarları gösterilmiştir. Bu çalışmada, hasar tutarının gamma dağılıma uyduğu varsayılmıştır. Burada yer alan katsayılar ve hasar frekansı ile Çizelge 4.12'de yer alan safi primler, hasar tutarı ve hasar getirme olasılığı ile çarpılarak hesaplanmıştır.

Çizelge 4.11. GDM (Gamma Hasar Tutarı)

	Tahmin	Std. Hata	t Değeri	Pr (> t)	
Sabit Terim	11,0043	0,0349	315,42	<2e-16	***
Basamak 1	-0,0752	0,0509	-1,48	0,139539	
Basamak 2	-0,2119	0,0582	-3,64	0,00027	***
Basamak 3	-0,2518	0,0651	-3,87	0,000109	***
Basamak 4	-0,2588	0,0657	-3,94	8,20E-05	***
Basamak 5	-0,3095	0,0436	-7,11	1,21E-12	***
ANKARA	-0,1122	0,0507	-2,21	0,027042	*
İSTANBUL	-0,0544	0,0403	-1,35	0,176934	
İZMİR	-0,2519	0,0643	-3,92	9,01E-05	***

Not: Anlamlılık. kodları: 0 '****' 0.001 '***' 0.01 '**' 0.05 '.' 0.1 '.' 1.

Kantil Prim Prensibi (KPP), Beklenen Değer Prim Prensibi (BDPP) yöntemine göre bazı avantajlara sahiptir. Çizelge 4.12'de 2'inci sütunda safi prim ya da beklenen hasar tutarını, 3'üncü sütunda %95 kantil değeri ile hesaplanan hasar tutarını, 4'üncü sütunda ise hasarın %95 kantil değeri ile beklenen prim tutarının çıkartıldığı ve risk grubu bazında arada kalan (K-SP) tutarları yer almaktadır. Safi primi hesaplamak için iki parçalı model kullanılmaktadır. $N_i \geq 0$ varsayımı altında, $E(S_i) = E(E(S_i|N_i))$ aşağıdaki eşitlik ile elde edilmektedir.

$$= P(N_i = 0)E(S_i|N_i = 0) + P(N_i > 0)E(S_i|N_i > 0) \quad (4.8)$$

$$= P(N_i > 0)E(S_i|N_i > 0)$$

$Prob(N_i > 0) = 1 - p_i$, LR yöntemi ile tahmin edilmiştir. İkinci bölümde ise $E(S_i|N_i > 0)$ durumu gamma regresyon ile tahmin edilmiştir.

Çizelge 4.12 ise hasar tutarlarına risk yüklemesi yapılmadan hesaplanan safi prim, %95 kantil değerine ilişkin kantil prim değerlerini, kantil prim değeri ile safi prim arasındaki farkı gösteren K-SP değerini, beklenen değer prim prensibi ve kantil prim prensibi ile hesaplanan tutarları göstermektedir.

Çizelge 4.12. Safi Prim, Beklenen Hasar, $\theta=0,95$ Kantil ve Safi Prim Arası Fark

Grup	Safi Prim	Kantil 95	K-SP	BDPP	KPP
1	19.717	87.142	67.425	21.215	21.740
2	14.825	63.134	48.309	15.952	16.274
3	15.595	59.850	44.254	16.780	16.923
4	17.687	81.819	64.132	19.031	19.611
5	15.385	77.945	62.561	16.553	17.261
6	13.617	57.338	43.721	14.652	14.929
7	13.243	49.609	36.367	14.249	14.334
8	11.568	50.641	39.074	12.446	12.740
9	12.169	51.702	39.533	13.093	13.355
10	13.801	62.900	49.099	14.849	15.274
11	10.625	46.750	36.125	11.432	11.709
12	10.924	45.848	34.924	11.754	11.971
13	10.333	39.912	29.580	11.118	11.220
14	8.213	34.361	26.147	8.838	8.998
15	10.679	36.718	26.039	11.491	11.460
16	7.928	41.227	33.298	8.531	8.927
17	7.337	27.025	19.689	7.894	7.927
18	6.913	29.467	22.554	7.439	7.590
19	9.779	28.844	19.065	10.522	10.351
20	7.353	22.834	15.481	7.911	7.817
21	7.735	23.756	16.021	8.322	8.215
22	8.772	24.468	15.696	9.439	9.243
23	6.754	21.008	14.255	7.267	7.181
24	6.568	17.814	11.246	7.067	6.905

Çizelge 4.13'te ise Çizelge 4.12'ye ilave olarak BDPP ve KPP ile hesaplanan prim tutarları üzerine BDPP'de kâr olarak gösterilirken %7.56 lık oran ile tüm risk grupları için benzer oran kullanılmıştır.

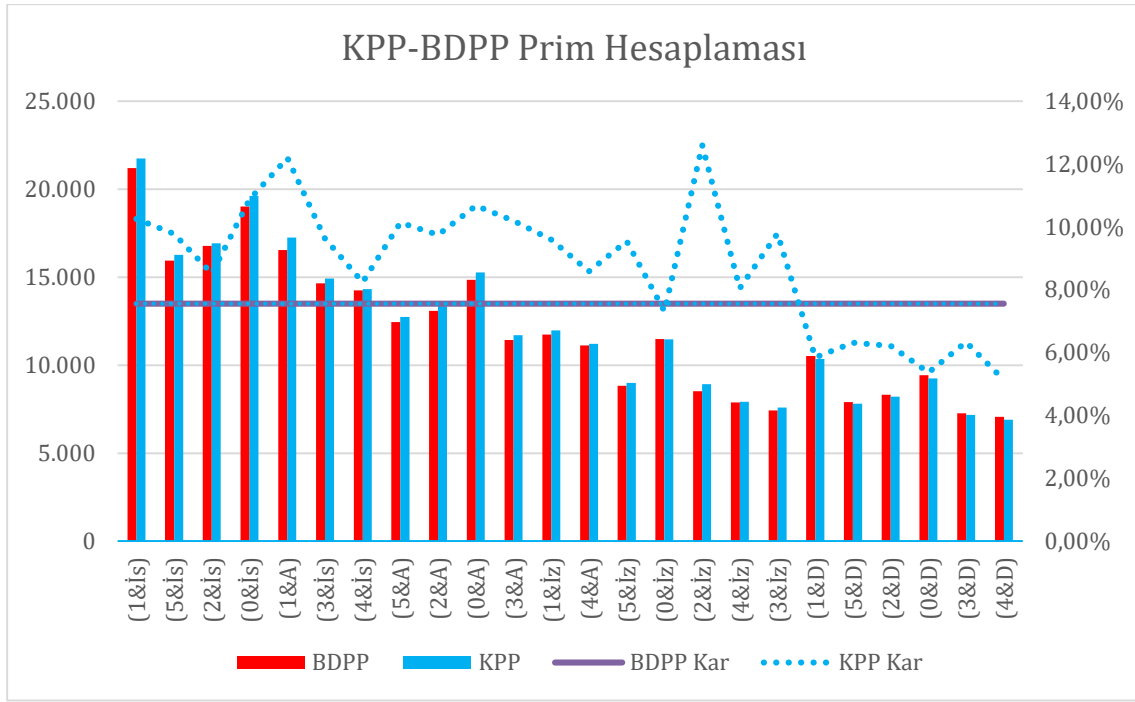
Çizelge 4.13. Safi Prim, BDPP, KPP, BDPP ve KPP Kâr Oranları

