

**KARAYOLU TRAFİK KAZALARINA
ETKİ EDEN RİSK FAKTÖRLERİNİN ÇOK
DURUMLU LOGİT MODELİ İLE ANALİZİ:
ERZURUM VE KARS İLLERİ ÖRNEĞİ**

Ali Kemal ÇELİK

Yüksek Lisans Tezi

İşletme Anabilim Dalı

Prof. Dr. Erkan OKTAY

2013

Her Hakkı Saklıdır

ATATÜRK ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İŞLETME ANABİLİM DALI

Ali Kemal ÇELİK

KARAYOLU TRAFİK KAZALARINA ETKİ EDEN RİSK
FAKTÖRLERİNİN ÇOK DURUMLU LOGİT MODELİ İLE
ANALİZİ: ERZURUM VE KARS İLLERİ ÖRNEĞİ

YÜKSEK LİSANS TEZİ

TEZ YÖNETİCİSİ
Prof. Dr. Erkan OKTAY

ERZURUM-2013



T.C.
ATATÜRK ÜNİVERSİTESİ
SOSYAL BİLİMLERİ ENSTİTÜSÜ

TEZ BEYAN FORMU



05/07/2013

SOSYAL BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

BİLDİRİM

Atatürk Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğine göre hazırlamış olduğum " Karayolu Trafik Kazalarına Etki Eden Risk Faktörlerinin Çok Durumlu Logit Modeli ile Analizi: Erzurum ve Kars İlleri Örneği " adlı tezin/raporun tamamen kendi çalışmam olduğunu ve her alıntıya kaynak gösterdiğimi taahhüt eder, tezimin/raporumun kağıt ve elektronik kopyalarının Atatürk Üniversitesi Sosyal Bilimler Enstitüsü arşivlerinde aşağıda belirttiğim koşullarda saklanmasına izin verdiğimi onaylarım:

Lisansüstü Eğitim-Öğretim yönetmeliğinin ilgili maddeleri uyarınca gereğinin yapılmasını arz ederim.

- ☐ Tezimin/Raporumun tamamı her yerden erişime açılabilir.
- ☐ Tezim/Raporum sadece Atatürk Üniversitesi yerleşkelerinden erişime açılabilir.
- ☐ Tezimin/Raporumun 3 yıl süreyle erişime açılmasını istemiyorum. Bu sürenin sonunda uzatma için başvuruda bulunmadığım takdirde, tezimin/raporumun tamamı her yerden erişime açılabilir.

05/07/2013

Ali Kemal ÇELİK

F-83/00/22.02.2012



**T.C.
ATATÜRK ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ**



TEZ KABUL TUTANAĞI

SOSYAL BİLİMLER ENSTİTÜSÜ MÜDÜRLÜĞÜNE

Prof. Dr. Erkan OKTAY danışmanlığında, Ali Kemal ÇELİK tarafından hazırlanan bu çalışma 05 / 07 / 2013 tarihinde aşağıdaki jüri tarafından İşletme Anabilim Dalı'nda Yüksek Lisans Tezi olarak kabul edilmiştir.

Başkan : Prof. Dr. Erkan OKTAY

Jüri Üyesi : Prof. Dr. Hüseyin ÖZER

Jüri Üyesi : Yrd. Doç. Dr. Emrah TALAŞ

İmza:

İmza:

İmza:

Yukarıdaki imzalar adı geçen öğretim üyelerine aittir. / /

Prof. Dr. Mustafa YILDIRIM
Enstitü Müdürü

F-85/00/22.02.2012

İÇİNDEKİLER

ÖZET.....	V
ABSTRACT.....	VI
KISALTMALAR DİZİNİ	VII
TABLolar DİZİNİ	VIII
ŞEKİLLER DİZİNİ	X
ÖNSÖZ.....	XI
GİRİŞ	1

BİRİNCİ BÖLÜM

KATEGORİK MODELLER

1.1. KATEGORİK DEĞİŞKENLER.....	5
1.2. KATEGORİK DEĞİŞKENLER İÇİN İSTATİSTİKSEL ÇIKARIMLAR	8
1.2.1.Binom Dağılımı	9
1.2.2.Çok Durumlu Dağılım.....	10
1.2.3.Olabilirlik Fonksiyonu ve En Yüksek Olabilirlik Tahmini.....	11
1.2.4.Wald ve Olabilirlik Oranı Testleri.....	13
1.2.5.Güven Aralığı.....	14
1.2.6.Marjinal Etkiler.....	14
1.3. KATEGORİK TERCİH MODELLERİ	17
1.4. TESADÜFİ FAYDA MODELİ	21

İKİNCİ BÖLÜM

İKİ DURUMLU MODELLER

2.1. İKİ DURUMLU MODELLERİN TANIMI	27
2.1.1. İki Durumlu Doğrusal Olasılık Modeli.....	28
2.1.1.1.İki Durumlu Doğrusal Olasılık Modelinin Tahmini.....	30
2.1.1.1.1. İki Durumlu Doğrusal Olasılık Modelinin Tartılı En Küçük Kareler Yöntemi İle Tahmini	30
2.1.1.1.2. İki Durumlu Doğrusal Olasılık Modelinin Kısıtlı En Küçük Kareler Yöntemi İle Tahmini	31

2.1.1.1.3. İki Durumlu Doğrusal Olasılık Modelinin En Yüksek Olabilirlik Yöntemi İle Tahmini.....	32
2.1.1.1.4. Doğrusal Olasılık Modelinin Tahmininde Karşılaşılan Sorunlar	33
2.1.1.1.4.1. Değişen Varyans (Heteroscedasticity) Sorunu	33
2.1.1.1.4.2. Normal Dağılım Sorunu	34
2.1.1.1.4.3. Uyum İyiliği Ölçüsü Olarak R ² 'nin Kuşkulu Değeri	34
2.1.1.1.4.4. Anlamsız Tahminler Sorunu.....	35
2.1.2.İki Durumlu Logit Modeli	36
2.1.2.1. İki Durumlu Logit Modelinin Tahmini	39
2.1.2.1.1. İki Durumlu Logit Modelinin En Küçük χ^2 Yöntemi ile Tahmini	40
2.1.2.1.2. İki Durumlu Logit Modelinin En Yüksek Olabilirlik Modeli İle Tahmini.....	40
2.1.3. İki Durumlu Probit Modeli.....	41
2.1.3.1. İki Durumlu Probit Modelinin Tahmini	43
2.1.3.1.1. İki Durumlu Probit Modelinin En Küçük Kareler Yöntemi ile Tahmini	44
2.1.3.1.2. İki Durumlu Probit Modelin En Yüksek Olabilirlik Yöntemi ile Tahmini.....	45
2.1.4. Tobit Modeli.....	46
2.1.4.1. Tobit Modelinin En Yüksek Olabilirlik Yöntemi ile Tahmini	48
2.1.5. İki Durumlu Modellerde Uyum İyiliği Ölçüleri.....	48
2.1.5.1. Pearson Ki-kare İstatistiği ve Log-Olabilirlik Oranı Testi.....	50
2.1.5.2. Akaike Bilgi Kriteri.....	51
2.1.5.3. Bayes Bilgi Kriteri	52

ÜÇÜNCÜ BÖLÜM

ÇOK DURUMLU MODELLER

3.1. ÇOK DURUMLU MODELLERİN TANIMI.....	53
3.1.1. Sıralı Modeller	56
3.1.1.1.Sıralı Probit Modeli.....	59
3.1.1.2.Sıralı Logit Modeli.....	60

3.1.2.Sıralı Olmayan Modeller.....	62
3.1.2.1.Çok Durumlu Doğrusal Olasılık Modeli	62
3.1.2.2.Çok Durumlu Logit Modeli	65
3.1.2.2.1.İlişkisiz Alternatiflerin Bağımsızlığı Varsayımı	65
3.1.2.2.2.Çok Durumlu Logit Modelinin Kurulması.....	68
3.1.2.2.3.Çok Durumlu Modellerde Uyum İyiliği.....	73
3.1.2.2.3.1.Hosmer-Lemeshow Testi	73
3.1.2.2.4.Çok Durumlu Logit Modelinin En Yüksek Olabilirlik Yöntemi İle Tahmini	76
3.1.3.Koşullu Logit Model.....	77
3.1.4.Genelleştirilmiş Uç Değer Modeli.....	78
3.1.5.Yuvalanmış (Nested) Logit Modeli.....	80
3.1.6.Çapraz Yuvalanmış Logit Modeli	82
3.1.7.Çok Durumlu Probit Modeli	83
3.1.8.Ardışık Modeller	85
3.1.9.Çok Değişkenli Modeller.....	87
3.1.9.1.Çok Değişkenli Logit Modeli	87
3.1.9.2.Çok Değişkenli Probit Modeli	88
3.1.9.3.Log-Doğrusal Model.....	91

DÖRDÜNCÜ BÖLÜM

ERZURUM VE KARS İLLERİNDE 2008-2012 YILLARINDA MEYDANA GELEN KARAYOLU TRAFİK KAZALARINA ETKİ EDEN RİSK FAKTÖRLERİNİN ÇOK DURUMLU LOGİT MODELİ İLE ANALİZİ

4.1. DÜNYA'DA KARAYOLU TRAFİK KAZALARI İSTATİSTİKLERİ	94
4.2. TÜRKİYE'DEKİ KARAYOLU TRAFİK KAZALARI İSTATİSTİKLERİ	96
4.3. ERZURUM VE KARS İLLERİNDEKİ KARAYOLU TRAFİK AĞLARINA AİT GENEL BİLGİLER	99
4.4. ERZURUM VE KARS İLLERİNDE MEYDANA GELEN KARAYOLU TRAFİK KAZALARI İSTATİSTİKLERİ (2007-2011).....	101
4.5. KARAYOLU TRAFİK KAZALARINA AİT VERİLERİN KATEGORİK MODELLER İLE ANALİZİ.....	102
4.5.1.Literatür Taraması.....	102

4.5.2.Uygulama.....	104
4.5.2.1.Araştırmanın Amacı, Kapsamı ve Önemi	104
4.5.2.2.Araştırmanın Yöntemi	105
4.5.2.3.Araştırmanın Veri Kümesi ve Kısıtları	105
4.5.2.4.Bulgular ve Değerlendirme.....	106
4.5.2.4.1.Değişkenlerin Tanımlanması ve Tanımlayıcı	
İstatistikler	106
4.5.2.4.2.Çok Durumlu Logit Modelinin Kurulması.....	109
4.5.2.4.3.Kurulan Modelin İlişkisiz Alternatiflerin Bağımsızlığı	
Varsayımını Sağlayıp Sağlamadığının Testi.....	109
4.5.2.4.4.Kurulan Modeldeki Değişkenlere Ait Çoklu Bağlantı	
Testi	110
4.5.2.4.5.Kurulan Modelin Uyum İyiliğinin Test Edilmesi.....	111
4.5.2.4.6.Bağımsız Değişkenler için Olabilirlik Oranı Testi	113
4.5.2.4.7.Bağımsız Değişkenler için Wald Testi.....	114
4.5.2.4.8.Kurulan Modelin Tahmini ve Katsayıların	
Yorumlanması	116
4.5.2.4.8.1.Referans Kategorisinin Değiştirilmesi ile Elde	
Edilen Bulguların Yorumlanması.....	120
4.5.2.4.9.Kurulan Modele ait Marjinal Etkiler	135
SONUÇ.....	140
KAYNAKLAR	145
EKLER.....	160
Ek1. 2918 Sayılı Karayolları Trafik Kanunu’nda Yer Alan ve Sıkça İhlal	
Edilen Trafik Kuralları (Asli Kusurlar ve Diğer Sürücü Kusurları)	160
ÖZGEÇMİŞ.....	163

ÖZET**YÜKSEK LİSANS TEZİ****KARAYOLU TRAFİK KAZALARINA ETKİ EDEN RİSK FAKTÖRLERİNİN
ÇOK DURUMLU LOGİT MODELİ İLE ANALİZİ: ERZURUM VE KARS
İLLERİ ÖRNEĞİ****Ali Kemal ÇELİK****Danışman: Prof. Dr. Erkan OKTAY****2013, Sayfa: 176****Jüri: Prof. Dr. Erkan OKTAY (Danışman)****Prof. Dr. Hüseyin ÖZER****Yrd. Doç. Dr. Emrah TALAŞ**

Bu tez çalışmasında karayolu trafik kazalarına etki eden risk faktörlerinin çok durumlu logit modeli yardımıyla analizi yapılmaya çalışılmıştır. Bu amaçla karayolu trafik kazası türü ölümlü, yaralanmalı ve maddi hasarlı olarak üç kategoriye ayrılmış ve bağımlı değişken olarak kabul edilen karayolu trafik kaza türüne göre Erzurum ve Kars illerinde 2008-2012 yılları arasında meydana gelen 9162 karayolu trafik kazası çok durumlu logit modeli ile analiz edilmiştir.

Bu çalışma dört bölümden oluşmaktadır. Birinci bölümde kategorik modeller ve kategorik değişkenlerin istatistiksel çıkarımları üzerine genel bilgiler verilmiştir. İkinci bölümde iki durumlu modeller, üçüncü bölümde ise çok durumlu modeller ile ilgili teorik altyapı irdelenmiştir. Dördüncü bölümde ise dünyada ve Türkiye’de karayolu trafik kazaları ile istatistiksel bilgiler verildikten sonra araştırma kapsamında elde edilen verilerin çok durumlu logit modeli kullanılarak analizi sunulmuştur. Sonuç bölümünde elde edilen bulgular genel olarak gözden geçirilerek araştırma kapsamındaki illerde karayolu trafik kazalarını artıran ve azaltan risk faktörlerinin hangi kapsamda değerlendirilmesi gerektiği ortaya konulmuştur.

Anahtar Kelimeler: İki Durumlu Modeller, Çok Durumlu Logit Modeli, Karayolu Trafik Kazası, Risk Faktörü

ABSTRACT**MASTER THESIS****THE ANALYSIS OF RISK FACTORS AFFECTING ROAD TRAFFIC
ACCIDENTS USING MULTINOMIAL LOGIT MODEL: THE CASE OF
ERZURUM AND KARS PROVINCES****Ali Kemal ÇELİK****Advisor: Prof. Dr. Erkan OKTAY****2013, Pages: 176****Jury: Prof. Dr. Erkan OKTAY (Advisor)****Prof. Dr. Hüseyin ÖZER****Assist. Prof. Dr. Emrah TALAŞ**

This thesis aims to analyze the risk factors affecting road traffic accidents using multinomial logit model. For this purpose, road traffic accidents were considered in three categories with respect to traffic accident type including fatal, non-fatal and property damage only accidents. By determining the road traffic accident type as a dependent variable; road traffic accidents, which were occurred in Erzurum and Kars provinces during 2008-2012, were analyzed using multinomial logit model.

The study consists of four sections. First section presents general information about categorical models and statistical inferences for categorical variables. Second and third sections examine the theoretical background relating to models for dichotomous and polytomous variables. Fourth section exhibits the analysis of the data using multinomial logit model within the scope of the study after introducing road traffic accidents statistics in the world and in Turkey. The conclusion section involves the review of the research findings and how these risk factors that increase and decrease the occurrence of road traffic accidents in Erzurum and Kars provinces can be assessed.

Keywords: Dichotomous Models, Multinomial Logit Model, Road Traffic Accident, Risk Factor.

KISALTMALAR DİZİNİ

AIC	: Akaike Bilgi Kriteri
BIC	: Bayes Bilgi Kriteri
G.A.	: Güven Aralığı
IIA	: İlişkisiz Alternatiflerin Bağımsızlığı Varsayımı
OR	: Bahis Oranı
RRR	: Göreli Risk Oranı
TÜİK	: Türkiye İstatistik Kurumu
vb.	: Ve benzeri

TABLOLAR DİZİNİ

Tablo 4. 1. Türkiye’deki Motorlu Kara Taşıtları Sayısı (2003-2013 Yılları Arası).....	96
Tablo 4. 2. Türkiye’de Meydana Gelen Karayolu Trafik Kazaları İstatistikleri (2003-2013 Yılları Arası) ¹	97
Tablo 4. 3. Karayolu Trafik Kazalarında Taşıt ve Nüfus Başına Düşen Ortalama Ölü ve Yaralı Sayısı (2002-2012 Yılları Arası)	98
Tablo 4. 4. Nüfusun Karayolu Trafik Kazalarından Etkilenme Oranı (2002-2012 Yılları Arası).....	99
Tablo 4. 5. Erzurum İlinde 2007-2011 Yılları Arasında Meydana Gelen Karayolu Trafik Kaza Sayıları	101
Tablo 4. 6. Kars İlinde 2007-2011 Yılları Arasında Meydana Gelen Karayolu Trafik Kaza Sayıları	101
Tablo 4. 7. Modele Dâhil Edilen Bağımsız Değişkenlere ait Tanımlamalar.....	106
Tablo 4. 8. İlişkisiz Alternatiflerin Bağımsızlığı Varsayımı Testi	110
Tablo 4. 9. Modeldeki Bağımsız Değişkenlere ait Varyans Şişkinlik Faktörleri.....	111
Tablo 4. 10. Kurulan Modelin Uygunluk Değerleri	112
Tablo 4. 11. Kurulan Modelin Uyum İyiliği Testi	112
Tablo 4. 12. Bağımsız Değişkenler için Olabilirlik Oranı Testi	113
Tablo 4. 13. Bağımsız Değişkenler için Wald Testi.....	115
Tablo 4. 14. Karayolu Trafik Kazalarını Etkileyen Risk Faktörlerinin Çok Durumlu Logit Modeli ile Analizi.....	116
Tablo 4. 15. Karayolu Trafik Kazasının Meydana Geldiği Mevsime Göre Bahis Oranlarının İkili Karşılaştırılması	121
Tablo 4. 16. Karayolu Trafik Kazasının Meydana Geldiği Saate Göre Bahis Oranlarının İkili Karşılaştırılması.....	122
Tablo 4. 17. Karayolu Trafik Kazasının Meydana Geldiği Haftaya Göre Bahis Oranlarının İkili Karşılaştırılması	123
Tablo 4. 18. Karayolu Trafik Kazasının Meydana Geldiği Gündeki Hava Durumuna Göre Bahis Oranlarının İkili Karşılaştırılması	123
Tablo 4. 19. Karayolu Trafik Kazasının Meydana Geldiği Güne Göre Bahis Oranlarının İkili Karşılaştırılması.....	124

Tablo 4. 20. Karayolu Trafik Kazasının Meydana Geldiği Yolda Yaya Kaldırımının Olup Olmamasına Göre Bahis Oranlarının İkili Karşılaştırılması	125
Tablo 4. 21. Karayolu Trafik Kazasının Meydana Geldiği Yolun Yüzeyine Göre Bahis Oranlarının İkili Karşılaştırılması	126
Tablo 4. 22. Karayolu Trafik Kazasının Meydana Geldiği Yolda Aydınlatma Olup Olmamasına Göre Bahis Oranlarının İkili Karşılaştırılması	127
Tablo 4. 23. Karayolu Trafik Kazasına Karışan Araç sayısına Göre Bahis Oranlarının İkili Karşılaştırılması.....	128
Tablo 4. 24. Karayolu Trafik Kazasına Karışan Araç Cinsine Göre Bahis Oranlarının İkili Karşılaştırılması.....	129
Tablo 4. 25. Karayolu Trafik Kazasına Karışan Araç Sürücüsünün Yaş Grubuna Göre Bahis Oranlarının İkili Karşılaştırılması	130
Tablo 4. 26. Karayolu Trafik Kazasına Karışan Araç Sürücüsünün Öğrenim Durumuna Göre Bahis Oranlarının İkili Karşılaştırılması	132
Tablo 4. 27. Karayolu Trafik Kazasına Karışan Araç Sürücüsünün Sürücü Belgesinin Sınıfına Göre Bahis Oranlarının İkili Karşılaştırılması.....	133
Tablo 4. 28. Karayolu Trafik Kazasına Karışan Araç Sürücüsünün İhlal Ettiği Trafik Kurallarına Göre Bahis Oranlarının İkili Karşılaştırılması	134
Tablo 4. 29. Ölümlü Kaza Türüne Ait Marjinal Etkiler	135
Tablo 4. 30. Yaralanmalı Kaza Türüne Ait Marjinal Etkiler	136
Tablo 4. 31. Maddi Hasarlı Kaza Türüne Ait Marjinal Etkiler	138

ŞEKİLLER DİZİNİ

Şekil 4. 1. Erzurum İline ait Karayolu Ağı Haritası	100
Şekil 4. 2. Kars İline ait Karayolu Ağı Haritası	100

ÖNSÖZ

Bu çalışmanın seçiminde ve yürütülmesinde yoğun iş yüküne rağmen bana yardımcı olan danışmanım Prof. Dr. Erkan OKTAY'a, Ekonometri Bölümü hocalarım Prof. Dr. Hüseyin ÖZER ve Yrd. Doç. Dr. Emrah TALAŞ'a, tezin hazırlanma süreci boyunca yardım ve desteklerini esirgemeyen Prof. Dr. Abdulbaki BİLGİÇ, Yrd. Doç. Dr. Ötüken SENER, Yrd. Doç. Dr. Ömer ALKAN, Arş. Gör. Abdulkerim KARAASLAN, Arş. Gör. Hayri ABAR, Arş. Gör. Miraç EREN, Arş. Gör. Murat KURUN ve Arş. Gör. Ahmet İlker AKBABA'ya teşekkürü bir borç bilirim. Ayrıca, trafik kazası tespit tutanaklarına erişebilmemi sağlayan ve çalışma süresince yakın ilgilerini gösteren başta Erzurum İl Emniyet Müdürü Sayın H. Turgut YILDIZ ve Kars İl Emniyet Müdürü Sayın Ercan ÇAKMAK olmak üzere; Erzurum Bölge Trafik Denetleme Şube Müdürlüğü, Kars Bölge Trafik Denetleme Şube Müdürlüğü, Erzurum Trafik Tescil ve Denetleme Şube Müdürlüğü ve Kars Trafik Tescil ve Denetleme Şube Müdürlüğü'nde görev yapan tüm personele; manevi desteklerini hep hissettiğim anneme ve babama teşekkürlerimi sunarım.

Erzurum – 2013**Ali Kemal ÇELİK**

GİRİŞ

Kategorik modeller, bağımlı değişkenin kategorik yapıda değerlendirildiği modeller olarak adlandırılmaktadır. Kategorik modellerin birçok farklı disipline ait uygulamalarından bahsetmek mümkün olabilmektedir. Kategorik modellerde bağımlı değişken yalnızca 0 ya da 1 değerini alıyor ise iki durumlu modeller, üç ya da daha fazla kategori söz konusu ise çok durumlu modeller ve kurulan modelde birden fazla kategorik bağımlı değişken ile birçok bağımsız değişken mevcut ise çok değişkenli modeller uygulama alanı bulmaktadır. Temelinde bireylerin olduğu bir kategorik modelin yapısı belirli varsayımlar altında karar vericiler, alternatiflere ait tutumların kümesi ve karar kuralını kapsamaktadır. Kategorik modeller, sözü edilen karar vericilerin birbirinden bağımsız ve tüm olasılıkları içeren bir sonlu alternatifler kümesinden elde edilen sonuçların olasılığını tahmin etmek için kullanılmaktadır. Araştırmacılar, sonuç değişkenine ait değerlerin mevcut bilgi ile tam olarak belirlenemediği durumlarda araştırma sürecine dâhil olmaktadır.

İki durumlu modeller; iki durumlu doğrusal olasılık modeli, iki durumlu logit modeli, iki durumlu probit modeli ve tobit modeli olarak sınıflandırılmaktadır. İki durumlu doğrusal olasılık modeli, en temel tanımıyla bir doğrusal regresyon modelinin iki durumlu bağımlı değişkene uygulanması olarak düşünülmektedir. İki durumlu doğrusal olasılık modelinin tahmininde karşılaşılan değişen varyans, normal dağılım, kuşkulu R^2 ve anlamsız tahminler sorunları Ağırlıklı En Küçük Kareler yöntemi ile çözülmeye çalışılsa bile bu yöntem de yeterli olmamakta ve bu sebeple alternatif modellerin türetilmesine ihtiyaç duyulmaktadır. Yalnızca 0 ya da 1 değerini alan bağımlı değişken ile bağımsız değişkenler arasındaki ilişkiyi birikimli standart normal dağılım kullanarak açıklayan iki durumlu probit modeli doğrusal olasılık modeline alternatif modeller arasında bulunmaktadır. İki durumlu probit modelinde tahmin yapabilmek için doğrusal olmayan normal birikimli dağılım fonksiyonunun doğrusal hale getirilmesi gerekmektedir. İki durumlu probit modelinin En Küçük Kareler yöntemi ile tahmininde karşılaşılan sorunlar nedeniyle en yüksek olasılık yöntemiyle tahmin edilmesi daha tutarlı sonuçlar vermektedir. İki durumlu doğrusal olasılık modeline alternatif olarak geliştirilen iki durumlu logit modelinde hata terimi, lojistik birikimli dağılıma sahip olmaktadır. Logit model kullanışlı yapısıyla literatürde en çok

tercih edilen iki durumlu modeldir. Logit modelin temelinde ve değerlendirilmesinde basitçe bir olayın gerçekleşme olasılığının gerçekleşmeme olasılığına bölünmesiyle elde edilen bahis oranı bulunmaktadır. Bahis oranının kullanılması ile hesaplama tüm gözlemler üzerinden yapılabilmekte ve diğer değişkenlerin aldığı değerlere bağımlı olmamaktadır. Gizli değişkenin sıfırdan büyük olduğu durumda gözlenebilir olduğunu varsayan alternatif bir diğer model de tobit modelidir. Diğer iki durumlu modellerinde olduğu gibi sınırlandırılmış tobit modelinin de tutarlı tahminleri en yüksek olasılık yöntemi ile elde edilebilmektedir.

Çoklu alternatiflerden kaçınmak mümkün olmadığından bağımlı değişkenin iki durumlu olduğu modeller ikiden fazla kategorinin söz konusu olduğu modellere de uygulanabilmektedir. Çok durumlu modellerin temel amacı, bahis değerlerini ortak değişkenlerin bir fonksiyonu olarak açıklayabilmektir. Çok durumlu modeller en genel sınıflama ile sıralı ve sıralı olmayan modeller olarak ayrılmaktadır. Sıralı modeller tanım gereği bağımlı değişkene ait kategorilerin doğal bir sıralamaya sahip olduğu modellerdir. Sıralı olmayan modeller ise bağımlı değişkene ait kategorilerin doğal bir sıralamaya sahip olmadığı modeller olup çok durumlu logit modeli ile bu modelin ilişkisiz alternatiflerin bağımsızlığı varsayımına esneklik getiren çok durumlu probit modelini, yuvalanmış modeli ve ardışık modelleri kapsamaktadır. Birden çok bağımlı ve bağımsız değişkenlerden oluşan çok değişkenli modeller içerisinde çok değişkenli logit ve probit modellerinin yanında log-doğrusal modeller de yer almaktadır.

Günümüzde, karayolu trafik kazalarının sebep olduğu olumsuz sonuçların dünyada üzerindeki ölümlerin ve yaralanmaların en dikkat çekici tetikleyicilerinden biri olduğu düşünülmektedir. Dünya Sağlık Örgütü'nün karayolu trafik güvenliği verilerine göre, dünya üzerinde her yıl yaklaşık 1 milyon 250 bin kişi karayolu trafik kazalarında yaşamını yitirmekte ve yaşamını yitiren kişiler arasında genç nüfusun çoğunlukta olması mevcut durumun vahamiyetini açık bir şekilde ortaya koymaktadır. Öte yandan, karayolu trafik kazalarının sebep olduğu ölümlerin % 91'si düşük ya da orta gelire sahip ülkelerde meydana gelmekte ve dünyada meydana gelen karayolu trafik kazasında ölen kişilerin yarısını yayalar, motosiklet ve bisiklet sürücüleri gibi savunmasız karayolu kullanıcıları oluşturmaktadır. 2007 yılından beri 88 ülkede karayolu trafik kazası ölümlerinde azalma olsa da özellikle düşük ve orta gelirli 87 ülkede ölümlerdeki artış

sürmektedir. Karayolu trafik kazalarının sözü edilen dramatik sonuçlarına rağmen dünyada az sayıda ülke karayolu trafik kazalarını önleme adına caydırıcı yasalara sahiptir.

Orta gelir düzeyine sahip bir ülke olarak Türkiye’de karayolu trafiğine çıkan motorlu araç sayısındaki büyük artış karayolu trafik kazalarının olumsuz etkilerini de beraberinde getirmektedir. Türkiye’de 2003-2013 yılları arasında, 1 milyondan fazlası ölümlü veya yaralanmalı ve yaklaşık 8 milyonu maddi hasarlı olmak üzere yaklaşık 9 milyon karayolu trafik kazası meydana gelmiştir. Sözü edilen kazalarda 43 binden fazla kişi yaşamını yitirmiş, yaklaşık 2 milyon kişi yaralanmıştır. Maddi ve manevi boyutları birlikte ele alındığında karayolu trafik kazalarının Türkiye için de ne kadar endişe verici sonuçlar doğurduğu aşikârdır. Tüm dünyada olduğu gibi Türkiye’de de karayolu trafik kazalarının olumsuz sonuçlarının kontrol alınması ve trafik güvenliğinin sağlanması adına etkin politikaların geliştirilmesi için karayolu trafik kazalarına sebep olan risk faktörlerinin tespit edilmesi gerekmektedir.

Bu tez çalışmasında karayolu trafik kazalarını etkileyen risk faktörlerinin istatistiksel olarak tespit edilmesi amaçlanmaktadır. Bu amaçla Erzurum ve Kars illerinde 2008-2012 yıllarında meydana gelen 9162 karayolu trafik kazasının istatistiksel analizinin yapılması yoluna gidilmektedir. Sırasıyla Karayolu 12. ve 18. Bölge Müdürlüklerinde yer alan Erzurum ve Kars illeri nüfuslarına oranla trafiğe kayıtlı araç sayısının ve bulundukları bölge ile kıyaslandıklarında nispeten uzun karayolu trafik ağlarıyla karayolu trafik kazalarının olumsuz etkilerini yansıtmaktadır. Buna rağmen bilindiği kadarıyla sözü edilen illerde karayolu trafik kazalarını etkileyen faktörlerin istatistiksel tespiti adına yapılan bilimsel bir çalışma bulunmamaktadır. Karayolu trafik kazasının ölümlü, yaralanmalı ve maddi hasarlı olmak üzere mevcut üç türü olduğundan kaza türü bağımlı değişken kabul edilerek karayolu trafik kazalarını etkilediği düşünülen risk faktörlerinin belirlenmesi adına çok durumlu logit modelinin uygulanmasına karar verilmiştir. İlgili ve benzer modelleri içine alan detaylı bir teorik altyapı bilgisi verildikten sonra uygulama bölümüne geçilmiş ve kurulan modelde maddi hasarlı karayolu trafik kazaları referans kategori olarak tercih edilip söz konusu model çoklu bağlantı, ilişkisiz alternatiflerin varsayımı, uyum iyiliği, olabilirlik oranı ve Wald oranı testlerine tabi tutulmuştur. Çalışmanın son bölümünde bağımlı değişkene

ait referans kategorilerin deęiřtirilmesi ile elde edilen bahis oranlarının deęerlendirilmesi ve kurulan modele ait marjinal etki sonuları sunulmaktadır. Sonu blmnde elde edilen bulgular ıřıęında ileriki yıllar iin arařtırma kapsamındaki illerde karayolu trafik kazalarının olumsuz etkilerinin azaltılması ve trafik gvenlięinin saęlanması adına izlenebilecek politikalar tartıřılmıřtır.

BİRİNCİ BÖLÜM

KATEGORİK MODELLER

1.1. KATEGORİK DEĞİŞKENLER

Değişkenler nicel (sayısal) ya da nitel (sayısal olmayan) özelliğe sahiptir. Nicel değişkenler kesikli ve sürekli değişkenler olmak üzere iki kısma ayrılmaktadır. Sürekli değişkenler, belirli bir aralıkta sonsuz sayıda değer alabilen değişkenler olarak tanımlanabilirken, kesikli değişkenler ise yalnızca belirli tamsayı değerleri alabilen değişkenler olarak ifade edilebilmektedir. Bu bağlamda, ölçülen ya da tartılan değişkenleri sürekli değişkenler, sayılabilen değişkenleri kesikli değişkenler olarak tanımlamak mümkündür. Örneğin, bir okuldaki öğrencilerin ağırlıkları belirli aralıkta tamsayı olmayan değerler olabildiği için sürekli değişken, bir sınıftaki öğrencilerin sayısı ancak belirli bir tamsayı değeri alabileceği için kesikli değişken olacaktır (Başar ve Oktay, 2010: 2). Bir başka deyişle, bir bağımlı değişken, iki ya da daha fazla değer aldığı anda, kategorik ya da kategorik olmayan değişken olarak sınıflandırılmaktadır (Maddala ve Flores-Lagunes, 2001: 367).

Kategorik bağımlı değişken, ölçek değeri kategori kümesinden oluşan değişkendir. Kategorik ölçümler, sosyal bilimlerde birçok konuda tutum ve düşünce ölçümünde yaygın olarak kullanılmaktadır. Kategorik değişkenler sosyal bilimlerde yaygın olarak kullanılmasına rağmen, bu değişkenlerin kullanım alanları yalnızca bu disiplinlerle sınırlandırılmaz. Kategorik değişkenlerin sıklıkla kullanıldığı diğer kullanım alanlarını iktisat (örneğin; alış-satış rakamları), pazarlama, (örneğin; tüketicilerin marka tercihleri), davranış bilimleri (örneğin; şizofreni, depresyon gibi ruhsal hastalık türleri), tıp (örneğin hasta sayısı), sosyoloji ve psikoloji (örneğin; insan sayısı), zooloji (örneğin; hayvan türleri sınıflamaları), eğitim (örneğin; öğrencilerin sınavda vermiş olduğu “doğru” ve “yanlış” cevap sayısı), mühendislik bilimleri, endüstriyel kalite kontrol, vb olarak sıralamak mümkündür (Agresti, 1996: 1-2; Leonard, 2000: 1).

Stevens (1946) ölçüm kavramını olayları veya birimleri kurallara göre numaralandırma olarak tarif etmiş ve sınıflama (nominal), sıralama (ordinal), aralık (interval) ve oran (ratio) ölçeği olmak üzere dört tür ölçek tanımlamıştır (Özer, 2004: 5). Kategorik değişkenler için sınıflama ve sıralama ölçekler olmak üzere iki temel

ölçek türü bulunmaktadır. Doğal bir sıralamaya sahip olmayan kategorik değişkenler, sınıflama ölçeği ile ölçülmüş kategorik değişkenler olarak tanımlanmaktadır. Sınıflama ölçeği, bir cisim veya olayı belirli bir isme göre diğer cisim veya olaydan ayırmak için kullanılmaktadır. Örneğin, hane halkı konut tercihleri (apartman, müstakil ev, lojman, diğer), bireylerin dinlediği favori müzik türleri (klasik, pop, rock), medeni durum (evli, bekar, dul) gibi değişkenleri sınıflama ölçeği ile ölçülmüş kategorik değişkenler olarak sınıflandırmak mümkündür. Sınıflama ölçeği ile ölçülmüş değişkenler, sayısal değerler, sözcükler ya da harfler ile temsil edilebilmektedir. Bu aşamada, kural olarak farklı sınıflar için aynı sayısal değerlerin ya da aynı sınıflar için farklı sayısal değerlerin belirlenmemesi gerekmektedir. Sınıflama ölçeği ile ölçülmüş veriler, üzerinde durulan özelliğin olması ya da olmamasına göre birbirinden farklı kategorilerde, üzerinde durulan özelliğin olması durumu 1, olmaması durumu 0 ile gösterilebilmektedir. Sınıflama ölçeği ile ölçülmüş değişkenler için kategorilerin keyfi sıralanması söz konusu olduğundan, bu değişkenler için kullanılacak istatistiki analiz bu sıralamaya bağlı olmayacaktır (Agresti, 2002: 2; Stevens, 1946: 679; Oktay, 2005: 2).

Sıralama ölçeği ile ölçülmüş kategorik değişkenlerde, kategoriler kendi aralarında sıralı olarak yer almaktadır. Bu tür kategorik değişkenler özellikle anket çalışmalarında yer alan ifadelerle ilgili görüşlerin alınmasında sıklıkla kullanılmaktadır. Anket çalışmalarında değişkenlerin sıralı olan kategorileri genellikle üç, beş ya da yedi kategoriyle sınırlanmakta ve “Likert” ölçekli değişkenler olarak adlandırılmaktadır. Likert ölçeğinde değişken kategorileri arasındaki uzaklıklar kesin olarak bilinmediğinden sıralama ölçeği olarak ele alınıp analiz edilmesi daha doğru bir yaklaşımdır. Sıralı ölçekli kategorik değişkenler, yalnızca araştırma konusu ifadelerin ölçülmesi dışında, değişkenlerin kategorileri arasında sıralamanın olması durumunda da kullanılmaktadır. Örneğin, araştırma konusu bireyin eğitim düzeyi; ilköğretim, lise, üniversite ve lisansüstü kategorileri ile ya da bireyin yaşı; (18-25), (26-35), (36-50) ve (50+) gibi yaş grubu kategorileriyle ölçüldüğünde kategorilerin arasında sıralama olmadığından değişkenlerin sıralı ölçekli olduğunu göstermektedir (Agresti, 2010: 2; Arıcıgil Çılan, 2009: 6).

Sıralı ölçekli değişkenler için sıralamanın değişmesi durumunda yapılan istatistik analizler farklı sonuçlar vermektedir. Bununla birlikte, sınıflama ölçeği ile ölçülmüş değişkenler için herhangi bir sıralama söz konusu olmadığından sıralı ölçekli

değişkenler için kullanılan yöntemler sınıflama ölçeği ile ölçülmüş değişkenler için kullanılamamaktadır. Sınıflama ölçeği ile ölçülmüş değişkenler için kullanılan ölçekler sıralı ölçekli yöntemler için kullanılabilir. Ancak, sınıflama ölçeği ile ölçülmüş değişkenler için kullanılan yöntemlerin sıralı ölçekli değişkenler için kullanılması halinde sıralamaya ilişkin bilgiler dikkate alınmamaktadır. Bu durum analiz aşamasında ciddi zayıflıklara sebep olabileceğinden, ölçek türü için en uygun yöntemin kullanılması önerilmektedir (Özer, 2005: 6).

Sınıflama veya sıralama ölçeği ile ölçülmüş iki değişkenin alt kategorilerini sağlama durumları frekans tablolarında gösterilebilmektedir. X bağımsız değişkeni ile Y bağımlı değişkeni sınıflama ya da sıralama ölçeği ile ölçülmüş ise X bağımsız değişkeni 1, 2, ..., I kategorilerinden ve Y bağımlı değişkeni 1, 2, ..., J kategorilerinden oluşacaktır. Bu şekilde X ve Y değişkenlerinin bileşik dağılımlarından bahsedilmektedir. X değişkeninin birinci kategorisi ile Y değişkeninin birinci kategorisini sağlayan gözlem sayısı n_{11} ile gösterilmektedir. Bu şekilde I adet bağımsız değişken kategorisi ile J adet bağımlı değişken kategorisi çapraz sınıflandırıldığında, $I \times J$ adet hücredeki n_{ij} frekansları elde edilecektir. Kontenjans tablolarındaki satır toplamaları, $n_{1+}, n_{2+}, \dots, n_{i+}, \dots, n_{I+}$ ve sütun toplamaları, $n_{+1}, n_{+2}, \dots, n_{+j}, \dots, n_{+J}$ ile gösterildiğinde, satır ve sütun toplamaları araştırmaya katılan toplam birim sayısı n değerini vermektedir (Oktay, 2005: 3, Reynolds, 1984: 13).

Kategorik değişkenler, objektif bir ölçü birimiyle ölçülememekte, çeşit yönünden farklılıklar göstermektedirler (Karagöz, 2010: 2). Regresyon çözümlemesinde bağımlı değişken özünde nitel olan değişkenlerden sık sık etkilendiği için, bu kategorik değişkenlerin açıklayıcı değişkenler arasına katılması gerekmektedir. Bu tür değişkenleri kategorik hale getirmenin bir yolu 0 ya da 1 değerlerini alan yapay değişkenler oluşturmaktır. 0-1 değerlerini alan bu değişkenlere gölge değişkenler denilmektedir. Gölge değişkenler ile ilgili literatürde gösterge değişkenler, iki değerli değişkenler, yapay değişkenler, sınıf değişkenleri, nicel değişkenler ve iki uçlu değişkenler gibi birçok almasıık adlandırmaya rastlamak mümkündür (Gujarati, 2011: 500).

Regresyon çözümlemelerinde gölge değişkenlerin kullanılması ile araştırmacılar; tahmin edilen olasılığın özellikleri hakkında fikir sahibi olabilmekte, söz

konusu olasılıktaki birim değişimin modele katkısını tahmin edebilmekte ve bağımsız değişkenlerdeki bir birim değişimin bağımlı değişkene etkisini oransal olarak yorumlayabilmektedir (Pampel, 2000: 1-2). Regresyon analizinde gölge değişken belirlenirken, çoklu bağlantı (collinearity) durumunu önlemek adına; eğer sınıflama ölçeği ile ölçülmüş bağımsız değişken k kategoriden oluşuyorsa, regresyon modelinin bir sabit terime (β_0) sahip olacağı göz önüne alınarak toplam $k-1$ adet gölge değişken tanımlanabilmektedir. Eğer regresyon analizi sabit terime sahip değilse, bazı durumlarda k adet gölge değişken de kullanılabilir (Kleinbaum, Kupper, Nizam, ve Muller, 2009: 151). Kategorik değişkenlerde sıfır değeri verilen özellik ile diğer özelliklerin karşılaştırılabilirliklerinde bir temel oluşturmaktadır. Burada, hangi özelliğin temel özellik olacağı, önsel bazı düşüncelerin etkili olduğu bir tercihi yansıtmaktadır ve gölge değişkenin 0 ve 1 değerini aldığı durumların değişmesi sonucu değiştirmemektedir. Bazı durumlarda, kategorik değişkenler model tahminlerinde bağımlı değişkenler yerine de kullanılmaktadır (Özer, 2005: 14).

Sosyal bilimlerde özellikle düşünceler ve tutumların ön planda olduğu sonuç değişkenleri çoğunlukla kategorik ölçekler ile ölçülmektedir. Önceki yıllarda bu tür veri kümelerinin modellenmesinde yaygın olarak En Küçük Kareler yöntemi kullanılmaktaydı ve sözü edilen veriler ya özgün ölçek yöntemi korunarak ya da varyansları daha sabit hale getirilerek değerlendirilmekteydi. Günümüzde ise kategorik değişkenlerin analizinde geliştirilen özel modeller tercih edilmektedir. Örneğin, iki durumlu değişkenler için lojistik regresyon; çok durumlu değişkenler için genelleştirilmiş logit modeller ve Poisson dağılımına sahip değişkenler için log-doğrusal modeller tercih edilebilmektedir (Agresti, Booth, Hobert ve Cappel, 2000: 28).

1.2. KATEGORİK DEĞİŞKENLER İÇİN İSTATİSTİKSEL ÇIKARIMLAR

Bir bilimsel çalışmada araştırmacılar uygun veri kümesi üzerinde gözlem yapabilmek için istatistiksel çıkarım yapma sorunuyla karşılaşmaktadır. Sözü edilen istatistiksel çıkarım süreci; gözlenemeyen veriler için hipotez kurma, modeldeki gözlenebilen veriler için çıkarımlarda bulunma ve gözlemden beklenen sonuçları ifade edebilme adına deneysel bir araştırma girişimi olmak üzere üç aşamadan oluşmaktadır. Araştırmacının kullandığı veriler beklentilerine uygun ise model gözlenemeyen verilerin

doğru bir gösterimi olarak kabul edilmektedir (Bock, 1975: 1). Bir ana kütle, ölçüm ve sonuçlar üzerindeki tüm olası gözlemlerin toplamını, sonlu bir ana kütle sonsuz olmayan tüm olası gözlemlerin sayısını ifade edilmektedir. Bir ana kütle sayısal özellikleri ise parametreler ile gösterilmektedir (Kmenta, 1997: 4). Daha önceden belirlenen bir veri kümesinde, bir olasılık dağılımı için olabilirlik fonksiyonu, bilinmeyen parametre olarak işlem gören söz konusu verilerin olasılığı olarak tanımlanmaktadır. Benzer şekilde, en yüksek olabilirlik tahmini ise olasılık fonksiyonunun en yüksek değerini belirlemeye çalışan parametre değeri olarak ifade edilmektedir. Doğal olarak gözlenen veriler ile ilişkili anılan parametre değeri en yüksek gerçekleşme olasılığına sahip olmaktadır. Olabilirlik fonksiyonunun en yüksek olmasını sağlayan parametre değeri aynı zamanda bu fonksiyona ait logaritma değerinin de en yüksek olmasına imkân tanımaktadır. Log-olabilirlik değeri bir toplam değeri olduğu sürece, söz konusu en yüksek log-olabilirlik değerini elde etmek daha kolay olacaktır. Zayıf düzenleyici koşullar altında, en yüksek olabilirlik tahmincileri, daha cazip özelliklere sahip olmaktadır. Örneğin; sözü edilen tahminciler büyük örneklerde normal dağılım göstermekte, parametre yakınsaması uygulandığında asimtotik tutarlılığa sahip olmakta ve büyük örneklerde diğer tahmin yöntemlerine göre daha küçük standart hata değerleri üreterek asimtotik etkinlik sağlamaktadır (Agresti, 2002: 9).

1.2.1. Binom Dağılımı

Tesadüfi bir değişken, tesadüfi bir deney sonucu elde edilen sayısal sonuçların özel olarak tanımlanması olarak ifade edilmektedir. Tesadüfi değişkenlere ait her bir değer birer gözlem olarak düşünülmektedir. Bir tesadüfi değişken için olasılıklara ait özellikler ise olasılık dağılımı olarak adlandırılmaktadır. Birçok tesadüfi değişken bir ortalama ve standart sapmaya sahip olup bir tesadüfi değişken esasına dayanan her bir olayın bir olasılığa sahip olduğu bilinmektedir (Siegel, 2012: 155). Bir binom deneyi, Bernoulli¹ sınamalarından oluşan deney olarak tanımlanmaktadır. Birbirinden bağımsız

¹ Bernoulli dağılımına sahip rastgele bir değişkene (örneğin X) ait fonksiyon eşitliği aşağıdaki gibidir:

$$f(x) = p^x(1-p)^{1-x}, \quad x = 0,1$$

Bernoulli dağılımının ortalaması μ ve varyansı σ^2 ;

$$\mu = E(X) = p$$

$$\sigma^2 = E(X^2) - [E(X)]^2 = p \cdot q = p \cdot (1-p)$$

şeklindedir (Akdeniz, 2013: 200).

n özdeş Bernoulli sınamasından başarılı olanların toplam sayısı X tesadüfi değişkeni; bir tek deneme için başarılı olma olasılığı P , başarısız olma olasılığı $1-P$ olmak üzere, tesadüfi binom değişkeni olarak ifade edilmektedir. Bu durumda; her bir deneme için başarı ve başarısızlık olmak üzere yalnızca iki sonucun olması, P olasılığının her bir deneme için aynı olması ve denemelerin birbirinden bağımsız olması gibi gereksinimler karşılanmaktadır. Böylece, X binom tesadüfi değişkeninin olasılık fonksiyonu,

$$f(x) = \binom{n}{x} p^x q^{n-x}, \quad x = 0, 1, 2, \dots, n \quad (1.1)$$

şeklinde ifade edilmekte ve bu dağılım binom dağılımı olarak adlandırılmaktadır. Binom dağılımının ortalaması, $\mu = E(X) = np$ ve varyansı, $\sigma^2 = E(X^2) - [E(X)]^2 = npq$ 'dir (Akdeniz, 2013: 200-201, 204).