Grup	Grup	Safi prim	BDPP	KPP	BDPP Kar	KPP Kar
1	(1&İs)	19.717	21.208	21.740	7,56%	10,26%
2	(5&İs)	14.825	15.946	16.274	7,56%	9,78%
3	(2&İs)	15.595	16.774	16.923	7,56%	8,51%
4	(0&İs)	17.687	19.025	19.611	7,56%	10,88%
5	(1&A)	15.385	16.548	17.261	7,56%	12,20%
6	(3&İs)	13.617	14.647	14.929	7,56%	9,63%
7	(4&İs)	13.243	14.244	14.334	7,56%	8,24%
8	(5&A)	11.568	12.442	12.740	7,56%	10,13%
9	(2&A)	12.169	13.088	13.355	7,56%	9,75%
10	(0&A)	13.801	14.844	15.274	7,56%	10,67%
11	(3&A)	10.625	11.428	11.709	7,56%	10,20%
12	(1&İz)	10.924	11.750	11.971	7,56%	9,59%
13	(4&A)	10.333	11.114	11.220	7,56%	8,59%
14	(5&İz)	8.213	8.834	8.998	7,56%	9,55%
15	(0&İz)	10.679	11.487	11.460	7,56%	7,31%
16	(2&İz)	7.928	8.528	8.927	7,56%	12,60%
17	(4&İz)	7.337	7.891	7.927	7,56%	8,05%
18	(3&İz)	6.913	7.436	7.590	7,56%	9,79%
19	(1&D)	9.779	10.518	10.351	7,56%	5,85%
20	(5&D)	7.353	7.909	7.817	7,56%	6,32%
21	(2&D)	7.735	8.319	8.215	7,56%	6,21%
22	(0&D)	8.772	9.435	9.243	7,56%	5,37%
23	(3&D)	6.754	7.264	7.181	7,56%	6,33%
24	(4&D)	6.568	7.064	6.905	7,56%	5,14%

KPP ile prim yüklemesinde KPP'ye ilişkin risk primlere kantil değerleri ile ilişkili K-SP'nin %3'ü olarak yansıtılmış olup bu durumda poliçe risk grubu bazında farklılaşmalar KPP ile hesaplanan primler Safi Prim üzerine yüklenen oranları poliçe risk grupları bazında farklılaştırmaktadır.

Çizelge 4.13'te ikinci kolonda yer alan BDPP belirli bir yükleme faktörü eklenerek hesaplanmaktadır. KPP'de ise risk yüklemesi (K ile SP) farkının farklı bir yüzdesi eklenerek hesaplanmaktadır. Bu durumda bütün kolonlarda yer alan rakamlar her zaman aynı yönde hareket etmemektedir. Düşük beklenen değerler her zaman düşük kantil değerler ile eşleşmez. Çizelge 4.13'te yer alan 3'üncü ve 5'inci risk grupları için safi prim benzer ve yakın değerler olup 15.595 TL ve 15.385 TL olarak elde edilmiştir. Bu iki risk grubu benzer risk yüklemesine sahiptir. Fakat burada dikkat edilirse %95 kantil değerlerinde Çizelge 4.12'de 5 numaralı risk grubu 77.945 TL olup, 3 numaralı risk grubu ise 59.850 TL ile daha düşük riske sahip olduğu gözlenmiştir. Eğer sigorta fiyatlamasında KPP hesaplaması kullanılırsa karşılaştırmalı riskler sigorta gruplarına ilişkin prim tutarı hesaplamasında sigorta primlerine yansıtılabilir. Bu durumda poliçe grubuna ilişkin fiyatlama 3'üncü Risk Grubu için 16.923 olup, 5'inci Risk grubu için 17.261 olarak farklılaştırılabilir. Başka bir ifade ile KPP yöntemi ile 5'inci grupta yer alan yüksek risk kıyaslanan 3'üncü grupta yer alan düşük riske göre 5'inci grubun fiyatlamasına yansıtılabilir.

Çizelge 4.13'te son iki sütunda yer alan BDPP'ye göre hesaplanan kâr ile KPP'ye göre hesaplanan kâr arasında anlamlı bir fark olup olmadığına ilişkin hipotez testi kurulmuştur. Hipotez testine ilişkin p değeri 0,0084 çıkmış olup %95 güven düzeyinde iki kâr oranı arasında anlamlı bir etkinin olduğu gözlemlenmiştir. KPP ile hesaplanan karın BDPP'ye göre hesaplanan kardan istatistiksel olarak anlamlı bir şekilde farklılaştığı gösterilmiştir.