1.2.2. Çok Durumlu Dağılım

Her bir n bağımsız ve özdeş sınama için c adet kategori bulunduğu varsayıldığında, i sınaması j kategorisinde yer alıyorsa $y_{ij} = 1$; belirtilen kategoride yer almıyorsa $y_{ij} = 0$ olacaktır. Bu bağlamda, $y_i = (y_{i1}, y_{i2}, \dots, y_{ic})$ değişkeni çok durumlu (multinomial) bir sınamayı ifade edecektir. Ek olarak, y_i değişken kümesinde yer alan y_{ic} değişkeni doğrusal olarak diğer değişkenlere bağımlı olacağından sınamaya katkısı olmayacaktır. Böylece, $n_j = \sum_i y_{ij}$ değişkeni j kategorisinde yer alan sınamaların sayısını ifade etmek üzere, bu sayılar $(n_1, n_2, \dots, n_c)$ çok durumlu dağılım göstermektedir. Benzer şekilde, $\pi_j = P(Y_{ij} = 1)$ değeri j kategorisindeki her bir sınama için çıktıların olasılığını ifade etmek üzere çok durumlu olasılık yığın fonksiyonu, (1.2) eşitliğindeki gibi gösterilebilecektir:

$$p(n_1, n_2, \dots, n_{c-1}) = \left(\frac{n!}{n_1! n_2! \dots n_{c-1}!} \right) \pi_1^{n_1} \pi_2^{n_2} \dots \pi_c^{n_c} \quad (1.2)$$

$\sum_j n_j = n$ olması durumunda, yukarıdaki (1.2) eşitliği $(c-1)$ boyutlu ve $n_c = n - (n_1 + \dots + n_{c-1})$ olmaktadır. Kolayca görülebileceği gibi, binom dağılımı, çok durumlu dağılımın $c = 2$ olduğu özel bir durumudur. Çok durumlu dağılımda, n_j değişkeni için beklenen değer, varyans ve ortak değişken eşitlikleri sırasıyla; $E(n_j) = n\pi_j$, $\text{var}(n_j) = n\pi_j(1-\pi_j)$ ve $\text{cov}(n_j, n_k) = -n\pi_j\pi_k$ şeklinde ifade edilmektedir (Agresti, 2002: 6-7). Çok durumlu dağılımların, tablolar ya da grafikler ile ifade edilmesi oldukça karmaşık olduğundan bu dağılımlar için bir örnek sonucu binom dağılımında olduğu gibi tek gibi

sayıyı değil, bütün frekans dağılımını göstermektedir. Kural olarak, verilen bir kesikli olasılık dağılımı özelleştirildiğinde, N büyüklüğündeki her bir muhtemel örnek frekans dağılımının olasılığı çok durumlu dağılım kuralına uymaktadır (Hays, 1994: 155).

1.2.3. Olabilirlik Fonksiyonu ve En Yüksek Olabilirlik Tahmini

Genel bir tahmin sorununa ait parametre değeri β ve en yüksek olabilirlik tahmin değeri $\hat{\beta}$, olabilirlik fonksiyonu $l(\beta)$, log-olabilirlik fonksiyonu $L(\beta) = \log[l(\beta)]$ olmak üzere, birçok modelde $L(\beta)$ içbükey bir şekle sahip olmakta ve $\hat{\beta}$ tahmin değeri ise türevin sıfır olduğu noktada yer almaktadır. Bu durumda, en yüksek olabilirlik tahmini, $\partial L(\beta) / \partial \beta = 0$ olduğunda elde edilen olabilirlik fonksiyonu çözümüne eşit olmaktadır. Bir olabilirlik fonksiyonunun parametre değerlerini içeren bölümüne “kernel” fonksiyonu denilmekte ve olabilirlik fonksiyonunun parametrelere göre en yüksek değerinin elde edilmesi amaçlandığında geri kalan değerlerin olabilirlik fonksiyonu ile ilişkisiz olduğu düşünülmektedir. Başka bir gösterim ile (1.1) eşitliğindeki gibi bir binom dağılımı düşünüldüğünde ve $\binom{n}{x}$ binom katsayısının bir π değerine bağlı olarak ulaşabilen en yüksek değere hiçbir etkisinin olmadığı varsayıldığında, bu değer ihmal edilmekte ve kernel değeri olabilirlik fonksiyonu olarak işlem görmektedir. Bu varsayımlar altında binom dağılımına ait log-olabilirlik fonksiyonu,

$$L(\pi) = \log[\pi^x (1 - \pi)^{n-x}] = x \log(\pi) + (n - x) \log(1 - \pi) \quad (1.3)$$

şeklinde yazılmakta ve bu eşitliğin π değerine göre türevi,

$$\partial L(\pi) / \partial \pi = x / \pi - (n - x) / (1 - \pi) = (x - n\pi) / (\pi(1 - \pi)) \quad (1.4)$$

eşitliğindeki gibi olmaktadır. Söz konusu türevin sıfıra eşitlenmesi ile n adet sınama için örnek başarı oranının çözümü $\hat{\pi} = y/n$ olan olabilirlik fonksiyonu elde edilmektedir. (1.4) eşitliğinden hareket ile $\partial^2 L(\pi) / \partial \pi^2$ değeri hesaplanarak beklenen değerinin alınması ile aşağıdaki (1.5) eşitliğine ulaşılmaktadır:

$$-E[\partial^2 L(\pi) / \partial \pi^2] = E[x / \pi^2 + (n - x) / (1 - \pi)^2] = n / [\pi(1 - \pi)] \quad (1.5)$$

Böylece, binom dağılımı gösteren $\hat{\pi}$ tahmin değerine ait beklenen ortalama ve varyans

değerleri sırasıyla $E(\hat{\pi}) = \pi$ ve $\sigma(\hat{\pi}) = \sqrt{\frac{\pi(1-\pi)}{n}}$ şeklinde olmaktadır (Agresti, 2002: 10-

11).

En yüksek olabilirlik yöntemi, doğrusal olmayan modeller için uygundur ve $\partial \log L / \partial \theta = 0$ olabilirlik eşitliğini sağlayan bir $\hat{\theta}$ en yüksek olabilirlik tahmincisi, ortalaması sıfır ve ortak varyans matrisi,

$$-E \frac{\partial^2 \log L(y; \theta_0)}{\partial \theta \partial \theta'} \quad (1.6)$$

ifadesinin tersine eşit olan asimtotik normal dağılıma sahip olmaktadır. Ortak varyans eşitliğindeki $L(y; \theta_0)$ değeri, $\theta = \theta_0$ olduğunda elde edilen olabilirlik fonksiyonunu temsil etmektedir. Uygulamada, (1.6) eşitliği yaklaşık olarak,

$$\frac{\partial^2 \log L(y; \theta_0)}{\partial \theta \partial \theta'} \quad (1.7)$$

şeklinde de gösterilebilmektedir (Chow, 1983: 230). En yüksek olabilirlik tahmincisi, örnek olabilirlik fonksiyonunun en yüksek değerini elde etmek için kullanılmakta ve bu sebeple olabilirlik fonksiyonunun,

$$\log L = \sum \log f(y; \theta) = \sum_i L_i \quad (1.8)$$

şeklinde logaritması alınmaktadır. Sözü edilen eşitliğinin θ parametresine göre türevi alındığında,

$$\frac{\partial \log L}{\partial \log \theta} = \sum_i \frac{\partial L_i}{\partial \theta} = \sum_i Z_i = \sum z(Y; \theta) \quad (1.9)$$

olmaktadır. Bu eşitliğin sıfıra eşitlenmesi $\sum z(y; T) = 0$ şeklinde olmakta ve bu ifadenin çözülmesi T değeri için en yüksek olabilirlik tahmincisini vermektedir (Goldberger, 1991: 134-135). En yüksek olabilirlik yöntemi ile tahmin edilen modeller parametrik modeller olmakta ve bu tür modelde parametre değerleri bilindiğinde bağımlı değişkeni temsil edebilecek gerekli bilgi elde edilebilmektedir. Hata terimlerinin normal dağıldığının varsayıldığı regresyon modelleri için en yüksek olabilirlik tahmincileri En Küçük Kareler tahmincilerine benzer işlev görebilmektedir. En yüksek olabilirlik tahmini regresyon modellerinden farklı olarak birçok uç modele de uygulanabilmekte ve tatminkâr asimtotik özelliklere sahip tahminler sağlayabilmektedir (Davidson ve MacKinnon, 1999: 393).

1.2.4. Wald ve Olabilirlik Oranı Testleri

Büyük örnekler için olabilirlik fonksiyonu çıkarımları yapabilmek için çeşitli standart yöntemlerin olduğu kabul edilmektedir. Sıfır hipotezinin $H_0: \beta = \beta_0$ şeklinde ifade edildiği düşünülerek bu testlerde en yüksek olabilirlik tahmincilerinin büyük örnek normalliği kullanılmaktadır. $\hat{\beta}$ tahmin değerine ait sıfır olmayan standart hatalar ile birlikte z test istatistiği,

$$z = (\hat{\beta} - \beta_0) / \text{Standart hata} \quad (1.10)$$

şeklinde tanımlanmakta ve $\beta = \beta_0$ iken yaklaşık standart normal dağılıma sahip olmaktadır. Böylece, z istatistiğine ait tek ya da iki yönlü P olasılık değerleri standart normal dağılım tablosu ile karşılaştırılarak elde edilebilmektedir. İki yönlü alternatifler için z^2 değeri 1 serbestlik derecesinde ki-kareli sıfır dağılımına sahip olmaktadır. Bu durumda, P olasılık değeri gözlenen değer üzerinde sağ kuyruklu ki-kare olasılık değerine eşit olmakta ve sıfır olmayan standart hatalar kullanılan bu istatistiğe Wald istatistiği denilmektedir. Sıfır hipotezine ait Wald istatistiğinin çok değişkenli uzantısı,

$$W = (\hat{\beta} - \beta_0) [\text{cov}(\hat{\beta})]^{-1} (\hat{\beta} - \beta_0) \quad (1.11)$$

şeklinde yazılmaktadır.

İkinci genel amaçlı yöntem, olabilirlik fonksiyonunu iki farklı orana ait en yüksek değerleri elde etmek için kullanılmaktadır. Sözü edilen oranlardan birincisi, H_0 hipotezi altında olası parametrelerin, ikincisi ise H_0 ya da H_1 hipotezlerinin doğru olduğu durumda daha büyük parametre kümelerinin en yüksek değerlerini elde etmeyi amaçlamaktadır. Buna göre, l_0 değeri sıfır hipotezi altında, l_1 değeri ise genel olarak (H_0 ya da H_1 hipotezleri için) olabilirlik fonksiyonunun en yüksek değeri olduğunda, l_0 daha sınırlı bir parametre kümesi ile ilgili olduğundan l_1 değerinden büyük olamayacak ve dolayısıyla $\Lambda = l_0 / l_1$ değeri her zaman birden küçük olacaktır. Ayrıca, $-2\log\Lambda$ değeri $n \rightarrow \infty$ giderken sınırlı sıfır ki-kareli dağılıma sahip olmakta ve serbestlik derecesi $H_0 \cup H_1$ hipotezleri altında elde edilen parametre uzaylarının boyutları arasındaki farka eşit olmaktadır. Böylece olabilirlik oranı istatistiği,

$$-2\log\Lambda = -2\log(l_0/l_1) = -2(L_0 - L_1) \quad (1.12)$$

ifadesine eşit olmaktadır. Bu eşitlikte L_0 ve L_1 değerleri en yüksek değeri elde edilen log-olabilirlik fonksiyonunu temsil etmektedir (Agresti, 2002: 11).

1.2.5. Güven Aralığı

Birim varyansa ve μ ortalamaya sahip normal dağılımlı bir ana kütleden elde edilen $(X_1, \dots, X_n)$ şeklinde tesadüfi bir örnek olduğu varsayıldığında, örnek ortalaması $\bar{X}$ değerinin μ ortalamasına göre normal dağıldığı ve varyans değerinin $1/n$ olduğu bilinmektedir. Benzer şekilde, normal dağılım tablosunda 0.95 olasılık ile

$$P \left[\bar{X} - \frac{1.96}{\sqrt{n}} \leq \mu \leq \bar{X} + \frac{1.96}{\sqrt{n}} \right] \quad (1.13)$$

olasılık değeri μ parametresi için güven aralığını temsil etmektedir. Ancak, varyans değeri σ^2 gibi bilinmeyen bir parametreye eşit olduğunda, (1.13) eşitliğindeki aralık değerinin limiti $\bar{X} \pm 1.96\sigma\sqrt{n}$ olacak ve örnek değerleri vasıtasıyla hesaplanamayacaktır. Bu durumda, payı sıfır ortalama ve σ^2 varyansa eşit olan,

$$\frac{(\bar{X}-\mu)\sqrt{n}}{\sqrt{(n-1)s^2}}\sqrt{n-1} = \frac{(\bar{X}-\mu)\sqrt{n}}{s} \quad (1.14)$$

oran değeri göz önüne alınmakta ve Student dağılım tablosu ışığında $n - 1$ serbestlik derecesinde, $1-\alpha$ güven katsayısı olmak üzere $\pm t_{\alpha/2}(n - 1)$ ifadesinin limit değerleri,

$$P[-t_{\alpha/2}(n - 1) \leq t(n - 1) \leq t_{\alpha/2}(n - 1)] = 1 - \alpha \quad (1.15)$$

olasılık değeri yardımıyla elde edilebilmektedir (Theil, 1971; 94). Sağ kuyruklu olasılık değerine ve standart normal dağılıma sahip bir α ifadesine ait z-değerinin z_α ile gösterildiği varsayıldığında, Wald güven aralığı, $|\hat{\beta} - \beta_0|/\text{Standart Hata} < z_{\alpha/2}$ için kurulan β_0 kümesinden oluşmakta ve $\hat{\beta} \pm z_{\alpha/2}$ aralığını vermektedir. Benzer şekilde, olabilirlik oranı temelinde güven aralığı, $-2[L(\beta_0) - L(\hat{\beta}_0)] < \chi_1^2(\alpha)$ için β_0 kümesinden oluşmaktadır (Agresti, 2002: 13).

1.2.6. Marjinal Etkiler

Marjinal etki bir niceliğin diğerine göre oran değişimini açıklamaktadır. Daha açık bir ifade ile marjinal etki, bağımsız değişkendeki bir birimlik değişimin bağımlı değişken üzerindeki değişimini göstermektedir. İleride detaylı bir şekilde açıklanacak logit ve probit modelleri için x_i üzerindeki bir birimlik değişim modeller parametreler

üzerinde doğrusal ise y_i^* üzerinde β büyüklüğünde bir değişime sebep olmaktadır. Benzer şekilde, logit ya da probit model için β değerinin x_i değerindeki değişimi ifade ettiği söylenebilmektedir. Örneğin bir logit ya da probit modelinin

$$y_i^* = \beta_0 + \beta_1 x_1 + \beta_2 d_i + \varepsilon_i \quad (1.16)$$

şeklinde tanımlandığı varsayıldığında, x_1 değişkenindeki bir birimlik değişimin y_i^* değişkeni üzerindeki etkisi,

$$\frac{\partial y_i^*}{\partial x_1} = \beta_1 \quad (1.17)$$

olarak ifade edilen bir parçalı türeve eşit olduğu görülmektedir. Kesikli bir değişken için marjinal etki,

$$E(y_i^* | d_i = 1) - E(y_i^* | d_i = 0) = \beta_2 \quad (1.18)$$

farkına eşit olmaktadır. Marjinal etkiler, literatürde en yaygın olarak olasılıklar cinsinden ifade edilmektedir. Bu bağlamda, k 'nci bağımsız değişkenin marjinal etkisi,

$$\frac{\partial P(y_i = 1 | x_i)}{\partial x_{ik}} = \frac{\partial F(x_i' \beta)}{\partial x_{ik}} = f(x_i' \beta) \beta_k \quad (1.19)$$

eşitliği ile gösterilmektedir. Bu eşitlikte $F(.)$ ifadesi birikimli dağılım ve $f(.)$ ifadesi yoğunluk fonksiyonunu temsil etmektedir. Ayrıca, bu nicelik yakınındaki x değerine göre başarı olasılığındaki değişimi açıklamaktadır (Powers ve Xie, 2000: 77).

1.2.7. Çoklu Bağlantı Sorunu

Klasik bir regresyon modelinin en önemli varsayımlarından biri modeldeki açıklayıcı değişkenlerden hiçbirinin birbirleriyle ya da açıklayıcı değişkenlerin doğrusal bileşenleriyle tam ilişki içerisinde olmamasıdır. Bu varsayım ihlal edildiğinde, değişkenler arasında kusursuz çoklu bağlantıdan söz edilmektedir. Öte yandan, tüm açıklayıcı değişkenlerin birbiri ile ilişkisiz olduğu durumlarda ise çoklu bağlantı eksikliği söz konusu olmaktadır. İki açıklayıcı değişkene sahip çoklu regresyon modelinde, kusursuz çoklu bağlantı,

$$x_{i2} = a + b x_{i3} \quad (1.20)$$

şeklinde yazılmaktadır. Bu eşitlikte a ile b sabit sayılar ve $b \neq 0$ olup bu durumda örnekteki iki açıklayıcı değişken arasında tam ilişki olduğu ifade edilmektedir. Sözü edilen model için En Küçük Kareler normal eşitlikleri,

$$m_{Y2} = \hat{\beta}_2 m_{22} + \hat{\beta}_3 m_{23} \quad (1.21)$$

$$m_{Y3} = \hat{\beta}_2 m_{23} + \hat{\beta}_3 m_{33} \quad (1.22)$$

şeklinde yazılmakta, ancak çoklu bağlantı eşitliğinden

$$m_{Y2} = b m_{Y3} \quad (1.23)$$

$$m_{22} = b^2 m_{33} \quad (1.24)$$

$$m_{23} = b m_{33} \quad (1.25)$$

olduğu bilindiğinden En Küçük Kareler eşitlikleri,

$$b m_{Y3} = b(\hat{\beta}_2 b m_{33} + \hat{\beta}_3 m_{33}) \quad (1.26)$$

$$m_{Y3} = \hat{\beta}_2 b m_{33} + \hat{\beta}_3 m_{33} \quad (1.27)$$

olarak yeniden yazılmaktadır. Böylece, iki eşitliğin bağımsız olmadığı ve $\hat{\beta}_2$ ile $\hat{\beta}_3$ tahmin parametrelerine ait çözüm kümelerinin kesinlik ifade etmediği görülmektedir. Bu bağlamda, yeniden yazılacak çoklu bağlantı eşitliği,

$$X_{i2} = a + b_3 X_{i3} + b_4 X_{i4} \quad (1.28)$$

şeklinde ifade edilmektedir. Bu eşitlikte a , b_3 , b_4 yine sabit sayılar olarak tanımlanmakta ve b_3 ile b_4 değerleri sıfırdan farklı olmaktadır (Kmenta, 1997: 430, 433).

Açıklayıcı değişkenler yüksek derecede ilişkili olduğunda, her bir değişkenin bağımlı değişken üzerindeki ayrı ayrı etkisini görebilmek zorlaşmaktadır. Bu sebeple kullanılan modelde ikiden fazla açıklayıcı değişken mevcut ise ilişkiler arası dereceyi görebilmek için basit ilişki analizi yerine R_i^2 değerlerinin kullanılması tavsiye edilmektedir. Bununla birlikte, çoklu bağlantının istatistiksel çıkarımlar için sorun teşkil edip etmediği R_i^2 değeri yerine standart hata ve t -değerleri gibi diğer faktörlere bağlı olmaktadır. Örneğin, R_i^2 değeri düşük olsa bile çoklu bağlantının model için bir sorun teşkil edip etmediği hususundaki değerlendirme yalnızca bu değere bakılarak

yapılamamaktadır. Modeldeki çoklu bağlantıyı ölçebilmek için tercih edilen ölçüm değerlerinden biri varyans şişkinlik faktörüdür ve bu değer,

$$VIF(\hat{\beta}_i) = \frac{1}{1 - R_i^2} \quad (1.29)$$

şeklinde ifade edilmektedir. Bu eşitlikteki R_i^2 değeri x_i değişkeni ile diğer açıklayıcı değişkenler arasındaki çoklu ilişki katsayısını temsil etmektedir (Maddala, 1992: 271-272).

1.3. KATEGORİK TERCİH MODELLERİ

Bir regresyon modeli, bağımlı değişkenin bir ya da daha fazla bağımsız değişkenin sayısal değerleri ile açıklanması olarak tanımlanmaktadır. Regresyon analizi, çok değişkenli yapısından dolayı istatistiki çözümlemelere çok daha fazla açıklayıcı güç sağlaması, bilgisayar paket programlarında uygulanmasının ve yorumlanmasının daha kolay olması gibi üstünlükleri sayesinde sosyal bilimlerde oldukça popüler bir yaklaşım olarak kabul edilmektedir (Aldrich ve Nelson, 1984: 9). Günümüzde regresyon yöntemleri, bir bağımlı değişken ile bir ya da birden fazla bağımsız değişken arasındaki ilişkiyi tanımlama amacındaki her türlü veri analizinin tamamlayıcı bir bileşeni olarak düşünülmektedir. Bu bağımsız değişkenler, çoğunlukla ortak değişkenler (covariates) olarak adlandırılmaktadır. Bunun yanında, bağımlı değişkenin, doğrusal regresyon modellerinde olduğu gibi sürekli olduğu varsayılmaktadır (Hosmer ve Lemeshow, 2000: 1). Regresyon analizinin temel amaçlarını; bağımsız değişkenlerin mevcut değerleri ile bağımlı değişkenin ortalama değerini tahmin etmek, bağımsız değişkenlerin bağımlı değişken üzerindeki etkisini araştırmak ve bağımsız değişkenlerin mevcut değerleri ile bağımlı değişkenin ortalama değerinin öngörüsünü yapmak ya da gelecekteki alacağı değeri tahmin etmek olarak sıralamak mümkündür (Tarı, 2012: 15).

Bir doğrusal regresyon modelinin En Küçük Kareler yöntemi ile tahmin edilebilmesi ve model katsayıları üzerine çıkarımlar yapılabilmesi için aşağıda sıralanan bir takım özel varsayımların sağlanması gerekmektedir:

1. Bir regresyon modelindeki bütün bağımsız değişkenler aralık, oran ya da iki değerli; bağımlı değişken ise aralık ya da oran ölçüm yöntemleri ile ölçülmelidir. Bağımsız değişken sürekli olmalı ve modeldeki bütün değişkenler hatasız ölçülmelidir.

2. Bağımlı değişken ile ilgili bütün belirleyiciler analize dâhil edilmeli, bağımsız değişken ile ilgisiz bütün belirleyiciler analizden çıkarılmalıdır. Bağımsız değişken ile belirleyiciler arasındaki ilişki doğrusal olmalıdır.
3. Hata teriminin beklenen değeri sıfır olmalıdır ($E(\varepsilon) = 0$).
4. Bağımsız değişkenlerin bütün değerleri için hata teriminin varyansı aynı ya da sabit olmalıdır.
5. Bağımsız değişkenin her bir değer kümesi için hata terimleri normal dağılıma sahip olmalıdır.
6. Bağımsız değişkenlerin farklı değerlerinden türetilen hata terimleri arasında herhangi bir ilişki olmamalıdır ($E(\varepsilon_i, \varepsilon_j) = 0$).
7. Hata terimleri ile bağımsız değişkenler arasında herhangi bir ilişki bulunmamalıdır ($E(\varepsilon_i, X_j) = 0$).
8. Çoklu regresyon durumunda, bağımsız değişkenlerden hiçbiri diğer bağımsız değişkenin doğrusal bir bileşimi olmamalıdır (Menard, 2002: 4-5).

Olasılık modelleri, regresyon modelleri gibi bağımlı değişkenin olasılık yoğunluk fonksiyonunu belirlemeyi amaçlamakta ve tesadüfî değişimin olasılık modellerdeki yeri olasılık modellerinin diğer modellerden farkını ortaya koymaktadır. Regresyon modellerinde tesadüfî değişim, modele gelişigüzel hata terimi tanımlanmasıyla açıklanmaya çalışılırken hata teriminin beklenen değeri üzerinden yorumlama yapılmaktadır. Bir başka ifade ile hata terimi, modelin sistematik bir bileşeni olarak görülmektedir. Bu durumun aksine, olasılık modellerinde tesadüfî değişim olmazsa olmaz bir olgu olarak görülmekte ve modelin kurulum aşamasında bir olasılık mekanizması olarak doğrudan yer almaktadır. Bu özellikleri ile olasılık modelleri toplam veriler yerine daha çok bireysel değişim üzerinde yoğunlaşmaktadır. (Cramer, 1986: 150).

Kategorik tercih modelleri, bağımlı değişkenin kesikli değerler aldığı regresyon modelleridir. Doğada davranış ile ilgili birçok tercih nitel olduğu için, kategorik tercih modellerinin iktisat biliminde çok sayıda uygulama alanından bahsetmek mümkündür. Örneğin, bir tüketici otomobil satın alma ya da satın almama kararı verirken; bir yolcu birçok uygun seçenek arasından ulaşım tercihi yaparken ya da bir çalışan kendisine önerilen bir işi kabul edip etmeme tercihinde bulunurken nitel tercihlerde bulunmaktadır. Kategorik tercih modelleri, biyometri gibi birçok farklı disipline ait

çalışmada daha geniş uygulama alanı bulsa dahi, bu çalışmaların genellikle iki durumlu tercih modeline yönelik çalışmalar olduğu göze çarpmaktadır. Günümüzde araştırmacılar, bir bağımsız değişkenin ikiden daha fazla kesikli değişken aldığı çok durumlu (multinomial) modeller ya da birden fazla kesikli bağımlı değişken ve birçok bağımsız değişkenden oluşan çok değişkenli (multivariate) modeller gibi daha karmaşık modeller ile sıklıkla ilgilenmektedir (Amemiya, 1985: 267). İstatistikçiler, yanıt değişkeninin değerinin diğer değişkenlere ait bilgi ile tam olarak belirlenemediği ve sabit değerlere sahip bağımsız değişkenler için tekrarlanan sınamaların ya da gözlemlerin aynı büyüklükte sonuçları işaret etmediği durumlarda, regresyon sürecine dâhil olmaktadır (Finney, 1971: 1).

Bir tercih kümesinin üç temel özelliğe sahip olduğu düşünülmektedir. Birinci özelliğe göre tercih kümesinde yer alan alternatiflerin karar vericinin bakış açısı çerçevesinde “karşılıklı dışlayıcılık” özelliğine sahip olduğu varsayılmaktadır. Bir başka ifade ile karar vericinin bir alternatifi tercih etmesi diğer herhangi bir alternatifi tercih edememesi anlamına gelmekte ve tercih kümesinden yalnızca bir alternatifi seçebilmektedir. İkinci özelliğe göre tercih kümesinin “kapsamlı” olması gerekmekte ve olası bütün alternatiflerin tercih kümesine dâhil edildiği düşünülmektedir. Şüphesiz ki, bu durumda da karar verici tercih kümesinden yalnızca bir tercih yapmaktadır. Üçüncü özelliğe göre ise tercih kümesinde yer alan alternatif sayısının sonlu bir sayı olması gerekmekte ve araştırmacıya mevcut alternatifleri sayabilme olanağı sağlanmaktadır (Train, 2003: 15).

Bir kategorik tercih modelinin yapısı, bir genel varsayımlar kümesi kullanılarak açıklanabilmektedir. Sözü edilen varsayımlar altında karar vericiler, alternatifler, tutumlar ve karar kuralı olmak üzere dört farklı ayırımdan bahsedilmektedir. Bir kategorik tercih modelinde karar vericiler, karar veren birimleri ve bu birimlerin özelliklerini ifade etmekte; alternatifler, karar vericiler için mevcut seçenekleri belirlemekte; tutumlar, bir alternatifi karar verici için faydalarını ve maliyetlerini ölçmekte ve karar kuralı, karar vericinin bir alternatifi tercih etmesi için izleyeceği süreci açıklamaktadır (Ben-Akiva ve Bierlaire, 1999: 6).

Kategorik tercih modelleri ile ilgili gözleme dayalı birçok çalışma basit tercih durumları temeline dayanmaktadır. Bu noktada, bir karar vericinin nispeten küçük ve iyi

tanımlanmış bir ana kütle ve alternatifler kümesi arasından beğenilerin türdeş olduğu varsayılarak bir alternatif tercih ettiği düşünülmektedir. Ancak, tercih modellerin genişletilmiş uygulamalarında bu basit tercih sürecinden ve iyi tanımlanmış alternatifler kümesinden çoğunlukla sapmalar olmakta; beğeni farklılıkları, gözlenemeyen tercih kümeleri ve karmaşık tercih süreçleri sorunları ile karşılaşmaktadır. Temel olarak sözü edilen başlıca sorunlar; beğeni farklılığı sorunu, büyük alternatif kümeleri arasından tercih yapma sorunu, çoklu davranışın söz konusu olduğu durumlarda tercih yapma sorunu, ardışık kararlar sorunu ve dışsal faktörlerin tercih kararlarını etkileme sorunu şeklinde sıralanmaktadır (Ben-Akiva vd., 1997: 274).

Geniş anlamıyla bir tercih davranışı; kullanılabilir bilgiye dayanan algılar ve inanışlardan ortaya çıkan ve bunun yanında duygular, tutumlar, dürtüler, tercihler gibi birçok faktörden etkilenen bir karar süreci olarak nitelendirilmektedir (Ben-Akiva vd. 1999: 188). Bir tercih davranışı araştırması; karar vericiler için uygun alternatif tercih kümeleri, karar vericilerin gözlenen tutumları, bireysel tercih-davranış modeli ve davranış özelliklerinin ana kütle içerisindeki dağılımı gibi bileşenlerden oluşmaktadır (McFadden, 1974: 106-107). Kategorik tercih modelleri, bir karar vericinin birbirinden bağımsız ve tüm olasılıkların listelendiği bir sonlu alternatifler kümesinden bir tercih yapma olasılığını tahmin etmektedir. Karar verici kavramı; bir birey, bir grup, bir hükümet ya da bir kurum olabilmesine rağmen, aksi ifade edilmedikçe “birey” olarak temsil edilmektedir. Kategorik tercih modelleri, bir ürüne olan toplam talebin bireyler tarafından yapılan tercihlerin kümesi olarak temsil edildiği talep modelleri ile ilişkilidir (Garrow, 2010: 15).

Tüketicilerin en yüksek faydayı sağlayan alternatifleri tercih edeceği düşünülmektedir. İhtimali davranış süreci, alternatifler ile ilgi faydaların belirli dağılımlara sahip tesadüfi değişkenler ile belirlenebileceği varsayımını taşımaktadır. Bu bağlamda, kaynak kısıtlamaları altında tüketici davranışlarını açıklayan modeller genellikle en yüksek faydanın elde edilmesine dayanmaktadır. Ancak, tüketici teorisi kategorik alternatif kümelerinin söz konusu olduğu ihtimali tercihleri açıklamak için yetersiz kalmaktadır (Ben-Akiva, 1973: 14-15). Tesadüfi fayda modelleri; bireylerin tercih davranışlarını daha iyi anlamaya yardımcı olmakta, bu tercihler ile ilgili yapılan tahminlerin hassasiyetini artırmakta, daha sade model yapılarının test edilmesine olanak

sağlamakta ve bu sayede uygulamada giderek yaygınlık kazanmaktadır (Walker ve Ben-Akiva, 2002: 303).

1.4. TESADÜFİ FAYDA MODELİ

İktisadi tercihler ile ilgili üç tür bilgiden söz etmek mümkündür. Bu bilgiler, küçük alternatif kümelerinden seçilen kontrollü sınamalar; alternatif kümeleri, pazar fiyatları ve tüketicinin mevcut parasal durumuna bağlı olarak belirlenen büyük alternatif kümelerinden seçilen tüketici tercih araştırmaları ve belirli bir coğrafi bölgedeki harcanan toplam malların, gelirlerin ve pazar fiyatlarının zaman serisi değerleri olarak sıralanmaktadır (Marschak, 1959: 312). Bu noktada, iktisat ve diğer birçok bilimin temel ilgi alanlarından biri, bireylerin tercih davranışlarını tespit edebilmek ve yorumlayabilmektir. Söz konusu karar süreçlerini bireylerin gözlenen davranışları ışığında değerlendirebilmek adına modeller ya da hipotezler kurulmaktadır. Ancak, bu davranışları etkileyen bütün faktörlerin gözlenmesi mümkün olmadığından araştırmacılar için mevcut durum, karmaşık bir süreç olmaya devam etmekte ve araştırmacıları bir örneklemden elde ettikleri verileri bir bireysel davranış modeli haline getirecek istatistiksel çıkarımlar yapmaya itmektedir (McFadden, 1974: 105).

Mantık çerçevesi içerisinde ve matematiksel kanunlara dayanan tercih teorileri incelenirken cebirsel-ihtimali teoriler ayrımı, belirli-belirsiz sonuçlar ayrımı ve nihayet basit-sıralı tercih sınamaları ayrımı olmak üzere üç farklı temel ayrımdan söz etmek mümkün olacaktır. İlgili sonuçların yalnızca 0, $\frac{1}{2}$ ve 1 olduğu teoriler, cebirsel tercih teorileri olarak kabul edilmektedir. Bu durum kural olarak doğru olsa bile, cebirsel teorilerin ihtimali tercih teorilerinin özel bir durumu olduğu anlamına gelmemektedir. Çünkü cebirsel modeller ihtimali modellerden farklı matematiksel araçlar yardımı ile çözümlenmektedir. Son zamanlarda, bazı yazarlar ihtimali tercih teorileri, “tesadüfi teoriler” olarak adlandırmakta ve bu tanımlama literatürde giderek yaygınlık kazanmaktadır. Bir başka ifade ile sürekli ya da kesikli zaman parametrelerini içeren bir olasılık süreci, tesadüfi süreç olarak adlandırılmaktadır. Bu bağlamda, tanım olarak zaman kavramının açıkça söz konusu olduğu teoriler tesadüfi teoriler, zaman kavramının olasılıkları değiştirmediği teoriler ise ihtimali teoriler olarak genişletilebilmektedir (Luce ve Suppes, 1965: 256).

Klasik tercih teorisi, birey davranışını iki aşamalı ve yinelemeli bir sürecin sonucu olarak açıklamaktadır. Birinci aşama, birey ve ilgili tercih kümesi arasındaki sorunlardan kaynaklanmaktadır. İkinci aşama, karar kuralı sınıflarının tanımlanması ve tercih kümesi yapısının oluşturulmasıdır (Manski, 1977: 8). Kategorik bağımlı değişkene sahip modellerde tercihler, sınırlı sayıda olası durum ve kesikli alternatifler arasından yapılmaktadır. Bireysel tercihlerin alternatif açıklaması, tesadüfi fayda teorisi aracılığı ile yapılmaktadır (Akın, 2002: 8-9). Tesadüfi fayda teorisi, tüketicilerin nadiren tatmin edilebilen, sıklıkla çok sınırlı ve gözlenemeyen heterojen tercihlerine açıklık getirmektedir (Mazzanti, 2003: 586). Tesadüfi fayda modelleri, olasılıklı tercih modellerinin bir alt kümesidir ve ilk olarak bireylerin davranış kalıplarında gözlenen tutarsızlıkları tanımlamak adına psikologlar tarafından geliştirilmiştir. Bu modellerin ekonometrik açıklaması McFadden tarafından yapılmıştır (Manski, 1977: 229).

McFadden, tüketicilerin en büyük algılanan faydayı elde etme amacıyla “rasyonel” oldukları aksiyomundan yola çıkmıştır. Bir başka deyişle, tüketiciler söz konusu amaca ulaşmak adına yapacakları harcamalarda iktisadi sınırlamalar olduğunun farkında olacaklardır. Ayrıca, algılanan fayda tesadüfi bir fonksiyon olarak varsayılmaktadır. Söz konusu seçime ait alternatiflerin $j = 1, \dots, J$ olarak tanımlandığı ve j alternatifinin, ölçülen tutumları ifade eden z_j ile (tesadüfi) faydayı ifade eden $u(z_j)$ ’nin bir vektörü olduğu kabul edildiğinde, tercih edilen alternatif i ’nin olasılığı,

$$P(i | z) = P [u \in E^J \mid u(z_i) \geq u(z_j), j \neq i] \quad (1.30)$$

şeklinde ifade edilmektedir. Burada, u fayda fonksiyonunun dağılımına ait düğüm olasılıklarının sıfıra eşit ve $z = (z_1, \dots, z_J)$ olduğu varsayılmıştır. Kesikli bir alternatif ile ilişkili tesadüfi fayda fonksiyonu $u(z_j)$, tüketicinin bütçesine ve sabit alternatif j ’ye göre mümkün olan en yüksek fayda değeri olarak yorumlanmaktadır (McFadden, 1980: 14-15). Nitekim tesadüfi fayda modellerinin yaygın olarak kullanılmalarının ve ilgi uyandıran modeller olmalarının altında da “her bir tüketicinin tesadüfi olmayan en yüksek faydaya ulaşma amacında olması” varsayımına imkân tanınması yatmaktadır (Brown ve Walker, 1989: 815). Örneğin, yukarıda açıklanan klasik iktisadi rasyonelliğe sahip tüketici bir yerleşim yeri tercihinde bulunurken mevcut her bir alternatifi değerlendirme eğiliminde olacaktır. Sözü edilen tüketici kapsamlı bir araştırma yaparak yerleşim yerinin ulaşılabilirliği, alışveriş merkezlerine, kamu kurumlarına ve okullara

uzaklığı; konut fiyatları, vergiler ve toplu taşıma ücretleri; konutların inşaat yılı, konuttaki oda sayısı gibi birçok alternatif arasından en yüksek faydayı elde edebileceği alternatifi tercih edecektir (McFadden, 1978: 75).

Tesadüfi fayda modeli, tesadüfi bileşenin fayda fonksiyonuna doğrudan dâhil edildiğini varsaymaktadır. Tesadüfi fayda modellerinin genel kabul görmüş yorumunda, bireyin kendi faydasını neyin etkileyeceğini bilmesine rağmen, ekonometrik sorgulayıcının tüm ilişkili faktörleri gözleyemeyeceği ifade edilmektedir (Brown ve Walker, 1989: 816). Bir başka ifade ile bu modellerde, analiz eden kişinin eksik bilgiye sahip olduğu ve bu sebeple belirsizliğin dikkate alınması gerektiği varsayılmaktadır. Bu modellerde; gözlenemeyen alternatif tutumlar, gözlenemeyen bireysel özellikler, ölçüm hataları ve temsili ya da araçsal değişkenler olmak üzere dört farklı belirsizlik kaynağı tanımlanmaktadır (Üçdoğruk, 2002: 9; Manski ve Lerman, 1977: 1979).

Bir tesadüfi fayda modelinde, modeli analiz eden kişinin tercihleri etkileyen bütün faktörleri gözlemleyebilmesi mümkün olmadığından en yüksek faydaya sahip gizli (latent) değişkenlerin belirlenmesi adına izlenecek en işlevsel yöntem tesadüfi ve tesadüfi olmayan iki bileşene sahip bir fayda fonksiyonunu göz önünde bulundurmak olacaktır. Sözü edilen fayda fonksiyonunda, tesadüfi bileşen tercih kararını etkileyen gözlenemeyen tutumları; tesadüfi olmayan bileşen ise bireylerin ölçülen tutumlarını yansıtmaktadır (Baltas ve Doyle, 2001: 115). Tercih edilen alternatiflerin olasılığını ifade eden (1.30) eşitliğinin ışığında, bir bireyin fayda fonksiyonu,

$$U = V(z, x) + \varepsilon(z, x) \quad (1.31)$$

şeklinde ifade edilmektedir. Bu eşitlikte, V değişkeni tesadüfi olmayan bir değişkendir ve toplumun “temsili” beğenilerine etki etmektedir. Bununla birlikte, ε değişkeni tesadüfi bir değişkendir ve sözü edilen bireyin kişisel beğenilerini yansıtmaktadır (McFadden, 1974: 108). (1.31) eşitliğinde sözü geçen U fayda fonksiyonunun aşağıdaki özelliklere sahip olduğu varsayılmaktadır:

1. U fayda fonksiyonu, sürekli bir fonksiyondur.
2. $U(a_j) = 0$ 'dır. Bu eşitlikte, a_j değişkeni, bireyin tercihte bulunduğu ürünün büyüklük, ağırlık, renk, miktar gibi temel özelliklerini temsil etmekte ve $a_j \neq 0$ olması durumunda ilişkisizlik özelliği göstermektedir.

3. $U(a_j + \Delta a_j) \geq u(a_j)$
4. U fayda fonksiyonunun tekdüze herhangi bir dönüştürülmüş şekli de bir fayda fonksiyonudur (Quandt, 1956: 512).

Bir A değeri, sonlu bir alternatif kümesini ve U değeri A kümesine bağlı bir fayda fonksiyonunu temsil etmek üzere, A kümesi içerisindeki her x değişkeni için $U(x)$ fonksiyonu tesadüfi bir değişken; U değeri ise A kümesi içerisinde yer alan bir tesadüfi vektör olacaktır. Tesadüfi değişkenler arasında herhangi bir bağımlılık varsayımının olmadığı düşünüldüğünde $U(x)$ fonksiyonun olasılığı aşağıdaki eşitlikteki gibi tanımlanmaktadır:

$$P[U(x) \geq U(y), y \in Y] = \int_{-\infty}^{\infty} P[U(x) = t, U(y) \leq t, y \in Y] dt \quad (1.32)$$

Tesadüfi değişkenlerin bağımsız olması durumunda, (1.32) eşitliğinin sağ tarafı aşağıdaki gibi daha basit hale gelecektir:

$$\int_{-\infty}^{\infty} P[U(x) = t] \prod_{y \in Y - \{x\}} P[U(y) \leq t] dt \quad (1.33)$$

(1.32) ve (1.33) eşitliklerini ve yukarıda bahsi geçen modeli özetlemek gerekirse, bir tesadüfi fayda modeli sonlu bir A alternatif kümesinin bütün alt kümeleri için tanımlanan tercih olasılıkları kümesidir. Bir başka ifade ile bir tesadüfi fayda modeli, $x \in Y \subset A$ olmak üzere, aşağıdaki gibi tanımlanmaktadır:

$$p_Y(x) = P[U(x) \geq U(y), y \in Y] \quad (1.34)$$

Yalnızca iki değerli tercih olasılıklar söz konusu olduğunda, (1.34) eşitliği aşağıdaki gibi yeniden yazılabilmektedir:

$$p(x, y) = P[U(x) \geq U(y)] \quad (1.35)$$

Tesadüfi fayda modelinde yer alan U tesadüfi vektörü bağımsız tesadüfi değişken bileşenlerinden oluşuyorsa, model bağımsız tesadüfi fayda modeli olarak adlandırılacaktır (Luce ve Suppes, 1965: 337-338).

A ve B gibi iki tercih alternatifinin sırasıyla n ve m sayıda temel özelliklere sahip oldukları varsayıldığında ve ayrıca sözü edilen iki alternatife ait her zaman tercih sürecine dâhil edilen r ve s gibi iki özelliğinin bulunduğu düşünüldüğünde, herhangi bir zamanda $n - r$ sayıda gerekli olmayan özelliğin modele dâhil edilme olasılığı

bulunmaktadır. Tüketicilerin gerekli olmayan bu özellikleri çeşitli şekillerde dikkate alabileceği bilindiğinden 2^{n-r} sayıda birleşik özelliğin düşünülmesi gerekmektedir. Bu durumda, belirli bir zamanda A alternatifinin fayda indeksi,

$$\sum_{i=1}^r w_i U(a_i) + w_j U(a_j) \quad j = r + 1, \dots, r + 2^{n-r} \quad (1.36)$$

şeklinde yazılmaktadır. Bu eşitlikte, a_j ($j \geq r + 1$) birleşik özellikleri ve w ağırlık değerlerini temsil etmekte olup toplam ağırlık değerleri,

$$\sum_{j=1}^{2^{n-r}+r} w_j = 1 \quad (1.37)$$

şeklinde tanımlanmaktadır. Benzer şekilde, B alternatifinin fayda indeksi,

$$\sum_{i=1}^s v_i U(b_i) + v_k U(b_k) \quad k = s + 1, \dots, s + 2^{m-s} \quad (1.38)$$

şeklinde yazılmaktadır ve v ağırlık değerlerini temsil etmektedir. En yüksek faydayı elde etmeyi amaçlayan bir tüketicinin tek bir j ve k değeri için A alternatifini tercih etme durumu ise,

$$\sum_{i=1}^r w_i U(a_i) + w_j U(a_j) > \sum_{i=1}^s v_i U(b_i) + v_k U(b_k) \quad (1.39)$$

şeklinde ifade edilebilmektedir (Quandt, 1956: 513-514).

Tercih modeli kurulurken öncelikle, bir bireyin tercih ettiği bir alternatifin ona sağlayacağı gözlenemeyen faydayı modellemek gerekmektedir (Özer, 2004: 70). Tesadüfi fayda modellerinde kesikli alternatiflere ait tercihler, ürün faydaları olarak da adlandırılan gizli (latent) indekslerin ortaya çıkarılması ile belirlenmektedir. Bireysel tercihler alternatiflerin özelliklerine ve tüketicilerin beğenilerine bağlı olmaktadır. Bu noktada, bir tesadüfi tercih modeli gözlenen özellikler ile tercihler arasında bir eşleştirme yapma görevi üstlenmektedir (Baltas ve Doyle, 2001: 115). Bir X tercih alternatifleri kümesinde yer alan gerçek değerli tesadüfi bir fonksiyon olan U değeri, her bir $x \neq y$ için,

$$P(U_x \geq U_y) = p_{xy} \quad (1.40)$$

olasılık koşulunu sağladığı durumda, iki değerli manada tesadüfi fayda indeksi olarak tanımlanmaktadır. Öte yandan, U değeri iki değerli manada tesadüfi fayda indeksini temsil etmek üzere, her bir $x \neq y$ için, $P(U_x = U_y) = 0$ olmaktadır (Marschak, 1959: 314).

Bir ihtimali tercih sistemi, biçimsel olarak $(I, Z, \xi, \mathcal{B}, S, P)$ şeklinde bir vektör vasıtasıyla tanımlanmaktadır. Sözü edilen vektörde I değeri indeks alternatifleri kümesini; Z değeri ölçülen tutumlara ait alternatifleri yansıtan vektörlerin bir evrensel kümesini; $\xi: I \rightarrow Z$ değeri gözlenen tutumlara ait alternatifleri gösteren bir eşleştirme değerini; $\mathcal{B}$ değeri I kümesinden elde edilen sonlu, boş olmayan tercih kümelerini; S değeri bireylere ait ölçülen özellikleri yansıtan vektörlerin bir evrensel kümesini ve $P: I \times \mathcal{B} \times S \rightarrow [0, 1]$ değeri ise tercih olasılığını ifade etmektedir. En yüksek tesadüfi fayda hipotezi ise biçimsel olarak (I, Z, ξ, S, μ) şeklinde bir vektör vasıtasıyla tanımlanmakta; (I, Z, ξ, S, μ) vektörü ihtimali tercih sisteminde olduğu gibi ifade edilmekte ve μ değeri ise I kümesine ait fayda fonksiyonların yer aldığı bir uzayda, $s \in S$ ifadesine bağlı olan bir olasılık ölçüm değerini temsil etmektedir. Sözü edilen μ olasılığı, $s \in S$ özelliğindeki bireyler topluluğuna ait beğenilerin dağılımını vermektedir. Bu bağlamda, $B = \{i_1, \dots, i_n\} \in \mathcal{B}$, $s \in S$ ve $k = 1, \dots, n$ için (I, Z, ξ, S, μ) ifadesine ait her bir tesadüfi fayda modeli ve $(B \in \mathcal{B})$ şeklindeki her bir tercih kümesi,

$$P(i_k | B, s) = \mu(\{U \in R^I \mid U(i_k) \geq U(i_j) \text{ } j = 1, \dots, n \text{ için}\}, s) \quad (1.41)$$

eşleştirmesi vasıtasıyla bir ihtimali tercih sistemi oluşturmaktadır. Sözü edilen varsayım ise tercihlerin en yüksek faydayı elde etme amacıyla yapıldığını açıklamaktadır (McFadden, 1981: 202-203).