Şekil 4.5. KPP-BDPP-Risk Prim Yükleme

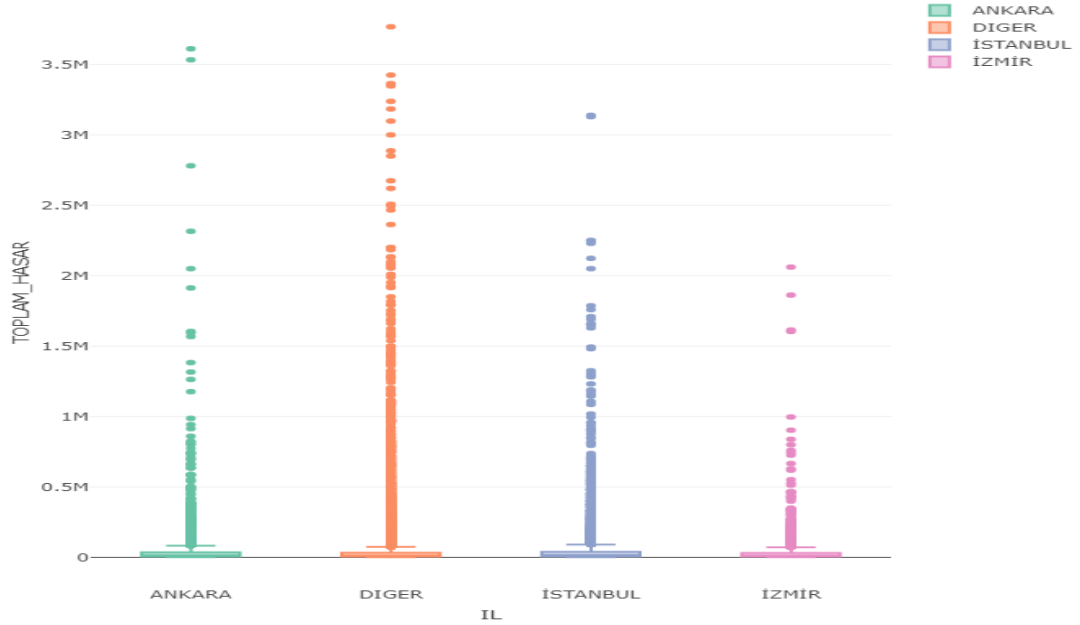
Şekil 4.5'te ise kırmızı bar BDPP ile hesaplanan ve Çizelge 4.13'te yer alan tutarları, mavi bar ile ise Çizelge 4.13'te yer alan KPP ile hesaplanan prim tutarlarını göstermektedir. Şekil 4.5 poliçe risk grubu için tutarları karşılaştırmaktadır. Yüksek risk grubunda KPP ile primlerin arttırılabileceği ve BDPP ile hesaplanan prim tutarından farklı olabileceği görülmektedir. Ayrıca BDPP kâr çizgisinin bütün poliçe risk grupları için benzer yükleme oranını kullandığı fakat KPP kâr ile gösterilen tutarın farklı poliçe risk grupları için farklılaştığı Şekil 4.5'te gözlenmektedir.

Çizelge 4.13'ün son iki kolonunda KPP ve BDPP ile yapılan prim hesaplaması gösterilmektedir. Burada KPP için (K-SP) değerinin %3 risk yükleme ve BDPP için %7,56 risk yükleme faktörleri kullanılmıştır. Fakat %3 lük risk yükleme faktörü kantil değer ile safi prim arasında yer alan farka göre hesaplanmıştır. Bu sayede kantil değeri prim fiyatlamasında poliçe risk grubuna göre farklı oranlarda yansıtılmaktadır. Bu durum ise risk yükleme yapılırken farklı prim tutarlarının farklı risk grupları çerçevesinde belirlenebileceğini göstermektedir. Bu durumda şirketin gelirinin her iki yöntemle hesaplanan prim tutarının benzer olması prensibi ile hareket edilmiştir. Fakat prim tutarları yüksek kantil değere sahip risk grupları için kantil prim prensibi ile farklılaştırılmak sureti ile daha farklı ve riski yansıtacak şekilde hesaplanabilir. Prim tutarları iki yöntemde karşılaştırıldığında yüksek kantil tutarları mevcut olan poliçelerde

BDPP ile karşılaştırıldığında primlerin daha yüksek olduğu, düşük kantil değerlerinde ise KPP hesaplamasının BDPP hesaplamasına göre daha düşük olduğu ve grubun riskinin poliçe fiyatlandırmasına yansıtıldığı gözlenmektedir.

4.3.1. Uç Değersiz Analiz

Analizde ayrıca, veriye ilişkin uç değerlerin modeli ne kadar etkilediği de incelenmiştir. Veri içerisinde bulunan 7.149.000 TL ve 4.847.734 TL'lik olan uç değerler çıkartılarak yapılan analizdeki sonuçlarının farklılığı araştırılmıştır. Şekil 4.6'da ise iki uç değer çıkartılması sonucu uç değersiz hasar tutarına ilişkin grafik yer almaktadır.



Şekil 4.6. Hasar Tutarları (İl Bazında Uç Değersiz)

Yapılan hesaplama sonucunda yüksek kantil tutarına sahip poliçe grubunun primlerinin BDP'ye göre hesaplanan primlere kıyasla daha yüksek çıktığı ve kantil değerine ilişkin riski poliçe grubuna yansıttığı gözlenmiştir. Uç değerlerin modelden çıkartılması ile kantil değerlerin, uç değerlerin çıkartılmadığı modele göre azaldığı gözlenmiştir. Bu durum uç değerlerin modelden çıkartılması durumunda KPP hesaplamasının etkisinin azalacağını göstermektedir.

5. SONUÇ VE ÖNERİ

Kantil regresyon, bir yanıt değişkeninin koşullu miktarlarını tahmin etmek için geleneksel regresyon analizini genişleten istatistiksel bir tekniktir. Aktüerya biliminde kantil regresyon, hasar tutarları, hasar oranları veya diğer risk ölçümleri gibi sigortayla ilgili değişkenlerin dağılımını analiz etmek ve modellemek için kullanılan önemli bir araçtır. Geleneksel regresyon, özellikle aykırı değerlerin bulunması veya dağılımın çarpık olması durumunda doğru bir modelleme imkânı sağlayamamaktadır. Kantil regresyon, aktüerlerin hasar tutarı dağılımının farklı yüzdelik dilimlerini ayrı ayrı modellemesine olanak tanıyarak riskin daha kapsamlı bir şekilde incelenmesine olanak sağlamaktadır. Sigortada kuyruk riski, sigorta şirketinin mali durumu üzerinde önemli bir etkiye sahip olabilecek katastrofik risk olarak da nitelendirebileceğim riski içermektedir. Aktüerler, yüzdelik dilimleri tahmin ederek dağılımın kuyruklarına odaklanabilir ve uç değerlerin daha iyi anlaşılmasına ve yönetilmesine yardımcı olabilir. Kantil regresyonun, dağılımsız bir yaklaşım olması ve uç değerlere karşı güçlü bir yapıya sahip olması nedeniyle oldukça ilgi çekici istatistiksel özelliklere sahiptir. Sigortacılıkta özellikle meydana gelen ciddi ve önemli hasar ödemelerinin tahmin edilmesinde kullanılabilir. Aynı zamanda bu yöntem sıra dışı hasar maliyetlerinin poliçe fiyatlandırmasını etkileyen faktörlerin değerlendirilmesinde de kullanılmaktadır.