İKİNCİ BÖLÜM

İKİ DURUMLU MODELLER

2.1. İKİ DURUMLU MODELLERİN TANIMI

Sosyal bilimler, bireylerin davranışlarını açıklamayı ve tahmin etmeyi amaçlamaktadır. Bu amaç doğrultusunda, pratikte bireylerin kararlarını ya da tercihlerini tahmin etme yoluna gidilmektedir. Birçok durumda, bireylerin tercihleri sürekli olasılıklar etrafında yinelenmektedir. Ancak birçok farklı durumda ise, söz konusu tercihler bir takım sınırlı olasılıklar ya da alternatiflerden oluşmaktadır. Bu tür tercihlere kesikli ya da iki durumlu tercihler denilmektedir (Hausman ve Wise, 1978: 403). Son yıllarda; meslek seçimi, iş gücüne katılma, tüketim mallarının marka ve satın alma seçimi, hane halkı ve iş yeri seçimi, alış veriş yeri seçimi gibi birçok tercih ve karar verme durumunda iki durumlu tercihler söz konusu olmaktadır. Bu tür durumlarda araştırmacılar, tercih olasılıklarının açıklayıcı değişkenlere ve bireysel tercih davranışları teorilerine bağlı olarak nasıl modelleneceği ve elde edilen gözlemlerin tahminlerinin nasıl yapılacağı gibi sorular ile karşı karşıya kalmaktadır (McFadden, 1976: 363-364).

İki durumlu tercih modellerinin karar birimleri temelde bireylerdir. Bu noktada, her birey için 1 (başarılı) ya da 0 (başarısız) olmak üzere tek bir sınama söz konusudur. Bu tür bir sınama, daha önce de ifade edildiği gibi Bernoulli sınaması olarak adlandırılmaktadır (Powers ve Xie, 2000: 41-42). Bir iki durumlu tercih modelinde, bağımlı değişken y_t , yalnızca 0 ve 1 değerlerini alabilmektedir. Dış kaynaklı ve önceden belirlenmiş değişkenlerden oluşan Ω_t bilgi kümesinde, $y_t = 1$ olması durumunda, P_t , bağımlı değişkenin olasılığını ifade etmekte ve iki durumlu tercih modeli, sözü edilen koşullu olasılığı aşağıdaki denklik vasıtasıyla tahmin etmektedir:

$$P_t \equiv P(y_t = 1 | \Omega_t) = E(y_t | \Omega_t) \quad (2.1)$$

Bu sayede; iki durumlu tercih modelleri, koşullu tahmin modellemesi olarak düşünülmektedir. Birçok bağımlı değişken için koşullu beklentileri modellemek için regresyon modeli kullanılsa dahi, mevcut durumda regresyon modelinden yararlanmak mantıklı olmayacaktır. X_t , hemen hemen her zaman bir sabit terim ya da eşdeğere sahip Ω_t bilgi kümesindeki değişkenlere ait k uzunluğunda bir sıra vektörü olmak üzere, bir doğrusal regresyon modeli $E(y_t | \Omega_t)$ beklenti fonksiyonunu $X_t\beta$ şeklinde açıklayacaktır.

Ancak, $E(y_t | \Omega_t)$ ifadesi bir olasılığı ifade ettiğinden ve bu sebeple $0 \leq E(y_t | \Omega_t) \leq 1$ eşitsizliğinin sağlanması gerektiğinden, bu tür bir doğrusal regresyon modeli bu eşitsizlik koşulunu sağlamada yetersiz kalacaktır. Belirli bir örneğin tüm gözlemleri için bu koşul sağlansa bile 0'dan küçük ya da 1'den büyük değer alabilecek tahmin edilen olasılık $X_t \hat{\beta}$ için, X_t değerlerini bulmak daima kolay olacaktır. Tahmin edilen olasılıkların negatif değer alması ya da 1'den büyük olmasının bir anlamı olmadığından, iki durumlu değişkenin koşullu beklentisini belirlemek için doğrusal regresyon modelini kullanmak kabul edilebilir bir seçim olmayacaktır. Bununla birlikte, bu tür bir regresyon modeli, ampirik araştırmanın daha sonraki aşamalarında faydalı bilgiler sağlayabilecektir (Davidson ve MacKinnon, 1999: 444).

İki durumlu tercih modellerinde amaç, tercih davranışlarının verilmiş bir setinden yararlanarak bireyin alternatifi yerine yapacağı belirli bir seçimin olasılığını belirleyebilmektir. Daha genel ifadeyle, iki durumlu tercih modellerinin temel kullanılma sebebi, bireyi tasvir eden davranışların bir veri kümesi ile o bireyin bir tercih yapma olasılığı arasındaki ilişkinin bulunmasıdır (İşyar, 1999: 258). İki durumlu değişkenler yalnızca 0 ya da 1 değerlerini alabildiklerinden ve ayrıca çoklu regresyon modelleri, bağımsız değişkenin beklenen değer etrafında seyreden dağılımının bu beklenen değer düzeyinden bağımsız olmasını varsaydığından iki durumlu değişkenler için uygun modeller olmayacaktır (Tobin, 1955: 2). Böylece, iki durumlu değişkenlere uygun alternatif modeller türetilmiştir. İki durumlu tercih modelleri; kendi arasında iki durumlu doğrusal olasılık modeli, iki durumlu probit modeli, iki durumlu logit modeli ve tamamlayıcı log-log model olarak dört gruba ayrılmaktadır (Long, 1997: 34).

2.1.1. İki Durumlu Doğrusal Olasılık Modeli

İki durumlu doğrusal olasılık modeli, doğrusal regresyon modelinin bir iki durumlu bağımlı değişkene uygulanmasıdır. Bu model, bir bireyin iki değerli tercih modelinde bir yapma olasılığının o bireyin bireysel özelliklerinin doğrusal bir fonksiyonu olduğunu varsaymaktadır (İşyar, 1999: 258). Bu bağlamda, bağımsız değişkenlerin değerleri veri iken, koşullu bağımlı değişkenin beklenen değerleri, meydana gelen olayın koşullu olasılıkları olarak açıklanabilmektedir. Bir açıklayıcı değişkenli doğrusal olasılık modelinde y_i , 0 veya 1 değerini alırken, $y_i = \beta x_i +$

u_i denkleminde hata terimleri $(1 - \beta x_i)$ ve $(-\beta x_i)$ değerlerini alabilmektedir (Akın, 2002: 16). Böylece, genel doğrusal regresyon modeli, koşullu beklenen değer olarak,

$$E(y_i | x_{i1}, \dots, x_{ip}) = \sum_{k=0}^p \beta_k x_{ik}, \quad i = 1, \dots, n \quad (2.2)$$

biçiminde yazılabilmektedir. Bu modelde açıklayıcı değişkenler üzerinde kısıt yok iken, y bağımlı değişkeni sürekli olmak zorundadır. Herhangi bir i 'nci gözlem için,

$$y_i = \sum_{k=0}^p \beta_k x_{ik} + u_i \quad (2.3)$$

eşitliğindeki gibi ifade edilen modelde, açıklayıcı değişkenler üzerinde bir kısıt olmadığından y_i bağımlı değişkeni, $-\infty$ ile $+\infty$ arasında tüm değerleri alabilmektedir (Tatlıdil, 2002: 289-290).

Bir doğrusal olasılık modelinin yapısı, (2.4) eşitliğinde özetlendiği gibi ifade edilebilmektedir:

$$y_i = x_i \beta + \varepsilon_i \quad (2.4)$$

Bu eşitlikte, x_i , i 'nci gözlem değerlerinin bir vektörünü; β , parametrelerin bir vektörünü ve ε , hata terimini ifade etmektedir. Ek olarak, iki durumlu modellerin genel mantığında olduğu gibi $y = 1$ bir olayın gerçekleşmesi, $y = 0$ bir olayın gerçekleşmemesi durumunu örneklemektedir. Tek bir bağımsız değişken olması durumunda, (2.4) eşitliği aşağıdaki gibi yeniden yazılabilecektir:

$$y_i = \alpha + \beta x_i + \varepsilon_i \quad (2.5)$$

Benzer şekilde, (2.5) eşitliğindeki x değerinin koşullu beklenen değeri, $E(y | x) = \alpha + \beta x_i$ olacaktır. Doğrusal regresyon modeli için koşullu beklenen değerler, (2.6) eşitliğinde gösterildiği gibi olmaktadır (Long, 1997: 35):

$$E(y_i | x_i) = [1 \times P(y_i = 1 | x_i)] + [0 \times P(y_i = 0 | x_i)] = P(y_i = 1 | x_i) \quad (2.6)$$

Doğrusal olasılık modellerinde, ana kütle oranları P_i ile örnek oranları p_i arasındaki ilişki, $i = 1, 2, \dots, T$ olmak üzere, $p_i = P_i + \varepsilon_i$ şeklinde varsayılmaktadır. Bu eşitlikte,

ε_i 'ler birbirinden bağımsız olup, ortalamaları sıfır, varyansları $\frac{P_i(1-P_i)}{n_i}$ dir ve binom dağılımına uymaktadır (Genceli, 1987: 217).

2.1.1.1. İki Durumlu Doğrusal Olasılık Modelinin Tahmini

Doğrusal olasılık modeli için hata terimlerinin normal dağılımı söz konusu olmamaktadır, çünkü Y_i 'nin farklı değer alması aşağıdaki gibi iki farklı hata terimi değeri ortaya çıkarmaktadır (Kutlar, 2009: 193):

$$u_i = Y_i - \beta_1 - \beta_2 X_i \quad (2.7)$$

$$Y_i = 1 \quad u_i = 1 - \beta_1 - \beta_2 X_i \quad (2.8)$$

$$Y_i = 0 \quad u_0 = \beta_1 - \beta_2 X_i \quad (2.9)$$

Gözlem sayısının fazla olması durumunda, doğrusal olasılık modellerinin parametrelerinin tahmini, klasik doğrusal regresyon modelinde olduğu gibi En Küçük Kareler yöntemi ile yapılabilmektedir (Güriş ve Çağlayan, 2010: 661). En Küçük Kareler yöntemi ile tahmin edilen olasılıklar birim aralık üzerinde sınırlanmamakta ve En Küçük Kareler tahmini değişen varyanslar sorunun var olduğu anlamına gelmektedir (Horrace ve Oaxaca, 2006: 321). Ayrıca, tahmin edilen parametreler en iyi doğrusal sapmasız tahminci olma özelliğini taşımadığından En Küçük Kareler yöntemi her ne kadar sapmasız tahminçiler verse de, diğer varsayımları sağlamadığı için kullanımı uygun olmayacaktır (Güriş ve Çağlayan, 2010: 662).

2.1.1.1.1. İki Durumlu Doğrusal Olasılık Modelinin Tartılı En Küçük Kareler Yöntemi İle Tahmini

D , T büyüklüğünde olasılıklı ya da olasılıklı olmayan pozitif değere sahip herhangi bir köşegen matris olmak üzere, $w_t = w_s$ eşitliğinin sağlandığı her durumda, D matrisinin t 'inci ve s 'inci köşegen elemanları aynı olmaktadır. Bu durumda, doğrusal olasılık modelindeki α değerinin Tartılı En Küçük Kareler tahminçilerinin oluşturduğu $\tilde{\alpha}(D)$ sınıfı aşağıdaki eşitlikteki gibi yazılabilecektir:

$$\tilde{\alpha}(D) = (W' D^{-1} W)^{-1} W' D^{-1} y \quad (2.10)$$

En Küçük Kareler ve Gauss-Markov tahminçileri olan $\tilde{\alpha}(I)$ ve $\tilde{\alpha}(A)$ değerleri de söz konusu sınıfa dâhil olacaktırlar (Amemiya, 1977: 645-646).

Klasik regresyon modellerinde değişen varyans problemini ortadan kaldırmak için uygulanan Tartılı En Küçük Kareler yöntemi, doğrusal olasılık modelinde de aynı amaçla, ancak iki aşamada uygulanmaktadır. Doğrusal olasılık modelinde parametreler önce En Küçük Kareler yöntemi ile tahmin edilir ve her bir gözlem değeri için bir tartı oluşturulur. $\sqrt{w_i} = \sqrt{P_i(1 - P_i)} = \sqrt{E(Y_i)[1 - E(Y_i)]}$ şeklindeki tartılar belirlendikten sonra regresyon denkleminin her iki tarafı bu tartıya bölünür. Sonuçta, elde edilen hata terimleri sabit varyansa sahip olacaktır (Güriş ve Çağlayan, 2010: 662). Bu tartılar, $\hat{\sigma}$ ile ifade edilmektedir. Ancak, klasik En Küçük Kareler yöntemi ile tahmin edilen $\hat{y}_i$ değişkeninin 1'den büyük ve 0'dan küçük değerleri hesaplanamayacaktır (Baltagi, 1997: 334).

Tartılı En Küçük Kareler yönteminde üç tür sorunla karşılaşmak mümkündür.

1. $\hat{y}_i(1 - \hat{y}_i)$ değeri $E(Y_i)[1 - E(Y_i)]$ ifadesinin tutarlı bir tahmincisi olduğu sürece, büyük örneklemelerde çok küçük olasılıkla olsa bile, $\hat{y}_i(1 - \hat{y}_i)$ tahmin değeri pratikte negatif olabilecektir.
2. Hata terimleri u_i 'ler normal dağılmadığı sürece, klasik anlamlılık testlerinin uygulanmasında sorunlarla karşılaşılabilir. Bu sorun, açıklayıcı değişkenlerin çok değişkenli normal dağılıma sahip olması ile ifade edilebilir.
3. Tartılı En Küçük Kareler eşitliğinin kendi içerisindeki en büyük eksikliği, $E(y_i | x_i)$ koşullu tahmin değerinin birçok durumda (0-1) sınırları dışında değer alabilmesidir (Maddala, 1992: 324).

2.1.1.1.2. İki Durumlu Doğrusal Olasılık Modelinin Kısıtlı En Küçük Kareler Yöntemi İle Tahmini

Bir En Küçük Kareler problemi için standart Lagrange amaç fonksiyonu eşitliği aşağıdaki gibi tanımlanmaktadır:

$$\text{Min}_{b^*} L(b^*, \lambda) = (y - Xb^*)'(y - Xb^*) + 2\lambda'(Rb^* - r) \quad (2.11)$$

Bu eşitlikte, b^* değişkeni Kısıtlı En Küçük Kareler tahmincisini; λ değişkeni, Lagrange çarpanlarının $q \times 1$ büyüklüğünde bir vektörünü ve R değişkeni ise $q \times k$ büyüklüğünde bir vektörü temsil etmektedir. Değişkenlerin birinci dereceden türevi alındığında aşağıdaki eşitlikler elde edilmektedir:

$$\partial L / \partial b^* = -2X'y + 2X' Xb^* + 2R' \lambda = 0 \quad (2.12)$$

$$\partial L / \partial \lambda = 2Rb^* - 2r = 0 \quad (2.13)$$

(2.12) ve (2.13) eşitlikleri, parçalı matris eşitliği şeklinde yazıldığında,

$$\begin{bmatrix} X'X & R' \\ R & 0 \end{bmatrix} \begin{bmatrix} b^* \\ \lambda \end{bmatrix} = \begin{bmatrix} X'y \\ r \end{bmatrix} \quad (2.14)$$

ifadesi elde edilmektedir. Nihai eşitliğin çözümüyle, Kısıtlı En Küçük Kareler tahmincisi olan b^* değişkenini hesaplamak mümkündür (Greene ve Seaks, 1991: 564).

Kısıtlı En Küçük Kareler yöntemi, bir ya da birkaç parametrenin kesin değeri ya da aralarındaki kesin doğrusal ilişki hakkında bilgi sahibi olunması durumunda kullanılmaktadır. Bu yöntemin kısıtlayıcı En Küçük Kareler yöntemi olarak tanımlanmasının sebebi bir kısıtlanmış ilişkiye klasik En Küçük Kareler yönteminin uygulanmasıdır. Diğer bir deyişle, bu yöntemde yapısal ilişkinin hata kareleri toplamı olan $\sum e^2$, $b = b^*$ kısıtlaması altında en küçüğe indirgenmektedir (Koutsoyiannis, 1992: 402-403). Kısıtlı En Küçük Kareler yönteminde amaç, yapılacak tahminlerin (0-1) sınırları arasına düşmesini sağlamaktır. Bu yöntemde, kısıtlı tahmincilerin varyansları kısıtsız tahmincilerle ait eşdeğerlerinden daha küçük olacaktır. Bu sayede, parametreler için anlamlılık testleri yaparken, kısıtsız tahmincilerin varyansları üst sınır olarak kullanılabilir. Kısıtlı En Küçük Kareler yöntemi ile yapılan tahminler sapmalı tahminler olacaktır (Güriş ve Çağlayan, 2010: 663).

2.1.1.1.3. İki Durumlu Doğrusal Olasılık Modelinin En Yüksek Olabilirlik Yöntemi İle Tahmini

İki durumlu doğrusal olasılık modelinin Tartılı En Küçük Kareler yöntemi ile tahmininde tanımlanan (2.10) eşitliğindeki değerler ışığında log-olabilirlik fonksiyonu aşağıdaki şekilde olacaktır:

$$L = \sum_{t=1}^T [y_t \log \alpha' w_t + (1 - y_t) \log (1 - \alpha' w_t)] \quad (2.15)$$

Log-olabilirlik fonksiyonunun α değerine göre türevinin alınması ile

$$\frac{\partial L}{\partial \alpha} = \sum [y_t (\alpha' w_t)^{-1} - (1 - y_t) (1 - \alpha' w_t)^{-1}] w_t \quad (2.16)$$

eşitliği elde edilmektedir ve tanım gereği en yüksek olabilirlik $\hat{\alpha}$ tahmincisi, (2.16) eşitliğinin sağ tarafının sıfıra eşitlendiği durumda bulunan α değerini ifade etmektedir. En yüksek olabilirlik tahmincisi, α_1 birinci tahminci ve α_2 ikinci tahminci olmak üzere, yinelemeli olarak skrolama yöntemi² ile tahmin edilebilmektedir. Bu noktada, α_2 ise aşağıdaki gibi tanımlanmaktadır:

$$\hat{\alpha}_2 = \{\sum [\hat{\alpha}'_1 w_t (1 - \hat{\alpha}'_1 w_t)]^{-1} w_t w'_t\}^{-1} \sum [\hat{\alpha}'_1 w_t (1 - \hat{\alpha}'_1 w_t)]^{-1} w_t y_t \quad (2.17)$$

Eğer sözü edilen yinelemeli yöntem $\hat{\alpha}$ en yüksek olabilirlik tahmincisi değerine yakınsar ise, $\hat{\Lambda}$ değeri t 'inci değeri $\hat{\alpha}' w_t (1 - \hat{\alpha}' w_t)$ ifadesine eşit olan bir köşegen matrisi olmak üzere, $\hat{\alpha} = \tilde{\alpha}(\hat{\Lambda})$ eşitliği elde edilmekte ve böylece en yüksek olabilirlik tahmincisi aynı zamanda En Küçük Kareler tahmincisinin bir elemanı olarak değerlendirilmektedir. Bu noktada aşağıdaki çıkarımları yapmak mümkün olmaktadır:

1. En yüksek olabilirlik ve tutarlı Tartılı En Küçük Kareler tahmincilerinin asimtotik dağılımları aynıdır
2. En yüksek olabilirlik ve En Küçük Kareler tahmincileri (ancak ve ancak) özdeş ise bu iki tahmincinin asimtotik dağılımları da aynı olacaktır (Amemiya, 1977: 646, 648).

2.1.1.1.4. Doğrusal Olasılık Modelinin Tahmininde Karşılaşılan Sorunlar

Doğrusal olasılık modelinin en önemli eksikliği, $x_i' \beta$ değerinin bir olasılık değerinde olması gerektiği gibi 0 ve 1 arasında sınırlandırılmış olmamasıdır. Bu eksiklik, $F = 1$ iken $F(x_i' \beta) > 1$ ve $F = 0$ iken $F(x_i' \beta) < 0$ şeklinde düzeltilebilmesine rağmen, söz konusu yöntem tepe noktalarında gerçekçi olmayan bükülmelere sebep olmaktadır (Amemiya, 1981: 1486). Bununla birlikte, doğrusal olasılık modelinde parametrelerin yorumlanması aşamasında doğrusal regresyon modelinin birçok varsayımı ihlal edilmektedir. Doğrusal olasılık modelinin tahmininde karşılaşılan sorunlardan birkaçını aşağıdaki gibi sıralamak mümkündür.

2.1.1.1.4.1. Değişen Varyans (Heteroscedasticity) Sorunu

Klasik doğrusal regresyon modelinin önemli varsayımlarından biri, açıklayıcı değişkenlerin seçilmiş değerlerinin koşullu sonucu olan her bir u_i bozucu teriminin

²Skrolama yöntemi, tüm türevlerin α_1 değerine göre alındığı kabul edildiğinde, $\hat{\alpha}_2 = \hat{\alpha}_1 - [E \left(\frac{\partial^2 L}{\partial \alpha \partial \alpha'} \right)]^{-1} \partial L \partial \alpha$ şeklinde tanımlanan yineleme olarak ifade edilmektedir.

varyansının σ^2 'ye eşit sabit bir sayı olduğunu ileri sürmesidir. Bu varsayım, sabit varyans (homoscedasticity) varsayımı olarak adlandırılmakta ve $i = 1, 2, \dots, n$ iken, $E(u_i^2) = \sigma^2$ eşitliği ile ifade edilmektedir. Ancak, X değerleri büyüdükçe, Y_i değerlerinin varyansı sabit kalmadığı için değişen varyans (heteroscedasticity) durumu ortaya çıkmakta ve bu durum $E(u_i^2) = \sigma_i^2$ eşitliği ile gösterilmektedir (Gujarati, 2011: 356). μ ortalamalı tesadüfi bir iki durumlu değişkenin varyansı $\mu(1-\mu)$ olacaktır. Belirlenen x değişkeninin beklenen değerinin $x\beta$ 'ye eşit olması durumunda, y değerinin koşullu varyansı aşağıdaki (2.18) eşitliğine göre, y 'nin beklenen değeri olan x 'e bağlı olacaktır.

$$\text{Var}(y | x) = P(y = 1 | x)[1 - P(y = 1 | x)] = x\beta(1 - x\beta) \quad (2.18)$$

Bu durum, varyans hatalarının sabit olmadığını işaret etmektedir. Doğrusal regresyon modeli değişen varyanslı olduğu için, En Küçük Kareler tahmincisi olan β değeri etkisiz kalacak, böylece standart hatalar sapmalı olacağından, elde edilen test istatistikleri hatalı sonuçlar verecektir (Long, 1997: 38). Olasılık değeri 0 ya da 1'e yakın gözlemler nispeten küçük varyanslı, buna karşılık olasılık değeri 0,5'e yakın gözlemler daha büyük varyanslı olacaktır. Değişen varyans sorunu, Tartılı En Küçük Kareler yöntemi vasıtasıyla çözülebilmektedir (Özer, 2005: 66).

2.1.1.1.4.2. Normal Dağılım Sorunu

x_* , y 'nin beklenen değeri x 'in özel bir değeri olarak kabul edildiğinde, y değeri yalnızca 0 ve 1 olabileceği için hata terimi ε , $\varepsilon_1 = 1 - E(y | x_*)$ ya da $\varepsilon_0 = 0 - E(y | x_*)$ değerlerine eşit olabilecektir. Görüldüğü gibi, modelin hata terimleri normal dağılım varsayımını ihlal edecektir (Long, 1997: 39). Standart En Küçük Kareler tahmin edicileri sapmasızlıklarını koruduklarından ve örneklem büyüklüğü sonsuza doğru artarken sözü edilen tahmin ediciler normal dağılıma uyma eğilimi gösterdiklerinden normallik varsayımının ihlal edilmesi, sapmasız en iyi tahminler yapılmasını engellememektedir (Gujarati, 2011: 543).

2.1.1.1.4.3. Uyum İyiği Ölçüsü Olarak R^2 'nin Kuşkulu Değeri

Geleneksel yöntemle hesaplanan R^2 değeri, sınırlanmamış ve 0-1 mantıksal sınırlar içerisinde sınırlanmış doğrusal olasılık modellerde 1'den çok küçük çıkma eğilimindedir. Bu sebeple, kategorik bağımlı değişkene sahip modellerde, belirlilik

katsayısının bir özetleme istatistiği olarak kullanılmasından kaçınılması gerektiği önerilmektedir (Gujarati, 2011: 546).

2.1.1.1.4.4. Anlamsız Tahminler Sorunu

Doğrusal regresyon modeli, y değerlerini negatif ya da birden büyük olarak tahmin etmektedir. Bu durum, $E(y | x) = P(y = 1 | x)$ şeklindeki olasılık tahminleri için anlamsız sonuçlara sebep olabilecektir (Long, 1997: 39). Bu sorunun ortaya çıkma sebebi, doğrusal olasılık modelinin söz konusu eşitlikteki değeri X 'in doğrusal bir fonksiyonu olduğunu ve dolayısıyla X_i 'nin marjinal etkisinin sabit kaldığını varsaymasıdır. Gerçek yaşamda karşılaşılan bu durum, P_i 'nin X 'in doğrusal olmayan bir fonksiyon olduğunu ve X_i arttıkça olasılık değerinin de 0-1 aralığının dışına çıkmayacak şekilde arttığını, X_i küçüldükçe 0'a giderek daha yavaş yaklaştığını ve X_i büyüdükçe 1'e daha yavaş yaklaştığını gösteren bir olasılık modeli ile daha iyi ortaya konulabilecektir (Özer, 2005: 69).

Doğrusal olasılık modelindeki sorunlara çözüm üretmek adına, kısıtlayıcı doğrusal olasılık modeli önerilmiştir. Bu model, doğrusal olasılık modeli eşitliği ile X değişkeni arasındaki doğrusal ilişkiyi koruyarak, anılan eşitliği aşağıda gösterildiği gibi birim aralıklara kısıtlamaktadır:

$$\pi = \begin{cases} 0 & 0 > \alpha + \beta X \text{ ise} \\ \alpha + \beta X & 0 \leq \alpha + \beta X \leq 1 \text{ ise} \\ 1 & \alpha + \beta X > 1 \text{ ise} \end{cases} \quad (2.19)$$

Doğrusal olasılık modelinde olduğu gibi, kısıtlayıcı doğrusal olasılık modelinin tahmininde de birçok sorunla karşılaşılmaktadır. En temelde, $\pi = 0$ ve $\pi = 1$ noktalarının eğim değerlerindeki ani değişiklikler anlamlı değildir. Bununla birlikte, anılan modeli tahmin ederken $\pi = \alpha + \beta X$ eşitliği, 0 ve 1 noktalarına bağlı olarak belirlendiği için, bu iki nokta arasına denk gelen X değerlerini bulma hususunda sorunlar yaşanmaktadır. Sonuç olarak, modelin tahmini kendi içerisinde değişkenlik göstermektedir. Birden çok X değeri olması durumunda, kısıtlı doğrusal olasılık modelinin tahmini çok daha zor olmaktadır (Fox, 1997: 442-443).

Doğrusal olasılık modelinin eksikliğini gidermek için koşullu olasılık durumunu koruyarak 0'dan küçük ve 1'den büyük değerler almayacak şekilde birer alt ve üst bükülmeye sahip olacak yeni bir doğrusal fonksiyon eşitliği yazmak gerekmektedir

(Horowitz ve Savin, 2001: 44). Bir ekonometrik modelde, Y bağımlı değişkeni sadece 0 ve 1 değerlerini alıyorsa, birikimli dağılım fonksiyonunu kullanarak tahmin yapılabilmektedir. Her hata terimi u için bir birikimli dağılım fonksiyonu söz konusu olduğundan, burada hangi birikimli dağılım fonksiyonunun kullanılacağına tespit edilmesi gerekmektedir. Uygulamalarda en çok kullanılan iki birikimli dağılım fonksiyonu, logit ve probit modellerdir (Akkaya ve Pazarlıoğlu, 1998: 87).

2.1.2. İki Durumlu Logit Modeli

Bağımlı bir modelde, bağımlı değişken sınıflama ölçeği ile ölçüldüğünde, En Küçük Kareler tekniğine alternatif olarak diskriminant ve lojistik regresyon modelleri yer almaktadır. Lojistik regresyon, bağımsız değişkenin normal dağılıma uyması ve bu değişkenin ortak değişkeninin her grup düzeyinde eşit olması gibi varsayımları gerektirmediği için bağımlı değişkeni iki ya da daha çok kategoriden oluşan bir modelde diskriminant analizine üstünlük kurmaktadır (Albayrak, 2008: 273). İlk olarak 1845 yılında Verhulst tarafından bir nüfus büyümesi modeli olarak ortaya atılan lojistik regresyon, 1920 yılında Pearl ve Reed tarafından tekrar ele alınmıştır. Logit model ve lojistik çıkarımlar Luce (1959) ve Berkson tarafından geliştirilmiştir (Cramer, 2003: 165; Train, 2003: 38), Lojistik regresyon, parametrik olmayan ileri bir istatistiksel yöntemdir ve bağımlı değişkenin mutlaka iki durumlu sonuca sahip olmasını gerektirmektedir. Lojistik regresyon yönteminde zorunluluk olmamakla birlikte, bağımsız değişkenler genellikle sürekli olmaktadır (Akgül ve Çevik, 2003: 390).

Lojistik regresyon modelini doğrusal regresyon modelinden ayıran en önemli fark, lojistik regresyon modelinde bağımlı değişkenin iki durumlu değişken olmasıdır. İki model arasındaki bu fark, hem parametrik modelin seçiminden hem de varsayımların belirlenmesinden etkilenmektedir (Hosmer ve Lemeshow, 2000: 1). Lojistik regresyon analizinin temel odak noktası, bireylerin hangi grubun üyesi olduğunu tahmin etmede kullanılacak bir regresyon denklemi oluşturmaktır. Bir başka ifadeyle, lojistik regresyon analizinde amaç bireylerin farklı gruplara ayrılarak sınıflandırılmasıdır (Çokluk, Şekercioğlu ve Büyüköztürk, 2010: 50). Günümüzde, iki durumlu ve çok bağımsız değişkenler ile ilgilenen birçok araştırmacı, gruplar arası farklara sahip veri setleri ile çalıştıklarından lojistik regresyonun kullanımı oldukça yaygın hale gelmiştir (Hardy, 1993: 76).

Bağımlı değişken iki durumlu olduğunda, tercih fonksiyonunun doğrusal olmayan yapıya sahip olması ve fonksiyon eğrisinin S-şeklinde seyretmesi beklenmektedir (Montgomery, Peck ve Vining, 2012: 423). Söz konusu S-şeklindeki eğrinin artış ya da azalış oranını β parametresinin işareti belirlemektedir ve $|\beta|$ değeri arttığında değişim oranı da artmaktadır (Agresti, 1996: 104). Hata teriminin (u) dağılımı³, probit modelinde standart normal birikimli dağılım şeklindeyken, logit modelinde lojistik birikimli dağılım söz konusu olmaktadır (Maddala, 1992: 327-328). Normal dağılım ile lojistik dağılım arasındaki en temel fark, lojistik dağılım eğrisinin kuyruklarında daha fazla ağırlık olmasıdır (Johnson ve DiNardo, 1997: 424).

Bir iki durumlu bağımlı değişken Y ve nicel bağımsız değişken X için, $\pi(x)$ değeri söz konusu bağımsız değişkenin x değerini aldığı durumdaki başarı olasılığını ifade etmek üzere; lojistik regresyon modeli, $\pi(x)$ olasılığının logit değeri için doğrusal bir yapıya sahip olmakta ve (2.20) eşitliğindeki gibi yazılabilmektedir:

$$\text{logit} [\pi(x)] = \log \left(\frac{\pi(x)}{1 - \pi(x)} \right) = \alpha + \beta x \quad (2.20)$$

Lojistik regresyon için doğrudan başarı olasılığını ifade eden alternatif bir formül aşağıdaki eşitlikteki gibi tanımlanmaktadır:

$$\pi(x) = \frac{\exp(\alpha + \beta x)}{1 + \exp(\alpha + \beta x)} \quad (2.21)$$

Bu eşitlikteki $\exp(x)$ değeri, e^x üstel değerine eşit olarak kabul edilmektedir. Söz konusu eşitlik, doğrusal yaklaşım ile yorumlandığında; modelde β değeri sıfıra eşitlendiğinde, (2.21) eşitliğinin sağ tarafı da sabit kalmaktadır. Bu durumda; $\pi(x)$ değeri, tüm x 'ler için eşit olacak, S-şeklindeki lojistik fonksiyon eğrisi yatay düz bir çizgi haline gelecek ve böylece Y bağımlı değişkeni, X değişkeninden bağımsız⁴ olacaktır (Agresti, 1996: 103-104). Ayrıca, (2.21) eşitliğindeki $\pi(x)$ küçük bir değer

³ Bir logit modelinde standart lojistik dağılıma sahip hata terimlerinin ortalaması sıfır ve varyans değeri $\pi^2/3$ 'tür. Bu değer in seçilme sebebi hata terimine ait birikimli dağılım fonksiyonun $\Lambda(u) = \frac{\exp(u)}{1 + \exp(u)}$ şeklinde olmasıdır (Long, 1997: 42).

⁴ Birçok tercih durumunda, tercihler üzerinde belirli bir dönemde etkili olan gözlenemeyen faktörlerin en azından bir ya da birkaç dönem boyunca etkilerini sürdürmeleri ve tercihlerin bağımlı olması özelliğinin belirli bir zaman süresince devam etmesi beklenmektedir. Logit model, belirli bir zaman süresince devam eden tercihlerin sonuçlarına uygulandığında bağımsızlık varsayımı dikkate alınacaktır (Train, 2003: 22).

aldığında, β değeri x_i 'nin birim oransal artışını; $1-\pi(x)$ değeri küçük bir değer aldığında ise β değeri x_i 'nin her birim artışındaki oransal azalışını ifade edecektir. Bütün $\pi(x)$ değerlerinin, $\pi(x)^0$ gibi bir değere yakın seyrettiği düşünüldüğünde, (2.21) eşitliği $\pi(x)$ ve x_i arasında $\beta[\pi(x)][1-\pi(x)]$ eğime sahip doğrusal bir ilişkiyi göstermiş olacaktır (Cox, 1958: 216). β değeri sıfırdan küçük ya da büyük olduğunda lojistik fonksiyonun birikimli dağılım eğrisi de sırasıyla azalma ya da artma eğilimi gösterecektir (Zelterman, 2006: 69).

Logaritmik dönüşüm ile birlikte, doğrusal regresyon modelinin daha arzu edilen özellikleri elde edilmiş olacaktır. Logit değeri, sahip olduğu parametreler ölçüsünde doğrusal ya da sürekli olabilmekte ve x değerinin aralığına göre $-\infty$ ile $+\infty$ arasında değerler alabilmektedir (Hosmer ve Lemeshow, 2000: 6). Bunun yanında, (2.20) eşitliğinde $\log\pi(x)$ değeri X değişkenine göre doğrusal olmakla olasılıklar sözü edilen değişkene göre doğrusal olmamaktadır. Bu özellik, olasılıkların X ile birlikte doğrusal olarak arttığı doğrusal olasılık modeli ile lojistik regresyon modeli arasındaki farkı ortaya koymaktadır (Gujarati, 2011: 555).

Lojistik regresyon sonuçlarını değerlendirmek ve ortaya konulan ilişkilerin yorumlanması adına yeni bir ölçüm değerinin tanımlanması gerekmekte ve bu ölçüm değerine bahis (odds) oranı denilmektedir. Yukarıda verilen (2.20) eşitliğinin sol tarafı log-bahis oranını, sağ tarafı ise bağımlı değişkeni etkilemesi muhtemel ortak değişkenler ya da risk faktörlerinin doğrusal bir fonksiyonunu ifade etmektedir (Zelterman, 2006: 69). Bu eşitlikte $x = 1$ durumundaki bahis değerinin, $x = 0$ durumundaki bahis değerine bölünmesiyle bahis oranı elde edilmektedir. Diğer bir ifadeyle bahis oranı (OR),

$$OR = \frac{\pi(1)/[1-\pi(1)]}{\pi(0)/[1-\pi(0)]} \quad (2.22)$$

şeklinde tanımlanacaktır. Bahis oranının temelinde görel risklerin (RR) yaklaşık ölçümü vardır. Görel risk ise $\pi(1)/\pi(0)$ oranına eşittir. (2.22) eşitliğinde görüldüğü gibi görel risk, $[1-\pi(1)]/[1-\pi(0)]$ oranının bir olması durumunu ifade etmektedir. Bu bağlamda, küçük olasılık değerleri için görel risk ve bahis oranı değerleri birbirine yakın olacaktır (Hosmer ve Lemeshow, 2000: 49-50).

Lojistik regresyon modeli ile ilgili bir diğer yaklaşım, modelin bahis değeri ve bahis oranı ile yorumlanmasıdır. Buna göre, söz konusu eşitlikteki model için bahis değerinin başarısı,

$$\frac{\pi(x)}{1 - \pi(x)} = \exp(\alpha + \beta x) = e^{\alpha}(e^{\beta})^x \quad (2.23)$$

şeklinde olacaktır. (2.23) eşitliğindeki üstel ilişki β değerinin yorumlanmasına olanak tanımaktadır. Eşitlikteki x değerindeki bir birimlik her artışa karşılık, bahis değerinde e^{β} kadar bir artışa karşılık gelmektedir. Bir başka ifadeyle, $x + 1$ düzeyindeki bahis değeri $x.(e^{\beta})$ değerine eşit olmaktadır. Bununla birlikte, $\beta = 0$ iken, $e^{\beta} = 1$ olacağından, x değerleri değiştiğinde bahis değeri değişmeyecektir (Agresti, 1996: 107). Bağımsız değişkendeki bir birim artışın olasılıktaki değişim üzerindeki etkisi,

$$\begin{aligned} \Delta P &= P(D = 1 | L_1) - P(D = 1 | L_0) \\ &= \exp(L_1) / [1 + \exp(L_1)] - \exp(L_0) / [1 + \exp(L_0)] \end{aligned} \quad (2.24)$$

şeklinde dir. Bu eşitlikte, L_0 ve $L_1 = L_0 + \beta_j$ değerleri sırasıyla x_j değişkenindeki birim değişimden önceki ve sonraki logit değerlerini temsil etmektedir (Petersen, 1985: 131).

Lojistik regresyon değerlendirmelerinde bahis oranı kullanılmasının iki üstünlüğü vardır. Birincisi, üstel değerler ışığında yapılan yorumlarda, söz konusu değerlerin 1'den büyük olması bahis değerini artırırken, 1'den küçük olması azaltmaktadır. İkincisi, hesaplama tüm gözlemler üzerinden yapılmakta, buna rağmen diğer değişkenlerin aldığı değerlere bağımlı olmamaktadır. Böylece, diğer değerlendirme yöntemlerinde karşılaşılan karmaşıklar önlenmektedir (Hoetker, 2007: 334).

2.1.2.1. İki Durumlu Logit Modelinin Tahmini

Logit modelinin tahmin edilebilmesi için tekrarlı veya tekrarsız gözlemlerin ayırt edilmesi gerekmektedir. Tekrarlı durumda seçilecek yöntemler genelleştirilmiş En Küçük Kareler ve En Küçük χ^2 yöntemleri, tekrarsız durumda seçilecek yöntem en yüksek olabilirlik yöntemidir. Logit modelleri En Küçük Kareler yöntemi ile tahmin etmek her ne kadar mümkün olsa da söz konusu yöntem ile tahmin edilen parametreler etkin olmayacaktır. Tartılı En Küçük Kareler yöntemi kullanıldığında ise etkinlik sorunu aşılmasına rağmen parametreler doğrusal olmayacaktır. (Güriş ve Çağlayan, 2010: 668).

2.1.2.1.1. İki Durumlu Logit Modelinin En Küçük χ^2 Yöntemi ile Tahmini

(2.20) eşitliğindeki α ve β değerlerinin logit modeli için en küçük χ^2 ile tahmini için aşağıdaki eşitlikler çözülmektedir:

$$\sum n_i \frac{\hat{p}_i q_i + \hat{q}_i p_i}{\hat{p}_i \hat{q}_i} (p_i - \hat{p}_i) = 0 \quad (2.25)$$

$$\sum n_i \frac{\hat{p}_i q_i + \hat{q}_i p_i}{\hat{p}_i \hat{q}_i} x_i (p_i - \hat{p}_i) = 0 \quad (2.26)$$

Bu eşitliklerde, n_i değeri x_i değişkenindeki gözlem sayısını; p_i değeri gözlenebilen n_i değerlerinin oranını; $\hat{p}_i$ değeri P_i olasılığının tahminini; q_i değeri $1-p_i$ değerini ve $\hat{q}_i$ değeri $1-P_i$ olasılığının tahminini temsil etmektedir (Berkson, 1955: 132).

2.1.2.1.2. İki Durumlu Logit Modelinin En Yüksek Olabilirlik Modeli İle Tahmini

Hata teriminin (u) dağılımı simetrik olduğunda ve $F(Z) = 1 - F(-Z)$ eşitliği sağlandığı sürece,

$$P_i = F \left(\beta_0 + \sum_{j=1}^k \beta_j x_{ij} \right) \quad (2.27)$$

olasılık eşitliği yazılabilmektedir. Her bir sınamada x_{ij} değerine göre değişen ve binom dağılımı gösteren gözlenen y_i değerleri, bu eşitlikteki olasılıkların gerçekleşme durumlarını ifade ettiği sürece; olabilirlik fonksiyonu,

$$L = \prod_{v_i=1} P_i \prod_{v_i=0} (1 - P_i) \quad (2.28)$$

şeklinde yazılabilmektedir (Maddala, 1992: 328). (2.20) eşitliğindeki α ve β değerlerinin logit modeli için en yüksek olabilirlik tahminleri aşağıdaki iki eşitliğin çözümü ile elde edilmektedir (Berkson, 1955: 132):

$$\sum n_i (p_i - \hat{p}_i) = 0 \quad (2.29)$$

$$\sum n_i x_i (p_i - \hat{p}_i) = 0 \quad (2.30)$$

2.1.3. İki Durumlu Probit Modeli

Probit modeli, ilk olarak 1920'lerde, canlı organizmaların uyarıcılara verdikleri kesikli yanıtları tarif etmek için kullanılmıştır. Kökeni biyoloji ve tıp bilimlerine dayanan bu modelin ilk uygulama alanı, bir tarım ilacının dozunun böcekler üzerinde uyarıcı olduğunu kanıtlamak adına yapılan bir biyolojik deney olmuştur (Cramer, 1986: 161). İki durumlu probit model; Aitchison ve Silvey (1957) tarafından sıralama ölçeği ile ölçülmüş bağımlı değişkenler için, Aitchison ve Bennet (1970) tarafından kategorik değişkenler için kullanılmak üzere genişletilmiştir (McKelvey ve Zavoina, 1975: 103). Tesadüfi bir değişken $-\infty$ ile $+\infty$ arasında bütün değerleri aldığı anda, bu değişkenin normal dağılıma sahip olduğu söylenebilmektedir. Matematiksel bir ifadeyle, anılan değişkene ait dağılım merkezinden herhangi bir d uzaklığındaki frekans değerinin logaritmasının dağılım merkezindeki d^2 uzaklığındaki bir frekansın logaritmasından küçük olması gerekmektedir. Bu durumda, dağılım simetrik olacak ve dağılım merkezindeki en büyük frekans değeri elde edilecektir (Finney, 1971: 9; Fisher, 1969: 12).

Probit analizi, birikimli standart normal dağılımı kullanarak, S şeklindeki ilişkiyi modellemek için alternatif bir iki durumlu tercih modeli yaklaşımıdır. Bu özelliği sayesinde, probit modeli, zaman zaman normit model olarak da adlandırılmaktadır. Probit analizinde, olasılığı belirleyen değişkenlerin doğrusal bir fonksiyonu olan Z değişkeni aşağıdaki eşitlikteki gibi tanımlanmaktadır:

$$Z = \beta_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (2.31)$$

Burada, standart hale getirilmiş birikimli normal dağılım olan $F(Z)$, Z 'nin herhangi bir değeri için bir olayın gerçekleşmesi olasılığını vermekte ve bu olasılık $p_i = F(Z_i)$ şeklinde ifade edilmektedir (Dougherty, 2007: 10).

Probit analizi, F fonksiyonunun çok değişkenli normal olması varsayımıyla türetilmiştir ve bu fonksiyonun standart normal hale getirilmesi probit modeli ile mümkündür. Bir probit modeli aşağıdaki eşitlikteki gibi tanımlanmaktadır:

$$P(y = 1) = \phi(X_i \beta) = \int_{-\infty}^{X_i \beta} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz \quad (2.32)$$

Bu eşitlikteki $\phi(\cdot)$ değeri, olasılığın 0 ile 1; ya da $\lim_{z \rightarrow +\infty} \phi(z) = 1$ ve $\lim_{z \rightarrow -\infty} \phi(z) = 0$ arasında olmasını sağlayan bir kısıtlayıcıdır (Johnston ve Dinardo, 1997: 418; Train, 2003: 20). Yukarıdaki (2.32) eşitliğindeki $X_i\beta$ değeri probit skoru ya da indeksi olarak adlandırılmaktadır. Probit katsayısı β , modelin tahminindeki bir birimlik yükselmenin probit skorunda yapacağı β standart sapmalık yükselmeyi ifade etmektedir (Kayış, 2008, 301). Probit modelinde tanımlanan F fonksiyonunun marjinal etkileri aşağıdaki eşitlikteki gibi tanımlanmaktadır (Baltagi, 1997: 335):

$$\partial \phi(x_i' \beta) / \partial x_k = \phi_i \beta_k \quad (2.33)$$

Yukarıdaki eşitlikte, parçalı etkilerin hesaplanmasının yanında ϕ değişkeni de değerlendirmeye alınarak, tahmin ediciler üzerinde etkisi olan değişkenler ile ilgili fikir sahibi olmak mümkündür. En yüksek olabilirlik modelinin değişmeyen varyans özelliği sayesinde, tahmin edilen parçalı etkiler ve olasılıklar kendiliğinden tutarlı ve asimtotik etkin olacaktır (Knapp ve Seaks, 1992: 406).

Probit modelin yaygın olarak kullanılma sebeplerinden bir tanesi, gözlenemeyen ya da gizli bir değişkene sahip bir modelden kolayca türetilbilmesidir. y_t^* , gözlenemeyen değişkenin,

$$y_t^* = X_t \beta + u_t, \quad u_t \sim \text{NBD}(0,1) \quad (2.34)$$

şeklinde tanımlandığı varsayılırsa; bu durumda, y_t^* değişkeninin yalnızca işareti hakkında fikir sahibi olmak mümkündür. y_t^* değişkeninin işareti vasıtasıyla da gözlenebilen iki durumlu değişken olan y_t 'nin değeri aşağıdaki ilişkiye göre tespit edilebilmektedir:

$$y_t^* \leq 0 \text{ ise } y_t = 0 \quad (2.35)$$

$$y_t^* > 0 \text{ ise } y_t = 1 \quad (2.36)$$

y_t^* değişkeninin yalnızca işareti tespit edilebildiği için, normal ve bağımsız dağıldığı kabul edilen u_t değerinin varyansını 1 olacak şekilde normalleştirmek mümkündür. Söz konusu varyans başka bir değer alırsa, örneğin σ^2 olursa, β , y_t^* ve u_t değerlerini σ^2 'ye bölmek gerekmektedir (Davidson and MacKinnon, 1999: 445-446). Probit modellerin tesadüfi tercih değişimine ve her türlü değişiklik kalıbına imkân vermesi ve panel veri setlerine de uygulanabilir olması, diğer kesikli tercihli modellere göre üstünlükleri

olarak kabul edilmektedir. Probit modelinin tek kısıtlayıcı varsayımı, bütün gözlenemeyen bileşenlerin normal dağılım göstermesi gerekliliğidir (Train, 2003: 101).