Bu çalışmada Türkiye’de faaliyette bulunan bir sigorta şirketinin kasko sigorta verileri üzerinden modele ilişkin uygulama gerçekleştirilmiştir. Bu çalışma ile toplam hasara ilişkin KR analizinin bireysel poliçe sahiplerinin en az bir ve daha fazla hasar getirme olasılıklarına önemli ölçüde bağlı olduğu gösterilmiştir. Analizde öncelikle her poliçe grubu için hasar getirmeme olasılığı tahmin edilmiştir. Bu problem için farklı sınıflandırma faktörlerine göre hasar getirmeme olasılığını belirleyen LR modeli uygulanmıştır. İAKR modelinde LR ile bulunan hasar getirmeme olasılıkları ikinci aşamada farklı poliçe gruplarının kantil değerlerinin hesaplanmasında kullanılmıştır.

KPP tekniği ise primlerin hesaplanmasında kullanılmıştır. KPP’nin BDPP’ye kıyasla poliçe grubuna ilişkin riski daha iyi yakaladığı gözlenmiştir. Bu açıdan değerlendirildiğinde, KPP’nin yalnızca belirli bir risk yükleme faktörünü bütün hasar değişkenlerine uygulanarak hesaplanan prim hesaplamasına göre daha etkin bir şekilde riskli poliçelerin belirlenmesi ve daha doğru fiyatlamanın yapılmasına imkân sağladığı

gösterilmiştir. Ayrıca IFRS 17 raporlama standardının uygulanması ve Solvency II'ye göre sigorta şirketlerinden sermaye yeterliliği hesaplanmasının istenmesi durumunda, KR ile risk yüklemesinin yapılması, riskin karşılıklara daha yüksek güven düzeyinde yansıtılmasına imkân sağlayabilir. KR ile prim hesaplanması sigorta şirketlerinin sermaye yeterlilik hesaplamalarının yetersiz olması durumuna karşı şirketler tarafından kullanılabilir. Kasko branşında uygulanan bu çalışma sağlık sigortalarında da uygulanarak güven düzeyi yüksek prim tutarları hesaplanabilir.

KAYNAKLAR

- [1] Türkiye Sigorta, 2022 Yılında Ekonomik ve Sektörel Gelişmeler, <https://www.turkiyesigorta.com.tr/faaliyetraporlari/turkiyesigorta/2022/files/f2.pdf> (Erişim tarihi: **4 Aralık 2023**).
- [2] Türkiye Sigorta, Reasürans ve Emeklilik Şirketleri Birliği, Kasko Prim Adet Raporu 2023-10, <https://tsb.org.tr/tr/istatistikler> (Erişim tarihi: **4 Aralık 2023**).
- [3] S. Haberman, A. E. Renshaw, Generalized Linear Models and Actuarial Science, *The Statistician*, 45(4) , 407–436, **1996**.
- [4] P. De Jong, G.Z. Heller, Generalized Linear Models for Insurance Data. Cambridge University Press, Cambridge, **2008**.
- [5] J. M. Stanton, Galton, Pearson, and the Peas: A Brief History of Linear Regression for Statistics Instructors, *Journal of Statistics Education*, 9(3): 10.1080/10691898.2001.11910537. Accessed 2021-12-19, **2001**.
- [6] K. Kumari, S. Yadav, Linear regression analysis study, *Journal of the Practice of Cardiovascular Sciences*, 4(1) 33-36, **2018**.
- [7] E. S. Pearson, *Mathematical Statistics and Data Analysis* (2nd ed.), Belmont, CA: Duxbury, **1938**.
- [8] J. A Nelder, R. W. M. Wedderburn, Generalized Linear Models, *Journal of the Royal Statistical Society, Series A*, 135(3), 370-384, **1972**.
- [9] P. McCullagh, J.A. Nelder, *Generalized Linear Models*. 2nd Ed. London: Chapman & Hall, London, **1989**.
- [10] A.E. Renshaw, Actuarial Graduation Practice and Generalised Linear and Non Linear Models, *Journal of the Institute of Actuaries*, 118(2), 295-312, **1991**.
- [11] A.E. Renshaw, R.J. Verrall, A Stochastic Model Underlying The Chain Ladder Technique, *Proceedings of the XXV ASTIN Colloquium, Cannes*, **1994**.