2.1.3.1. İki Durumlu Probit Modelinin Tahmini

Probit modelinde doğrusal olmayan normal birikimli dağılım fonksiyonu ile tahmin yapılmaktadır. Söz konusu tahminler, seçeneklerden “en çok fayda sağlayanı seçilir” prensibine dayanmaktadır ve tahmin yapılabilmesi için fonksiyonun doğrusal hale getirilmesi gerekmektedir (Güriş, Çağlayan ve Güriş, 2011: 383). Probit model, birikimli normal dağılımın tersi alınarak doğrusal hale getirilmektedir. $F^{-1}(P_i)$ ifadesi, $\hat{P}_i$ etrafında Taylor serisine göre açıldığında, aşağıdaki denklik elde edilmektedir:

$$\begin{aligned} F^{-1}(\hat{P}_i) &\equiv \beta' X_i + \left. \frac{\delta F^{-1}}{\delta P_i} \right|_{P_i} (\hat{P}_i - P_i) \\ &= \beta' X_i + \frac{1}{f[F^{-1}(P_i)]} (\hat{P}_i - P_i) \\ &\equiv \beta' X_i + v_i \end{aligned} \quad (2.37)$$

$F^{-1}(\hat{P}_i) = \beta' X_i + v_i$ denklemi, değişen varyanslı hata terimli doğrusal bir regresyon denklemini vermektedir. $\hat{P}_i$ değeri, $V(v_i)$ eşitliğine konularak β parametresi aşağıdaki gibi tahmin edilmektedir:

$$\hat{\beta} = \left[\sum_{i=1}^N \frac{n_i f^2[F^{-1}(\hat{P}_i)]}{\hat{P}_i (1 - \hat{P}_i)} \cdot x_i x_i' \right]^{-1} \cdot \sum_{i=1}^N \frac{n_i f^2[F^{-1}(\hat{P}_i)]}{\hat{P}_i (1 - \hat{P}_i)} \cdot x_i F^{-1}(\hat{P}_i) \quad (2.38)$$

Söz konusu bu süreç, probit modelin gruplandırılmış verilerde ağırlıklı En Küçük Kareler yöntemi ile tahminini ifade etmektedir (Akın, 2002: 27). Büyük örnekler için, $F^{-1}(P_i)$ ifadesi,

$$F^{-1}(P_i) = \beta_0 + \beta_1 X_i + \frac{\varepsilon_i}{f(\beta_0 + \beta_1 X_i)} \quad (2.39)$$

şeklinde olacaktır. Bu eşitlikte, β_0 ve β_1 değerleri yerine $\hat{\beta}_0$, $\hat{\beta}_1$ ve P_i değerleri kullanılarak tahmin yapılabilmektedir ve elde edilen tahminler tutarlı tahminler olacaktır. Bu durumda, hata teriminin beklenen değeri sıfır ve varyansı $\frac{P_i(1-P_i)}{n_i[f(\beta_0 + \beta_1 X_i)]^2}$ şeklindedir. (Güriş ve Çağlayan, 2010: 674-675).

2.1.3.1.1. İki Durumlu Probit Modelinin En Küçük Kareler Yöntemi ile Tahmini

Probit model için doğrusal olmayan En Küçük Kareler, araç değişkenler ve en yüksek olasılık yöntemleri olmak üzere üç farklı tahmin ediciden söz etmek mümkündür. Doğrusal olmayan En Küçük Kareler yönteminde amaç,

$$u_i = y_i - F_i, \quad F_i = F(v_i), \quad v_i = c_1 + c_2 x_i \quad (2.40)$$

olmak üzere, $\phi = \phi(c_1, c_2) = \sum_i u_i^2$ ölçütünün en küçük değerini elde etmektir. $f(.)$ fonksiyonu, standart normal olasılık yoğunluk fonksiyonu olarak kabul varsayıldığında, bu fonksiyon,

$$f_i = f(v_i) = \partial F_i / \partial v_i \quad (2.41)$$

şeklinde ifade edilebilecektir. Bu durumda, c_1 ve c_2 değişkenlerine göre türevler,

$$\partial \phi / \partial c_1 = 2 \sum_i u_i (\partial u_i / \partial c_1) = -2 \sum_i u_i f_i \quad (2.42)$$

$$\partial \phi / \partial c_2 = 2 \sum_i u_i (\partial u_i / \partial c_2) = -2 \sum_i u_i f_i x_i \quad (2.43)$$

olarak tanımlanacaktır. Böylece, birinci dereceden koşullu türevler aşağıdaki eşitliklerdeki gibi sıfıra eşit olacaktır:

$$\sum_i u_i f_i = 0, \quad \sum_i u_i f_i x_i = 0 \quad (2.44)$$

Sonuç olarak, θ_1 ve θ_2 gibi iki doğrusal olmayan tahmin edici değeri elde etmek için anılan eşitlikler çözülmelidir. Benzer şekilde, araç değişkenler vasıtasıyla probit model tahmininde, c_1 ve c_2 değişkenlerinin, $\sum_i u_i / n = 0$ ve $\sum_i x_i u_i / n = 0$ koşulunu sağlayan değerlerini hesaplanmak gerekmektedir. Önceki yöntemden farklı eşitlikler çözülerek, yeni θ_1 ve θ_2 tahmin edici değerleri elde edilmektedir (Goldberger, 1991: 144-145).

Probit analizinin tahmininde En Küçük Kareler yöntemi kullanıldığında bazı sorunlar ortaya çıkmaktadır:

- 1) Açıklayıcı değişken iki değerli olduğunda, hata terimi normal dağılım göstermediğinden, parametrelerin dağılımı da normal dağılım göstermemektedir.

Özellikle, örnek çapı (serbestlik derecesi) küçük olduğunda, parametrelerin anlamlılık testleri doğru sonuçlar vermeyebilecektir.

- 2) Y bağımlı değişkeni için yapılacak tahminler, X değerlerine bağlı olarak 0 ve 1'in dışında değerler alabilmektedir. Bu durumda, negatif değerli tahminler ve 1'den büyük tahminler, sırasıyla 0 ve 1 değerlerine eşitlenmektedir.
- 3) $Y = \beta X + \varepsilon$ modelinde, açıklayıcı değişkenin hata teriminden bağımsız olduğu kabul edilmektedir. Ancak, Y bağımlı değişkeni 0 ve 1 değerlerini alabildiğinden, Y 'nin beklenen değeri $E(Y) = bX$ ifadesine göre fark terimi ve hata teriminin varyansı,

$$Y_i = 0 \text{ için } e_i = -bX \quad (2.45)$$

$$Y_i = 1 \text{ için } e_i = 1 - bX \quad (2.46)$$

$$E(ee') = E(e^2) = (X'b)(1-X'b) = E(Y)[1-E(Y)] \quad (2.47)$$

olacaktır. Böylece, hata terimleri Y 'nin beklenen değeri ve X 'e bağımlı olacağından değişen varyans sorunu ortaya çıkacaktır (Öztürkcan, 2009: 77-78).

2.1.3.1.2. İki Durumlu Probit Modelin En Yüksek Olabilirlik Yöntemi ile Tahmini

Probit modelinde verilen X değeri için Y değerinin koşullu dağılımı tayin edilebildiği için, en yüksek olabilirlik yöntemi ile de tahmin mümkün olmaktadır. Y değişkeni, $E(Y | X) = F(\theta_1 + \theta_2 X)$ koşulunu sağlayan bir iki durumlu değişken olduğundan, bu değişken $F(\theta_1 + \theta_2 X)$ parametresi ile birlikte Bernoulli dağılımına uymaktadır. θ_1 ve θ_2 parametreleri X değişkeninin marjinal dağılımında yer almadıkları için, en yüksek olabilirliği elde edebilmek için, en yüksek koşullu olabilirliğin değerine bakılmaktadır (Goldberger, 1991: 144-145).

Probit modellerin için en yüksek olabilirlik yöntemi kullanıldığında, doğrusal olmayan parametreler tahmin edileceğinden, bu tahminler ancak Newton-Raphson ve Skorum Yöntemi gibi tekrarlı yöntemleri kullanılarak yapılabilmektedir. En yüksek olabilirlik ve düzeltilmiş En Küçük Kareler tahmincileri asimtotik özelliğe sahiptir. En yüksek olabilirlik yöntemi ile elde edilen tahmincilerin tek tek anlamlılığı test edilmek istenildiğinde, t testinden faydalanılmakta ve F testi yerine Wald veya en yüksek olabilirlik oranı kullanılmaktadır (Güriş ve Çağlayan, 2010: 675). Probit

modeller için, en yüksek olabilirlik fonksiyonu aşağıdaki gibi tanımlanmaktadır (DeMaris, 2004: 255):

$$L(\beta | y, x) = \prod_{i=1}^n \phi(z_i)^{y_i} [1 - \phi(z_i)]^{1-y_i} \quad (2.48)$$

En yüksek olabilirlik yöntemi ile probit modelinin parametrelerinin tutarlı tahminlerini elde etmek mümkün olmasına rağmen, söz konusu modeli kullanmanın teorik gerekçesinin de kısıtlı olması sebebiyle, logit modeller araştırmacılar tarafından daha cazip görülmektedir (İşyar, 1999: 267).

2.1.4. Tobit Modeli

İstatistiksel analiz yapılırken bazı durumlarda veriler, sansürleme (censoring) ya da kesme (truncation) sebebi ile sınırlandırılmış olabilmektedir. Sansürlenmiş veri ile bağımlı değişkenin değerlerinin bir değer aralığı ile sınırlı olması durumunda; kesilmiş veri ile ise örnekte yer alan birimlerin bir ana kütleinin alt örneğinde seçildiği durumlarda karşılaşılmaktadır. Sansürlü veride tüm birimler gözlenebilmesine rağmen, yalnızca her bir birimin değerinin belirli bir değerden küçük ya da büyük olduğu; kesilmiş veride ise örnek birimlerinin yalnızca bir alt grubun gözlem değerlerinin tümü bilinmektedir (Güriş ve Çağlayan, 2010: 679).

Sıfır değeri etrafında kesikli sıçrama özelliğine sahip bir bağımlı değişken, sınırlı bağımsız değişken olarak tanımlanmaktadır. Sınırlı bağımlı değişkene sahip modellerin ilk uygulamalarından birisi, Tobin (1958) tarafından hane halkı harcamalarının otomobil fiyatları üzerindeki etkisini tespit etmek adına ortaya koyduğu tobit modelidir. Bu sayede literatürde tobit modeli, Tobin'in probit modeli ya da sınırlı regresyon modeli olarak da adlandırılmaktadır (Ramanathan, 1989: 610). Tobit modeli en genel ifadeyle, daha öncede logit ve probit modellerde bahsedilen y_i^* gizli değişkeninin sıfırdan büyük olduğu durumda gözlenebilir, sıfırdan küçük olduğu durumlarda gözlenemez olduğunu varsaymakta ve bu durumda, gözlenen y_i değişkenini aşağıdaki eşitlikteki gibi tanımlamaktadır:

$$y_i = \begin{cases} y_i^* = \beta x_i + u_i, & \text{eğer } y_i^* > 0 \text{ ise} \\ 0, & \text{eğer } y_i^* \leq 0 \text{ ise} \end{cases} \quad (2.49)$$

$$u_t \sim \text{NBD}(0, \sigma^2)$$

Tobit modelinde amaç β ve σ değerlerini tahmin etmek olup, y^* değerlerinin bazı gözlemleri görülemediği için probit modelinin sansürlü normal regresyonu olarak görülmektedir (Maddala, 1992: 333).

Tobin (1958)'e göre; W değişkeni, alt sınırı L olan bir sınırlı bağımlı değişken ve Y değişkeni, sınırlı bağımlı değişken ile ilişkili bağımsız değişkenlerinin doğrusal bir birleşimi $(X_1, X_2, \dots, X_m)$ olmak üzere, hane halklarının davranışları W değişkenine göre farklılık göstermektedir. Bu durumda, bağımsız değişkenlerdeki ve alt sınırdaki değişimler bütünüyle hesaba katılamamaktadır ve bütün diğer değişimler; ortalaması sıfır, standart sapması σ ve normal dağılım gösteren tesadüfi bir ϵ değeri ile ifade edilmektedir. Böylece hane halkı davranışının aşağıdaki eşitliklerdeki gibi ifade edilebileceği varsayılmaktadır:

$$W = L \quad (2.50)$$

$$W = Y - \epsilon \quad (2.51)$$

$P(x)$ fonksiyonu, x değişkeni için birikimli birim-normal dağılım fonksiyonunun değerini; $Q(x) = 1 - P(x)$ şeklindeki eşitliği ve $Z(x)$ fonksiyonu, x değişkeni için $P(x)$ fonksiyonu için olasılık yoğunluk fonksiyonunu ifade etmek üzere; $W-L$ ifadesinin dağılımı, ϵ değerinin dağılımından aşağıdaki şekilde türetilenmektedir:

$$Pr(W = L \mid Y, L) = Pr(\epsilon > Y - L) = Q\{(Y - L) / \sigma\}, \quad (2.52)$$

$$Pr(W > x \geq L \mid Y) = Pr(\epsilon < Y - x) = P\{(Y - x) / \sigma\} \quad (2.53)$$

Sonuç olarak, W değişkeni için birikimli dağılım fonksiyonu aşağıdaki eşitliklerdeki gibi olacaktır:

$$F(x; Y, L) = 0 \quad (x < L), \quad (2.54)$$

$$F(L; Y, L) = Q\{(Y - L) / \sigma\}, \quad (2.55)$$

$$F(x; Y, L) = Q\{(Y - L) / \sigma\}, \quad (x > L) \quad (2.56)$$

Söz konusu birikimli dağılım fonksiyonu için olasılık yoğunluk fonksiyonu, (2.57) eşitliğindeki gibidir:

$$f(x; Y, L) = \frac{1}{\sigma} Z\{(Y - x) / \sigma\} \quad (2.57)$$

Son olarak, W değişkeninin beklenen değeri,

$$E(W; Y, L) = LQ\{(Y - L) / \sigma\} + \int_L^\infty \frac{x}{\sigma} Z\{(Y - x) / \sigma\} dx \quad (2.58)$$

$$= LQ\{(Y - L) / \sigma\} + Y \int_{-\infty}^{(Y-L)/\sigma} Z(x) dx + \sigma \int_{-\infty}^{(Y-L)/\sigma} -xZ(x) dx$$

eşitliğindeki gibi ifade edilecektir (Tobin, 1958: 25-26).

2.1.4.1. Tobit Modelinin En Yüksek Olabilirlik Yöntemi ile Tahmini

Tobit modelinin olabilirlik fonksiyonu aşağıdaki eşitlikteki gibi yazılabilmektedir:

$$L = \prod_0 [1 - \phi(x_i' \beta / \sigma)] \prod_1 \phi\left(\frac{x_i' \beta}{\sigma}\right) \quad (2.59)$$

$$= \prod_1 \phi(x_i' \beta / \sigma)^{-1} \sigma^{-1} \phi[(y_i - x_i' \beta) / \sigma]$$

Yukarıdaki (2.59) eşitliğinin sağ tarafındaki ilk ifade probit modelinin, ikinci ifade ise kesme tobit modelin olabilirlik fonksiyonunu yansıtmaktadır. Bununla birlikte, $\alpha = \beta/\sigma$ ifadesinin en yüksek olabilirlik tahmincisi $\hat{\alpha}$ değeri, (2.59) eşitliğindeki iki ifadenin en yüksek logaritmalarının alınması ile elde edilmektedir. Ancak, bu yöntem sayesinde yalnızca β/σ ifadesi tahmin edilebilmekte, β ve σ değerleri ayrı ayrı tahmin edilememektedir. En yüksek olabilirlik tahmincisi β ve σ değerlerini içeren olabilirlik fonksiyonunun bir bölümünü göz ardı ettiğinden, söz konusu tahminci tam etkin olamamaktadır (Amemiya, 1985: 366). Tobit modelin log-olabilirlik fonksiyonu ise,

$$\sum_{y_i|y_i=0} \ln \left[1 - \phi\left(\frac{x_i \beta}{\sigma}\right) \right] + \sum_{y_i|y_i>0} \left[\ln \frac{1}{\sqrt{2\pi\sigma^2}} - \frac{1}{2} \frac{(y_i - x_i \beta)^2}{\sigma^2} \right] \quad (2.60)$$

şeklinde (Johnston ve DiNardo, 1997: 437).

2.1.5. İki Durumlu Modellerde Uyum İyiliği Ölçüleri

Tahmin edilen modelin değerlendirilmesi yapılırken modelin kurulma aşamasında en iyi bilgiye sahip bir araştırmacının modelin tahmin aşaması için gerekli değişkenleri içerdiği ve sözü edilen değişkenlerin modele doğru fonksiyonel yöntemler ile dâhil edildiği varsayılmaktadır. Bu şekilde araştırmacılar, kurdukları modelin sonuç değişkenini tasvir etme etkinliğini tespit edebilmektedir. Bu durum literatürde yaygın olarak uyum iyiliği olarak bilinmektedir (Hosmer ve Lemeshow, 2000: 143). Bir başka ifade ile uyum iyiliği ölçüsü, bir modelin gözlenen veriler ile arasındaki hassasiyet

derecesini gösteren özetlenmiş bir istatistik değeri olarak tanımlanmaktadır. Bağımlı değişkenin kategorik olması durumunda, bahsi geçen hassasiyet derecesi ya hesaplanan olasılıklar ile gözlenen sonuç frekansları arasındaki uygunluk değeri ya da modelin gözlenen sonuçları öngörebilme düzeyi üzerinden değerlendirilebilmektedir (Maddala, 1983: 37-38).

Bir y parametresinin $y' = (y_1, y_2, y_3, \dots, y_n)$ olmak üzere, sonuç değişkenine ait gözlenen örnek değerlerinin vektörel şekli olduğu ve benzer şekilde bir $\hat{y}$ parametresinin $\hat{y}' = (\hat{y}_1, \hat{y}_2, \hat{y}_3, \hat{y}_n)$ olmak üzere, bir modelde tahmin edilen değerleri, bir başka ifadeyle uygun değerleri, temsil ettiği varsayıldığında, y ve $\hat{y}$ arasındaki mesafe değerini gösteren özet ölçüm değerleri sayısal olarak küçük değerler ise ve $i = 1, 2, 3, \dots, n$ olmak üzere $(y_i, \hat{y}_i)$ şeklinde gösterilen her bir iki değer söz konusu özet ölçüm değerlerine sistematik olmayan katkısı modelin hata yapısına göre daha az ise bahsi geçen modelin uygun bir model olduğu söylenebilmektedir (Hosmer ve Lemeshow, 2000: 143).

Ortak değişken örüntüleri bir modeldeki ortak değişken değerlerinin bir kümesini tarif etmektedir. Bununla birlikte, kullanılacak testlere ait serbestlik dereceleri modeldeki parametre sayılarının farklarına göre belirlenmekte ortak değişken sayısını göz ardı etmektedir. Ortak değişken örüntülerinin sayısı modelin uygunluğu değerlendirilirken dikkate alınmaktadır. Bir modelin uyum iyiliği değerlendirilirken ortak değişkenler tarafından belirlenen uygun değişkenlerin toplamı dikkate alınmaktadır. Kurulan bir modelin $x' = (x_1, x_2, x_3, \dots, x_p)$ şeklinde p bağımsız değişkene sahip olduğu ve gözlenen x değişkenleri arasındaki farklı değerlerin sayısının J gibi bir parametre ile gösterildiği varsayıldığında bazı örnekler x değişkeninin aynı değerine sahip ise bu durumda $J < n$ olmaktadır. Bu noktada, $x = x_j$ değerine sahip örneklerin sayısı $m_j (j = 1, 2, 3, \dots, J)$ ile temsil edildiğinde $\sum m_j = n$ olmaktadır. Benzer şekilde, m_j değerleri arasındaki pozitif sonuçlar y_j ile temsil edildiğinde $\sum y_j = n_1$ olacak ve bu eşitlik $y = 1$ değerine sahip toplam örnek sayısını ifade edecektir. Burada, uyum iyiliği istatistiğinin dağılımı n değerinin büyük olmasına imkân tanıyarak elde edilmektedir. Ortak değişken örüntülerinin sayısı n değeri ile birlikte artıyorsa her bir m_j değeri küçük olacaktır. Yalnızca n değerinin büyük olması koşulunu temel alan dağılım sonuçları n -asimtotik, m_j değerinin büyük olduğu dağılım sonuçları

m -asimtotik dağılım olarak adlandırılmaktadır. Bu asimtotik dağılımlar arasındaki farklar özet uyum iyiliği istatistiklerinin detaylı olarak değerlendirilmesi ile açıklanabilmektedir (Hosmer ve Lemeshow, 2000: 144-145).

2.1.5.1. Pearson Ki-kare İstatistiği ve Log-Olabilirlik Oranı Testi

Belirli bir ortak değişken için Pearson kalıntı değeri aşağıdaki eşitlikteki gibi tanımlanmaktadır:

$$r(y_j, \hat{\pi}_j) = \frac{(y_j - m_j \hat{\pi}_j)}{\sqrt{m_j \hat{\pi}_j (1 - \hat{\pi}_j)}} \quad (2.61)$$

Bu kalıntı değerine göre kurulan Pearson ki-kare istatistiği,

$$X^2 = \sum_{j=1}^J r(y_j, \hat{\pi}_j)^2 \quad (2.62)$$

şeklinde ifade edilmektedir. Öte yandan, $y_j = 0$ olduğu durumda, kalıntılara ait sapma değeri,

$$d(y_j, \hat{\pi}_j) = -\sqrt{2m_j \mid \ln(1 - \hat{\pi}_j) \mid} \quad (2.63)$$

eşitliği ile gösterilmekte ve benzer şekilde $y_j = m_j$ olduğunda sözü edilen sapma değeri,

$$d(y_j, \hat{\pi}_j) = \sqrt{2m_j \mid \ln(\hat{\pi}_j) \mid} \quad (2.64)$$

olmaktadır. Sonuç olarak kalıntılara ait özet istatistiği ise,

$$D = \sum_{j=1}^J d(y_j, \hat{\pi}_j)^2 \quad (2.65)$$

şeklinde yazılmaktadır. Buna göre, kurulan modelin her yönden doğru olduğu varsayımı altında, X^2 ve D istatistiklerine ait dağılımlar $J-(p+1)$ serbestlik derecesinde ki-kare dağılımına bağlı olmak zorundadır (Hosmer ve Lemeshow, 2000: 146).

Uyum iyiliğinin ölçümünde yaygın olarak kullanılan bir diğer test istatistik değeri,

$$G^2 = 2 \sum_{i=1}^k X_i \ln(X_i/n\pi x_{0i}) \quad (2.66)$$

şeklinde tanımlanan log-olabilirlik oranı istatistiğidir. Sözü edilen iki istatistiğin hangisinin en iyi ölçüm olduğunu belirlemek adına birçok araştırmacı; sıfır hipotezi altında sonlu örnek karşılaştırmaları yaparak, sınıf aralıklarının değiştirilmesi ile farklı alternatif hipotezler altında asimtotik ve sonlu örnekler için güç karşılaştırması yaparak, sıfır ve alternatif hipotezleri altında k değerinin artırılması ile asimtotik dağılım teorisini uygulayarak ve bilinmeyen parametrelerin tahminine izin vererek etkinlik karşılaştırması yapmıştır (Cressie ve Read, 1984: 441).

2.1.5.2. Akaike Bilgi Kriteri

Bir model, olabilirlik oranı testi dikkate alınarak incelendiğinde eğer, olabilirlik oranı test istatistiği d gibi bir değerden büyükse $L(\theta)$ olabilirlik fonksiyonu $L(\alpha)$ fonksiyonuna tercih edilmektedir. Söz konusu d değeri,

$$P[\text{LRT} > d \mid L(\alpha)] = c \quad (2.67)$$

olduğu durumda, % 5 gibi bir sabit değer göz önünde bulundurularak tespit edilmektedir. Birçok model söz konusu olduğunda, Akaike (1973) tarafından önerilen kayıp fonksiyonu,

$$W(\theta_0, \hat{\alpha}) = -\frac{2}{T} \int \left[\log \frac{L(\hat{\alpha})}{L(\theta_0)} \right] L(\theta_0) dx \quad (2.68)$$

şeklinde tanımlanmaktadır. Bu eşitlikte $\hat{\alpha}$ integraldeki sabit terimi temsil etmektedir. Ancak, $W(\theta_0, \hat{\alpha}) \geq W(\theta_0, \theta_0) = 0$ olduğu için yukarıdaki eşitlik önemli bir kayıp fonksiyon değerine sahip olabilmektedir. Bununla birlikte, W değeri θ_0 bilinmeyen parametresine bağlı olduğundan W yerine öngörücü bir değer tanımlanması gerekmektedir. Yukarıdaki eşitliği basitleştirme adına, sözü edilen eşitlik yerine Akaike bilgi kriteri (AIC) diye adlandırılan,

$$\text{AIC} = -\frac{2}{T} \log L(\hat{\alpha}) + \frac{2p}{T} \quad (2.69)$$

şeklindeki öngörü değeri tanımlanmıştır. Bu eşitlikte, p değeri α vektörünün boyutunu temsil etmektedir. AIC değerinin en küçük olduğu model en uygun model olarak kabul edilmektedir (Amemiya, 1985: 146-147).