- [12] D. Andersen, S. Feldblum, C. Modlin, D. Schirmacher, E. Schirmacher, N. Thandi, *A Practitioner's Guide to Generalized Linear Models*, Second Edition, CAS Study Note, **2005**.
- [13] A. Heras, I. Moreno, J. L. Vilar-Zanon, An Application Of Two-Stage Quantile Regression To Insurance Ratemaking, *Scandinavian Actuarial Journal* 753–69, **2018**.
- [14] J. Berkson, Application of The Logistic Function To Bio-Assay, *Journal of the American Statistical Association*, 39(227), 357-365, **1944**.
- [15] J. Berkson, Why I Prefer Logits to probits. *Biometrics*, 7(4), 327-339, **1951**.
- [16] D. Finney, *Probit Analysis*. Cambridge University Press, **1971**.
- [17] Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. *Applied Logistic Regression*. New York, Wiley, **2013**.
- [18] R. L. Prentice, A Generalization of the Probit and Logit Methods for Dose response Curves, *Sage Publications, Biometrics* 32(4), 761-768, **1976**.
- [19] H.J. Aldrich, F. D. Nelson, *Linear Probability Logit and Probit Models*, Sage Publications, London, S. 95-96, **1984**.
- [20] S. Menard, *Applied Logistic regression Analysis*, Sage Publications, **2002**.
- [21] R. Koenker, K.F. Hallock, Quantile regression, *Journal of Economic Perspectives* 15 (4), 143–156, **2001**.
- [22] R. Koenker, A. F. Machado, Goodness of Fit and Related Inference Processes for Quantile Regression, *Journal of the American Statistical Association*, 94(448), 1296–1310, **1999**.
- [23] R. Koenker, *Quantile Regression*. Cambridge University Press, Cambridge, **2005**.
- [24] M. Crouhy, et al., *Risk Management*. McGraw-Hill, New York, **2001**.
- [25] P. Jorion, *Value at Risk: The New Benchmark for Managing Financial Risk*. McGraw-Hill, New York, **2007**.
- [26] R.F. Engle, S. Manganelli, *CAViaR: Conditional Value At Risk By Quantile Regression*, National Bureau of Economic Research. Working Paper No. 7341, **1999**.

- [27] E. Portnoy, Regression-Quantile Graduation Of Australian Life Tables, 1946–1992. *Insurance: Mathematics and Economics* 21 (2), 163–172, **1997**.
- [28] D.G.W. Pitt, Regression Quantile Analysis Of Claim Termination Rates For Income Protection Insurance, *Annals of Actuarial Science* 1 (II), 345–357, **2006**.
- [29] R. Abduramanov, A. Kudryavtsev, The Method Of Quantile Regression, A New Approach To Actuarial Mathematics. In: 11th International Congress Insurance: Mathematics and Economics, July 10–12, 2007, Piraeus, Greece. Book of Abstracts, 56–57, **2007**.
- [30] A. A. Kudryavtsev, Using Quantile Regression For Rate-Making, *Insurance: Mathematics and Economics*, 296-304, Elsevier, **2009**.
- [31] F. Baione, D. Biancalana, An Individual Risk Model for Premium Calculation Based on Quantile: A Comparison between Generalized Linear Models and Quantile Regression, *North American Actuarial Journal*, **2019**.
- [32] D. G. Kleinbaum, M. Klein, Logistic Regression, Statistics for Biology and Health, Third Edition, Springer, Atlanta, **2010**.
- [33] D. Cook, P. Dixon, W. Duckworth, M. Kaiser, K. Koehler K, W. Q. Meeker, W. R. Stephenson Binary response and Logistic Regression Analysis, 1–23, **2001**.
- [34] P. D. Allison, Logistic Regression Using The SAS System, 2. Edition, Cary: SAS Institute, **2000**.
- [35] H. Tatlıdil, Uygulamalı Çok Değişkenli İstatistiksel Analiz, Ankara: Cem Web Ofset, **1996**.
- [36] S. Chatterjee, A.S. Hadi, 2006 Regression Analysis by Example. 4th Edition, John Wiley & Sons, Hoboken, **2006**.
- [37] G. Başarır, Uygulamalı Parametrik Olmayan İstatistiksel Yöntemler, İstanbul: Pamaş, **2001**.
- [38] R. Koenker, G. Bassett, Regression Quantiles, *Econometrica* 46 (1), 33–50, **1978**.
- [39] W. G. Gilchrist, Statistical Modelling with Quantile Functions. Florida: Chapman & Hall/CRC, **2000**.
- [40] E. Parzen, Quantile Probability and Statistical Data Modeling, *Statistical Science*, 19(4), 652-662, **2004**.

- [41] L. Hao, D. Naiman, *Quantile Regression*. Sage, **2007**.
- [42] M. P. Doan, *The Roles of Systematic Skewness and Systematic Kurtosis in Asset Pricing*. (Doktora Tezi). Australia: Rmit University, **2011**.
- [43] A. K. Ally, *Quantile-Based Methods for Prediction, Risk Measurement and Inference*. (Doktora Tezi). London: Brunel University, **2010**.
- [44] D. L. Corey, G. Phelps, D. L. Ball, J. Demonte, D. Harrison, Explaining Variation in Instructional Time an Application of Quantile Regression, *Educational Evaluation and Policy Analysis*, 34(2), 146-163, **2012**.
- [45] Z. Zhang, P. Zhu, An Alternative Simple Quantile Regression Estimator, *Economic Letters*, 118(1), 163-166, **2013**.
- [46] B. S. Cade, J. D. Richards, A Permutation Test for Quantile Regression, *Journal of Agricultural, Biological, and Environmental Statistics*, 11(1), 106-126, **2006**.
- [47] P. S. Martins, P. T. Pereira, Does Education Reduce Wage Inequality? Quantile Regression Evidence from 16 Countries, *Labour Economics*, 11(3), 355-371, **2004**.
- [48] M. Buchinsky, Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research, *The Journal of Human Resources*, 33(1), 90-101, **1998**.
- [49] A. Coad, R. Rao, "Innovation and Firm Growth in High-Tech Sectors: A Quantile Regression Approach". *Research Policy*, 1-35, **2007**.
- [50] Z. Landsman, M. Sherris, Risk measures and insurance premium principles, *Insurance: Mathematics and Economics*, 29(1), 103–115, **2001**.
- [51] R. Kaas, M. Goovaerts, J. Dhaene, M. Denuit, *Modern Actuarial Risk Theory: Using R*, volume 128, Springer Science & Business Media, **2008**.
- [52] P. Liebwein, Risk Models for Capital Adequacy: Applications in the context of solvency ii and beyond, *The Geneva Papers on Risk and Insurance-Issues and Practice*, 31(3), pp. 528–550, **2006**.
- [53] P. Artzner, F. Delbaen, J.M. Eber, D. Heath, Coherent measures of risk, *Mathematical Finance*, 9(3), 203–228, **1999**.

- [54] Y. K. Tse, *Nonlife Actuarial Models Theory, Methods and Evaluation*, Cambridge University Press, Newyork, **2009**.
- [55] S. Uryasev, S. Sarykalin, G. Serraino, Uryasev, S., *Value-at-Risk vs. Conditional Value-at Risk in Risk Management and Optimization*, Conference Paper, INFORMS, DOI: 10.1287/educ.1080.0052, **2008**.
- [56] R Core Team *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>, **2022**.

EKLER

EK 1 – Kantil Regresyon Tahminleri

Bağımlı değişken hasarın oluşması durumunda ortaya çıkan toplam hasardır. Hasar getirmeme olasılığı LR model ile hesaplanmıştır. Katsayılar %95 olasılık değeri için tahmin edilmiştir.

Model 1 $\theta_* = 0,95 = 0,8661$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	11,4382	0,03209	356,39569	0,000000
BASAMAK_KODU1	-0,11823	0,0423	-2,79507	0,005190
BASAMAK_KODU2	-0,32049	0,04884	-6,56142	0,000000
BASAMAK_KODU3	-0,33623	0,06249	-5,38024	0,000000
BASAMAK_KODU4	-0,46404	0,05312	-8,73553	0,000000
BASAMAK_KODU5	-0,38731	0,0366	-10,58315	0,000000
ILANKARA	0,04964	0,04108	1,20852	0,226850
İLİSTANBUL	0,05532	0,03102	1,78354	0,074510
İLİZMİR	-0,13734	0,04159	-3,30264	0,000960

Model 2 $\theta_* = 0,95 = 0,8591$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	11,37823	0,03229	352,42091	0,000000
BASAMAK_KODU1	-0,1296	0,04377	-2,96115	0,003070
BASAMAK_KODU2	-0,32388	0,05329	-6,07816	0,000000
BASAMAK_KODU3	-0,35156	0,05714	-6,15289	0,000000
BASAMAK_KODU4	-0,48762	0,06318	-7,71854	0,000000
BASAMAK_KODU5	-0,40334	0,03751	-10,7519	0,000000
ILANKARA	0,07993	0,04241	1,88491	0,059450
İLİSTANBUL	0,07812	0,03201	2,4402	0,014680
İLİZMİR	-0,10853	0,05146	-2,10909	0,034940

Model 3 $\theta_* = 0,95 = 0,8523$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	11,31526	0,03137	360,73857	0,000000
BASAMAK_KODU1	-0,11646	0,04566	-2,55043	0,010760
BASAMAK_KODU2	-0,31567	0,05078	-6,21654	0,000000
BASAMAK_KODU3	-0,36621	0,05859	-6,25034	0,000000
BASAMAK_KODU4	-0,47988	0,05905	-8,12672	0,000000
BASAMAK_KODU5	-0,38848	0,03821	-10,16784	0,000000
ILANKARA	0,08186	0,04323	1,89361	0,058280
İLİSTANBUL	0,09198	0,03387	2,71575	0,006620
İLİZMİR	-0,08689	0,05351	-1,62368	0,104450

Model 4 $\theta_* = 0,95 = 0,8390$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	11,21539	0,03189	351,74481	0,000000
BASAMAK_KODU1	-0,13017	0,04978	-2,61512	0,008920
BASAMAK_KODU2	-0,32284	0,0502	-6,43129	0,000000
BASAMAK_KODU3	-0,34987	0,04718	-7,4159	0,000000
BASAMAK_KODU4	-0,47833	0,05243	-9,12315	0,000000
BASAMAK_KODU5	-0,39061	0,03824	-10,21411	0,000000
ILANKARA	0,08233	0,0431	1,90991	0,056150
İLİSTANBUL	0,09688	0,03227	3,00188	0,002690
İLİZMİR	-0,10225	0,05779	-1,76949	0,076820

Model 5 $\theta_* = 0,95 = 0,8380$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	11,20971	0,03192	351,18438	0,000000
BASAMAK_KODU1	-0,13456	0,04971	-2,70682	0,006800
BASAMAK_KODU2	-0,31978	0,0501	-6,38292	0,000000
BASAMAK_KODU3	-0,34527	0,04656	-7,41597	0,000000
BASAMAK_KODU4	-0,47433	0,04953	-9,57626	0,000000
BASAMAK_KODU5	-0,39155	0,03828	-10,22882	0,000000
ILANKARA	0,08376	0,04313	1,94203	0,052140
İLİSTANBUL	0,09619	0,03152	3,0515	0,002280
İLİZMİR	-0,10398	0,05684	-1,82948	0,067340

Model 6 $\theta_* = 0,95 = 0,8375$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	11,20581	0,032	350,20856	0,000000
BASAMAK_KODU1	-0,13918	0,05051	-2,75515	0,005870
BASAMAK_KODU2	-0,31738	0,05041	-6,29587	0,000000
BASAMAK_KODU3	-0,34678	0,04688	-7,39648	0,000000
BASAMAK_KODU4	-0,47349	0,04982	-9,50364	0,000000
BASAMAK_KODU5	-0,39417	0,03872	-10,18001	0,000000
ILANKARA	0,08458	0,04364	1,93814	0,052610
İLİSTANBUL	0,09769	0,0322	3,0343	0,002410
İLİZMİR	-0,105	0,05749	-1,82627	0,067820

Model 7 $\theta_* = 0,95 = 0,8340$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	11,18252	0,03316	337,27457	0,000000
BASAMAK_KODU1	-0,14499	0,04967	-2,919	0,003510
BASAMAK_KODU2	-0,34082	0,05091	-6,69462	0,000000
BASAMAK_KODU3	-0,33929	0,04904	-6,91858	0,000000
BASAMAK_KODU4	-0,47401	0,04804	-9,86634	0,000000
BASAMAK_KODU5	-0,39502	0,03931	-10,04774	0,000000
ILANKARA	0,08373	0,0451	1,85663	0,063370
İLİSTANBUL	0,10342	0,03188	3,24441	0,001180
İLİZMİR	-0,10027	0,05095	-1,96815	0,049060

Model 8 $\theta_* = 0,95 = 0,8295$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	11,14268	0,03304	337,25304	0,000000
BASAMAK_KODU1	-0,15387	0,05022	-3,06399	0,002190
BASAMAK_KODU2	-0,3229	0,04725	-6,83459	0,000000
BASAMAK_KODU3	-0,32341	0,04833	-6,69198	0,000000
BASAMAK_KODU4	-0,45578	0,04719	-9,65774	0,000000
BASAMAK_KODU5	-0,39112	0,03889	-10,05704	0,000000
ILANKARA	0,08096	0,04111	1,96939	0,048920
İLİSTANBUL	0,11048	0,03209	3,4432	0,000580
İLİZMİR	-0,10146	0,04723	-2,1484	0,031690

Model 9 $\theta_* = 0,95 = 0,8213$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	11,0678	0,03138	352,70531	0,000000
BASAMAK_KODU1	-0,1447	0,0467	-3,09865	0,001950
BASAMAK_KODU2	-0,30342	0,0456	-6,65363	0,000000
BASAMAK_KODU3	-0,30432	0,04721	-6,44571	0,000000
BASAMAK_KODU4	-0,42378	0,04943	-8,57362	0,000000
BASAMAK_KODU5	-0,3704	0,03681	-10,06165	0,000000
ILANKARA	0,08887	0,0424	2,09563	0,036120
İLİSTANBUL	0,12238	0,0303	4,03856	0,000050
İLİZMİR	-0,08023	0,04391	-1,82704	0,067700

Model 10 $\theta_* = 0,95 = 0,8053$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	10,95612	0,02842	385,51663	0,000000
BASAMAK_KODU1	-0,14599	0,04377	-3,33538	0,000850
BASAMAK_KODU2	-0,27989	0,04087	-6,84754	0,000000
BASAMAK_KODU3	-0,28437	0,05111	-5,56346	0,000000
BASAMAK_KODU4	-0,42424	0,05543	-7,6538	0,000000
BASAMAK_KODU5	-0,36252	0,03453	-10,49893	0,000000
ILANKARA	0,09318	0,0404	2,30646	0,021090
İLİSTANBUL	0,13899	0,02999	4,63398	0,000000
İLİZMİR	-0,04859	0,04756	-1,02157	0,306990

Model 11 $\theta_* = 0,95 = 0,8034$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	10,9402	0,02802	390,46648	0,000000
BASAMAK_KODU1	-0,14765	0,04267	-3,4606	0,000540
BASAMAK_KODU2	-0,27096	0,03993	-6,78661	0,000000
BASAMAK_KODU3	-0,28805	0,04761	-6,04964	0,000000
BASAMAK_KODU4	-0,42629	0,05444	-7,8304	0,000000
BASAMAK_KODU5	-0,36236	0,03447	-10,51179	0,000000
ILANKARA	0,10041	0,03857	2,60353	0,009230
İLİSTANBUL	0,14387	0,02957	4,86503	0,000000
İLİZMİR	-0,05052	0,04737	-1,06646	0,286220

Model 12 $\theta_* = 0,95 = 0,8016$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	10,92681	0,02789	391,81571	0,000000
BASAMAK_KODU1	-0,1431	0,04306	-3,32308	0,000890
BASAMAK_KODU2	-0,26338	0,03909	-6,73745	0,000000
BASAMAK_KODU3	-0,28116	0,04661	-6,03219	0,000000
BASAMAK_KODU4	-0,42298	0,05327	-7,93979	0,000000
BASAMAK_KODU5	-0,36203	0,03415	-10,60169	0,000000
ILANKARA	0,10659	0,0373	2,85739	0,004270
İLİSTANBUL	0,14759	0,02931	5,03581	0,000000
İLİZMİR	-0,05063	0,04734	-1,06952	0,284840

Model 13 $\theta_* = 0,95 = 0,7992$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	10,91509	0,02757	395,88731	0,000000
BASAMAK_KODU1	-0,15883	0,04241	-3,7448	0,000180
BASAMAK_KODU2	-0,26393	0,03769	-7,00314	0,000000
BASAMAK_KODU3	-0,28909	0,04804	-6,01787	0,000000
BASAMAK_KODU4	-0,42665	0,05348	-7,97708	0,000000
BASAMAK_KODU5	-0,36679	0,03378	-10,85866	0,000000
ILANKARA	0,106	0,03764	2,81628	0,004860
İLİSTANBUL	0,14934	0,02863	5,21694	0,000000
İLİZMİR	-0,04839	0,04776	-1,01323	0,310960

Model 14 $\theta_* = 0,95 = 0,7912$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	10,86719	0,02863	379,60839	0,000000
BASAMAK_KODU1	-0,17381	0,0424	-4,09909	0,000040
BASAMAK_KODU2	-0,26088	0,04085	-6,3868	0,000000
BASAMAK_KODU3	-0,29234	0,04828	-6,05575	0,000000
BASAMAK_KODU4	-0,43752	0,05256	-8,32384	0,000000
BASAMAK_KODU5	-0,36364	0,03414	-10,65232	0,000000
ILANKARA	0,10729	0,03836	2,79691	0,005160
İLİSTANBUL	0,16081	0,02825	5,69182	0,000000
İLİZMİR	-0,05888	0,05054	-1,1651	0,243980

Model 15 $\theta_* = 0,95 = 0,7812$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	10,80251	0,02861	377,6343	0,000000
BASAMAK_KODU1	-0,16277	0,04079	-3,99056	0,000070
BASAMAK_KODU2	-0,2497	0,04163	-5,99856	0,000000
BASAMAK_KODU3	-0,28609	0,04902	-5,83654	0,000000
BASAMAK_KODU4	-0,43116	0,04746	-9,08556	0,000000
BASAMAK_KODU5	-0,3609	0,03415	-10,56904	0,000000
ILANKARA	0,11742	0,03638	3,22718	0,001250
İLİSTANBUL	0,1693	0,02833	5,97609	0,000000
İLİZMİR	-0,04178	0,05027	-0,83102	0,405970

Model 16 $\theta_* = 0,95 = 0,7615$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	10,67443	0,02985	357,60553	0,000000
BASAMAK_KODU1	-0,15618	0,04154	-3,76003	0,000170
BASAMAK_KODU2	-0,22482	0,04308	-5,21875	0,000000
BASAMAK_KODU3	-0,269	0,04554	-5,90676	0,000000
BASAMAK_KODU4	-0,39088	0,04147	-9,42554	0,000000
BASAMAK_KODU5	-0,3496	0,035	-9,9879	0,000000
ILANKARA	0,13827	0,03579	3,86297	0,000110
İLİSTANBUL	0,18727	0,02764	6,7751	0,000000
İLİZMİR	-0,04759	0,0456	-1,04346	0,296740

Model 17 $\theta_* = 0,95 = 0,7541$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	10,62591	0,02965	358,31856	0,000000
BASAMAK_KODU1	-0,15057	0,04012	-3,75323	0,000170
BASAMAK_KODU2	-0,21764	0,0429	-5,07284	0,000000
BASAMAK_KODU3	-0,26972	0,04478	-6,02329	0,000000
BASAMAK_KODU4	-0,38338	0,04187	-9,15554	0,000000
BASAMAK_KODU5	-0,33843	0,0346	-9,78124	0,000000
ILANKARA	0,13775	0,03521	3,91206	0,000090
İLİSTANBUL	0,19613	0,02748	7,13756	0,000000
İLİZMİR	-0,038	0,04277	-0,88854	0,374260

Model 18 $\theta_* = 0,95 = 0,7372$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	10,51659	0,02771	379,57771	0,000000
BASAMAK_KODU1	-0,14391	0,03827	-3,76071	0,000170
BASAMAK_KODU2	-0,19585	0,04167	-4,7003	0,000000
BASAMAK_KODU3	-0,21676	0,04075	-5,31972	0,000000
BASAMAK_KODU4	-0,33715	0,03964	-8,50501	0,000000
BASAMAK_KODU5	-0,32067	0,03169	-10,11808	0,000000
ILANKARA	0,1393	0,0338	4,12073	0,000040
İLİSTANBUL	0,2156	0,02589	8,32625	0,000000
İLİZMİR	-0,00879	0,03433	-0,25601	0,797940

Model 19 $\theta_* = 0,95 = 0,7148$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	10,39354	0,02651	392,11396	0,000000
BASAMAK_KODU1	-0,1239	0,03585	-3,45547	0,000550
BASAMAK_KODU2	-0,1771	0,04045	-4,3779	0,000010
BASAMAK_KODU3	-0,19199	0,04178	-4,59533	0,000000
BASAMAK_KODU4	-0,3051	0,04185	-7,28995	0,000000
BASAMAK_KODU5	-0,29256	0,03064	-9,54921	0,000000
ILANKARA	0,14518	0,03121	4,65094	0,000000
İLİSTANBUL	0,21872	0,02603	8,40249	0,000000
İLİZMİR	0,01695	0,04108	0,41272	0,679810

Model 20 $\theta_* = 0,95 = 0,7000$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	10,31572	0,02544	405,53933	0,000000
BASAMAK_KODU1	-0,11428	0,03367	-3,39434	0,000690
BASAMAK_KODU2	-0,17158	0,03845	-4,4619	0,000010
BASAMAK_KODU3	-0,1765	0,04145	-4,25758	0,000020
BASAMAK_KODU4	-0,29071	0,04123	-7,0503	0,000000
BASAMAK_KODU5	-0,27973	0,02991	-9,35309	0,000000
ILANKARA	0,15515	0,02904	5,34291	0,000000
İLİSTANBUL	0,22135	0,02512	8,81338	0,000000
İLİZMİR	-0,00498	0,04154	-0,11987	0,904590

Model 21 $\theta_* = 0,95 = 0,6855$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	10,24471	0,02453	417,68873	0,000000
BASAMAK_KODU1	-0,11034	0,03199	-3,44887	0,000560
BASAMAK_KODU2	-0,16912	0,04246	-3,98313	0,000070
BASAMAK_KODU3	-0,17401	0,03859	-4,50969	0,000010
BASAMAK_KODU4	-0,27311	0,04023	-6,78924	0,000000
BASAMAK_KODU5	-0,27541	0,02943	-9,35722	0,000000
ILANKARA	0,17056	0,03164	5,3904	0,000000
İLİSTANBUL	0,23499	0,02402	9,78153	0,000000
İLİZMİR	0,00283	0,04052	0,06989	0,944280

Model 22 $\theta_* = 0,95 = 0,6573$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	10,10514	0,02531	399,28791	0,000000
BASAMAK_KODU1	-0,08533	0,03306	-2,5813	0,009850
BASAMAK_KODU2	-0,18238	0,04255	-4,28603	0,000020
BASAMAK_KODU3	-0,13962	0,03936	-3,54771	0,000390
BASAMAK_KODU4	-0,25788	0,04396	-5,86658	0,000000
BASAMAK_KODU5	-0,26795	0,02992	-8,95472	0,000000
ILANKARA	0,17172	0,03525	4,8714	0,000000
İLİSTANBUL	0,26223	0,02481	10,56761	0,000000
İLİZMİR	0,03392	0,03606	0,94055	0,346940

Model 23 $\theta_* = 0,95 = 0,6539$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	10,08801	0,02547	396,12782	0,000000
BASAMAK_KODU1	-0,08413	0,03429	-2,45359	0,014150
BASAMAK_KODU2	-0,18345	0,04209	-4,3589	0,000010
BASAMAK_KODU3	-0,13535	0,04086	-3,31247	0,000930
BASAMAK_KODU4	-0,25694	0,04384	-5,86019	0,000000
BASAMAK_KODU5	-0,26518	0,02983	-8,88974	0,000000
ILANKARA	0,17467	0,03631	4,80992	0,000000
İLİSTANBUL	0,26545	0,02456	10,80933	0,000000
İLİZMİR	0,03522	0,03811	0,92414	0,355420

Model 24 $\theta_* = 0,95 = 0,6466$

	Değer	Std. Hata	t değeri	Pr(> t)
(Sabit Terim)	10,05669	0,02495	403,14784	0,000000
BASAMAK_KODU1	-0,07834	0,03439	-2,2784	0,022710
BASAMAK_KODU2	-0,17722	0,04048	-4,37821	0,000010
BASAMAK_KODU3	-0,13094	0,04047	-3,23527	0,001220
BASAMAK_KODU4	-0,26895	0,0439	-6,12576	0,000000
BASAMAK_KODU5	-0,25979	0,02922	-8,89001	0,000000
ILANKARA	0,16795	0,03629	4,62754	0,000000
İLİSTANBUL	0,26934	0,02421	11,12465	0,000000
İLİZMİR	0,03754	0,03792	0,98991	0,322230