2.1.5.3. Bayes Bilgi Kriteri

Bayes bilgi kriteri, bir modelin genel uygunluğunu değerlendirmek ve Bayesçi karşılaştırma yapmak adına bir ölçü olarak geliştirilmiştir. Buna göre, M_1 ve M_2 gibi iki model için M_2 modelinin M_1 modeline göre ardışık bahis oranları,

$$\frac{P(M_2 \mid \text{Gözlenen veriler})}{P(M_1 \mid \text{Gözlenen veriler})} \quad (2.70)$$

şeklinde ifade edilmektedir. Bu noktada, eğer M_2 modelinin gözlenen verilere göre olasılığı M_1 modeli için eşdeğer olasılıktan büyük ise M_2 modeli tercih edilecektir. Sözü edilen iki modele ait önceki bahis değerlerine ait oranın 1 olduğu varsayımı altında, ardışık olasılıkların aşağıdaki eşitlikteki gibi gösterilen Bayes faktörüne eşit olduğunu göstermek için Bayes teoremi kullanılabilir:

$$\frac{P(\text{Gözlenen veriler} \mid M_2)}{P(\text{Gözlenen veriler} \mid M_1)} \quad (2.71)$$

Bunun yanında, eğer iki model de doğru model olarak seçilemiyor ise Bayes faktörü daha iyi örnek tahminlerini veren modeli tespit edebilecektir. N sayıda gözlemin, $D(M_k)$ sapmalı ve M_S modeli karşılaştırılan bir M_k modelinin olduğu varsayıldığında, df_k değeri, örnek büyüklüğü ile M_k modelindeki parametre sayısının farkını ifade etmek üzere birinci Bayes bilgi kriteri (BIC),

$$\text{BIC}_k = D(M_k) - df_k \ln N \quad (2.72)$$

olmaktadır. Bu durumda, M_S modeline ait bilgi kriteri sıfıra eşit olduğu sürece $\text{BIC}_k > 0$ olduğunda M_S modeli tercih edilecektir. İkinci bir Bayes bilgi kriteri, olabilirlik oranı ki-kare temeline dayanmakta ve aşağıdaki eşitlikteki gibi yazılmaktadır:

$$\text{BIC}'_k = -G^2(M_k) + df'_k \ln N \quad (2.73)$$

Bu eşitlikte df'_k değeri parametreler yerine eğilimin yönünü belirleyen değişken sayısını vermektedir (Long, 1997: 110-11)

ÜÇÜNCÜ BÖLÜM

ÇOK DURUMLU MODELLER

3.1.ÇOK DURUMLU MODELLERİN TANIMI

Bireylerin tercih davranışının tesadüfi olduğu varsayıldığında, bir bireyin üç ya da daha fazla alternatifler arasından ve her bir olası alternatif kümesinden yaptığı tercihler ile ilgili elde ettiği eksiksiz veriler, bireyin hangi tercihi yaptığını tespit etmek için yeterli görünmemektedir. Çok alternatifli tercih sorunundan kaçınmak mümkün olmadığından, iktisatçılar tesadüfi modellerin bu özelliğine dikkat çekmişler ve özellikle kusurlu ayrımcılık kavramına karşı yapılan eleştirilerde söz konusu özelliği önemli bir kaynak olarak kullanmışlardır (Luce, 1959: 3). Kategorik bağımlı değişkene sahip model yapısının tasnifi yapılırken bağımsız değişkeni etkileyen faktörlerin olasılıkların dikkate alınması ve alternatif modeller için uygun tahmin tekniklerinin tespit edilmesi gerekmektedir. Bununla birlikte, kategorik bağımlı değişkene sahip doğrusal modeller, çoklu ya da birbirinden farklı birtakım olaylar için ortak sonuçlar içeren kategorik değişkenler ile ölçülmüş ise, söz konusu modelleri tahmin etmek güçleşmektedir (Hanushek ve Jackson, 1977: 187).

Bağımlı değişkenin iki durumlu olduğu modeller, anılan değişken için sınıflama ölçeği ile ölçülmüş değerlere sahip ikiden fazla durumun söz konusu olduğu modellere de kolaylıkla uygulanabilmektedir. Bu tür modellere çok durumlu modeller ya da multinomial regresyon modelleri denilmektedir. Çok durumlu modellerde genel amaç, durumların bahis değerlerini ortak değişkenlerin bir fonksiyonu olarak modellemek ve sonuçları farklı durumların tercihleri için bahis değerleri açısından açıklayabilmektir. Çok durumlu modellerde, modelin ve yöntemlerin genişletilmesi ve açıklanması için bağımlı değişken düzeylerinin sayısı hususunda bir sınırlama olmamasına rağmen, kolaylık sağlaması açısından, bağımlı değişkenlerin genellikle üç kategori değişkeni ile ifade edilmesi tavsiye edilmektedir. Üçten fazla kategori ile çalışılması kavram olarak bir sorun teşkil etmese de kategorilerin sayısı arttıkça, araştırmacı tanımlama ve simgeleme sorunları ile karşı karşıya kalabilecektir (Hosmer ve Lemeshow, 2000: 260).

Üç alternatife sahip bir durumda tercihte bulunan bireylerin fayda fonksiyonları aşağıdaki eşitliklerdeki gibi tanımlanmaktadır:

$$U_{1i} = x'_{1i}\beta + u_{1i} \quad (3.1)$$

$$U_{2i} = x'_{2i}\beta + u_{2i} \quad (3.2)$$

$$U_{3i} = x'_{3i}\beta + u_{3i} \quad (3.3)$$

Bu eşitliklerde; u_{1i} , u_{2i} , u_{3i} değerlerinin bağımsız özdeş dağıldığı ve kural olarak en yüksek faydayı sağlayan alternatifin tercih edileceği varsayılmaktadır. Benzer şekilde, eğer i 'nci bireyin kesikli tercihi y_i değişkeni ile ifade edilirse, sözü edilen tercihlere ait çok durumlu model aşağıdaki gibi tanımlanmaktadır (Amemiya, 1994: 336):

$$P(y_i = 1) = P(U_{1i} > U_{2i}, U_{1i} > U_{3i}) \quad (3.4)$$

$$P(y_i = 2) = P(U_{2i} > U_{1i}, U_{2i} > U_{3i}), \quad 1, 2, \dots, n \quad (3.5)$$

Sonuç değerleri, 0, 1, 2, $\dots$, m olmak üzere $m+1$ olası kategorik değere sahip bir çok durumlu modelde, $m + 1$ olası gözlem arasından herhangi bir i gözlemine ait olasılıkları genel olarak

$$P(y_i = 0) = 1 / 1 + \sum_{j=1}^m \exp[x'_i \beta^{(j)}] \quad (3.6)$$

$$P(y_i = 1) = \exp[x'_i \beta^{(1)}] / 1 + \sum_{j=1}^m \exp[x'_i \beta^{(j)}] \quad (3.7)$$

$$P(y_i = m) = \exp[x'_i \beta^{(m)}] / 1 + \sum_{j=1}^m \exp[x'_i \beta^{(j)}] \quad (3.8)$$

eşitliklerindeki gibi modellemek mümkündür (Montgomery vd., 2012: 443).

Daha genel bir ifade ile çok durumlu bir modeli (3.9) eşitliğindeki gibi yeniden tanımlamak mümkün olmaktadır:

$$P(y_i = j) = F_{ij}(x^*, \theta), \quad i = 1, 2, \dots, n \quad \text{ve} \quad j = 1, 2, \dots, m \quad (3.9)$$

Bu eşitlikte x^* ve θ değerleri sırasıyla bağımsız değişkenlerin ve parametrelerin birer vektörü olarak ifade edilmektedir. Birçok uygulamada bireyler farklı tercih kümeleri ile karşı karşıya kaldığından m değerinin i değerine bağlı olduğu düşünülmektedir. Yukarıdaki modelde θ değerinin en yüksek olabilirlik tahmincisini belirleyebilmek adına $\sum_{i=1}^n (m + 1)$ ikili değişkenlerini aşağıdaki gibi tanımlamak kolaylık sağlayacaktır:

$$y_{ij} = 1 \text{ eğer } y_i = j \text{ ise} \quad (3.10)$$

$$y_{ij} = 0 \text{ eğer } y_i \neq j \text{ ise} \quad (3.11)$$

Böylece, log-olabilirlik fonksiyonu,

$$\log L = \sum_{i=1}^n \sum_{j=1}^m y_{ij} \log F_{ij} \quad (3.12)$$

şeklinde yazılmakta ve en yüksek olabilirlik tahmincisi $\hat{\theta}$ değeri $\partial \log L / \partial \theta = 0$ eşitliğinin çözülmesi ile elde edilmektedir. Bununla birlikte, daha özel bir ifade ile H sabit bir tam sayı değerini temsil etmek üzere bağımsız değişkenlerin vektörü T farklı değer aldığında başlangıç modeli,

$$P(y_i = j) = F_j(x'_{t1}\beta_1, x'_{t2}\beta_2, \dots, x'_{tH}\beta_H) \quad (3.13)$$

$$i = 1, 2, \dots, T \text{ ve } j = 1, 2, \dots, m$$

şeklinde ifade edilirse, β parametresinin en küçük χ^2 tahmincisi bu yeni eşitlikteki m eşitliğin her bir t için dönüştürülmesi ve H sayıda $x'_{t1}\beta_1, x'_{t2}\beta_2, \dots, x'_{tH}\beta_H$ ifadenin çözülmesi ile elde edilmektedir (Amemiya, 1985: 287-290).

Sınıflama ölçeği ile ölçülen değişkenlerin kullanıldığı bir modelde, bağımlı değişkenin sıralı olması ve ihmal edilen bilgi olması durumunda, model tahmininde etkinlik kaybı söz konusu olabilecektir. Diğer taraftan, sıralı değişkenlere uygulanabilen bir yöntem, sınıflama ölçeği ile ölçülmüş bir bağımsız değişkene uygulandığında ise tahminler, sapmalı ya anlamsız olabilmektedir. Bağımlı değişkenin sıralı olup olmadığı ile ilgili şüphe duyulması durumunda, sınıflama ölçeği ile ölçülmüş çıktılar için kullanılan modellerdeki potansiyel etkinlik kaybının göz önünde bulundurulması daha önemli olmaktadır (Long, 1997: 149). Sıralı ve sıralı olmayan modellerin analizi için farklı tekniklere gereksinim duyulmaktadır. Sıralı modeller, hem logit hem de probit modeller ile tahmin edilebilmekte iken, sıralı olmayan modellerin tahmini çoğunlukla logit modeller ile yapılmaktadır. Sıralı olmayan modellerin probit modeller ile tahmini mümkün olsa bile, uygulaması nispeten zor olan bu modeller genellikle tercih edilmemektedir. Bu sebeple, sıralı olmayan çoklu tercih modelleri, çok durumlu logit modeller ile çözümlenmektedir (Borooah, 2002: 2).

3.1.1. Sıralı Modeller

Bazı çok durumlu tercih modelleri, doğası gereği sıralı değişkenlerden oluşmaktadır. Örneğin; tahvil derecelendirmeleri, katılımcıların görüşlerinin değerlendirildiği araştırmalar, askeri personelin yeteneklerine göre görev sınıflarına atanması, belirli programlar için oy verme sonuçları, bir tüketicinin tercih ettiği sigorta poliçesinin kapsamı ve istihdam gibi değişkenler sıralı değişkenler olarak tanımlanmaktadır (Greene, 2008: 831). Kategorik tercih modellerinin yaygınlık kazanması ile birlikte doğası gereği sıralı değişkenlerden oluşan modellerin literatürdeki uygulama alanları genişlemeye başlamıştır (Small, 1987: 409). Sıralı kategorik değişkenler kesin ölçümlerin her zaman mümkün olmadığı psikoloji, sosyoloji, tıp ve sosyal bilimler gibi birçok çalışma alanı için oldukça önem arz etmektedir (Anderson ve Philips, 1981: 22; Long, 1997: 114). Kategorik veri analizinde sınıflama ölçeği ile ölçülmüş değişkenlerin yerine sıralı kategorik değişkenlerin kullanılmasının üstünlüklerinden bazılarını şu şekilde sıralamak mümkündür:

1. Sıralı değişkene sahip verilerin analizinde kullanılan ölçümler; korelasyon, eğim, ortalama gibi nicel değişkenler için kullanılan ölçümler ile benzerlik göstermektedir.
2. Sıralı değişkene sahip verilerin analizleri sınıflama ölçeği ile ölçülmüş verilerin analizine göre daha kapsamlı, daha hassas ve yorumlanması daha kolay analizler olarak kabul edilmektedir.
3. Sıralı değişkene sahip verilerin analizlerinde kullanılan yöntemler sınıflama ölçeği ile ölçülmüş veriler için kullanılan yöntemlere göre daha güçlü yöntemler olarak kabul edilmektedir.
4. Sıralı modellerin uyum iyiliği ölçümünde sınıflama modellerine göre daha az parametre test edilmektedir ve hesaplama kolaylığı sağlamaktır (Agresti, 2010: 3).

Bir değişken sıralama ölçeği ile ölçüldüğünde, anılan değişkene ait kategoriler düşük değerden yüksek değere kadar sıralanmasına rağmen ardışık kategoriler arasındaki mesafeler bilinmemektedir. Bununla birlikte, bir değişkenin sıralı olabilmesi anılan değişkenin kesin olarak sıralı modeller ile analiz edileceği anlamına gelmemektedir. Belirli bir amaç için sıralı olan bir değişken başka bir amaç için sıralı

olmayabilecektir (Long, 1997: 114-115). Sıralı modellerin temelinde tek bir indeks fonksiyonuna sahip gizli değişkenlerin ve sabit kesişim noktalarının göz önünde bulundurulması yatmaktadır. Bu bağlamda, sözü edilen yaklaşımların oldukça kısıtlayıcı oldukları düşünülmekte ve eğilimin yönünü belirleyen değişkenlerin kategorik sonuç olasılıkları üzerindeki etkilerinin esnek bir şekilde tahmin edilemediği tartışılmaktadır. Bu sebeple sıralı değişkenler için bu kısıtlamaların giderilmesini sağlayan alternatif modellerin de geliştirildiği görülmektedir (Boes ve Winkelmann, 2006: 167).

Her bir birey için yalnızca bir olayın gerçekleştiği N sayıda bireyin ($i = 1, \dots, N$) olduğu, karşılıklı dışlayıcılık özelliğini sağlayan ve tüm olasılıkları kapsayan sözü edilen olayın $M > 2$ sayıda ($j = 1, \dots, M$) “doğal sıraya sahip” sonucunun olduğu varsayıldığında, Y_i değişkeninin aldığı değerler i bireyi için gözlenen sonuçları temsil etmek üzere, tasarlanan bu sıralı modele ait bağımsız değişken hem kesikli hem de sıralı bir değişken olmaktadır. Örneğin, model sonuçları 1, 2 ve 3 gibi sayısal değerler ile ifade edilirse, doğrusal regresyon modelinde bu sıralı değerlerin birbirinden farkı olmayacaktır. Öte yandan, sıralı bağımlı değişkene ait ekonometrik ilişkinin çok durumlu sıralı modeller ile tahmin edilmesi sayesinde doğrusal regresyondan farklı olarak verilerin doğal sırası dikkate alınacak ve tutarlı tahminler elde edilebilecektir (Borooah, 2002: 4-5). Aynı zamanda, sözü edilen doğal sıralama, alternatiflere ait birbiri ile bağlantılı gözlenemeyen tesadüfi fayda bileşenleri arasında çok yakın ortak değişken ilişkisi olduğu anlamına gelmektedir. Böylece, gözlenen sonuçların bilinmeyen parametreler içerisinde doğrusal özelliğe sahip gizli değişken aralıklarını yansıttığı durumda sıralı modeller vasıtası ile analiz yöntemleri tercih edilmektedir (Small, 1987: 409-410).

Ölçek kategorilerinin aralıklı olduğu sürekli bir ölçüm düşünüldüğünde ve k kategoriden oluşan bir ölçek için s_j kategorisi x_{j-1} ile x_j aralığında olmak üzere, bu kategoride yer alan noktaların x_j ($j = 1, 2, \dots, k$) ile temsil edildiği varsayılmaktadır. Bununla birlikte, m adet gözlem grubu olduğu ve sürekli dağılım fonksiyonunun $P_i(x_i)$ ile temsil edildiği varsayıldığında, s_j kategorisindeki i gözlem grubunda yer alan bir gözlem grubunun olasılığı,

$$P_i(x_i) - P_i(x_{j-1}), \quad i = 1, 2, \dots, m; \quad j = 1, 2, \dots, k \quad (3.14)$$

şeklinde tanımlanmaktadır (Snell, 1964: 593). Sıralı tercih modellerinin genel tanımı (3.14) eşitliğinden hareket ile aşağıdaki eşitlikteki gibi yapılmaktadır:

$$P(y = j|x, \theta) = P(s_j) \quad (3.15)$$

Bu eşitlikteki olasılık değeri ile ardışık aralıkların sonlu bir değeri olan s_j dizisi, x ve θ eşitliklerine bağlı olarak ölçülmektedir. Sıralı tercih modelinde, bir y değişkenin aldığı değerler, reel sayılar kümesinin bir bölümünde yer almaktadır. Bununla birlikte, sıralı tercih modelleri birçok uygulamada F gibi bir dağılım fonksiyonu için aşağıdaki basitleştirilmiş eşitlik ile ifade edilmektedir:

$$P(y = j|x, \alpha, \beta) = F(\alpha_{j+1} - x'\beta) - F(\alpha_j - x'\beta) \quad (3.16)$$

$$j = 0, 1, \dots, m, \quad \alpha_0 = -\infty, \quad \alpha_j \leq \alpha_{j+1}, \quad \alpha_{m+1} = +\infty$$

Anılan eşitlikte, $F = \phi$ olması durumunda sıralı probit modeli, $F = A$ olması durumunda ise sıralı logit modeli tanımlanmış olacaktır. Yukarıdaki eşitlik, gözlenemeyen, sürekli ve tesadüfi y^* değişkeninin aşağıdaki koşulu sağlamasından hareket ile yazılabilmektedir:

$$y = j \quad \text{ancak ve ancak} \quad \alpha_j < y^* < \alpha_{j+1} \quad (3.17)$$

$$j = 0, 1, \dots, m$$

Bir başka ifade ile $y^* - x'\beta$ ifadesine ait dağılım fonksiyonunun yukarıdaki eşitlikteki F fonksiyonu olması durumunda (3.16) eşitliği sağlanmış olmaktadır (Amemiya, 1985: 292-293).

Sıralı logit ve probit modellerinin analizi noktasında dikkate alınması gereken önemli yaklaşımlardan birinin de “paralel eğimler” varsayımı olduğu düşünülmektedir. Söz konusu varsayıma göre, bir sıralı modelde bir bireyin sıralı kategorilerde yer alma olasılığını etkileyen bir değişken mevcut ise bu değişkeni farklı sonuçlar ile ilişkilendiren katsayıların tüm sonuçlar için aynı olması gerekmektedir. Bu varsayımının sağlanmadığı durumda, belirli bir sonuç değeri ile ilişkili eğim katsayısı başka sonuçlar için farklılık gösterecek ve bu sebeple anılan model için sıralı modeller uygun bir çözüm olmayacaktır (Boorah, 2002: 6). Sıralı modeller biyoistatistik gibi bazı farklı disiplinlere ait çalışmalarda orantılı bahis modeli, paralel regresyon modeli ve gruplandırılmış sürekli model olarak farklı tanımlamalar ile adlandırılmıştır. Tahmin

edilebileceği gibi orantılı bahis modelinde, modelin bahis oranına göre yorumlanmasına odaklanılmaktadır. Bununla birlikte, paralel regresyon modelinde model tarafından türetilen olasılık eğrilerinin yapısı hakkında üstü kapalı bir varsayıma; gruplandırılmış sürekli modelde ise sürekli değişken ile gözlenen ve gruplandırılmış değişken arasındaki ilişkiye dikkat çekilmektedir (Long, 1997: 116).

3.1.1.1.Sıralı Probit Modeli

Bir Z değişkeni, Y^* gibi bir gizli değişkenden kaynaklanan M ($R_1, \dots, R_M$) adet yanıt kategorisine sahip bir kategorik değişken olmak üzere, $M + 1$ adet genişletilmiş gerçek değerlerin ortalama değerlerinin $\mu_0, \mu_1, \dots, \mu_M$ ($\mu_0 = -\infty, \mu_M = +\infty, \mu_0 \leq \mu_1 \leq \dots \leq \mu_M$) şeklinde tanımlandığı varsayıldığında, Z kategorik değişkenine ait aşağıdaki ifade elde edilecektir:

$$1 \leq j \leq M \text{ için, } Z_j \in R_i \leftrightarrow \mu_{i-1} < Y_j \leq \mu_i \quad (3.18)$$

Z gözlenen bağımlı değişkeninin, sıralı kategorik bir değişken olduğu durumda, anılan değişken gölge değişkenler ile aşağıdaki eşitlikteki gibi ifade edilebilmektedir:

$$1 \leq j \leq n, 1 \leq k \leq M \text{ için, } Z_{jk} = \begin{cases} 1 & \text{eğer } Z_j \in R_k \\ 0 & \text{değilse} \end{cases} \quad (3.19)$$

Hata terimi ε değerinin çok değişkenli normal dağılıma sahip olduğu varsayıldığında, herhangi bir $1 \leq j \leq n, 1 \leq k \leq M$ için ve $\phi(t)$ birikimli standart normal yoğunluk fonksiyonu olmak üzere, Z değişkeninin olasılık fonksiyonu aşağıdaki gibi yazılmaktadır:

$$P[Z_{jk} = 1] = P[Z_j \in R_k] = \phi \left[\frac{(\mu_k - \sum_{i=0}^K \beta_i X_{ij})}{\sigma} \right] - \phi \left[\frac{(\mu_{k-1} - \sum_{i=0}^K \beta_i X_{ij})}{\sigma} \right] \quad (3.20)$$

Bu eşitlikte, $\mu_1 = 0$ ve $\sigma = 1$ değerleri yerine konulduğunda, tahmin edilecek nihai sıralı probit modeli aşağıdaki eşitlikteki gibi olacaktır:

$$P[Z_{jk} = 1] = \phi \left[\mu_k - \sum_{i=0}^K \beta_i X_{ij} \right] - \phi \left[\mu_{k-1} - \sum_{i=0}^K \beta_i X_{ij} \right] \quad (3.21)$$

Sözü edilen modelde, araştırmacının temel sorunu $Q = M + K - 1$ parametresinin tutarlı tahmincilerini tespit etmektir (McKelvey ve Zavoina, 1975: 106-107).

Bir sıralı probit modelinde, herhangi bir gözlenen sonuç için yazılabilecek genel olasılık değeri aşağıdaki eşitlikteki gibi tanımlanmaktadır:

$$P(y_i = m|x_i) = F(\tau_m - x_i\beta) - F(\tau_{m-1} - x_i\beta) \quad (3.22)$$

Bu eşitlikteki τ değerleri kesim noktalarını, m kategori sayısını ifade etmektedir. Ayrıca, sözü edilen eşitlik ışığında $P(y = 1 | x)$ olasılığı hesaplanırken $F(\tau_0 - x\beta) = F(-\infty - x\beta) = 0$ olduğundan eşitliğin sağ tarafı sıfır olacak ve $P(y = J | x)$ olasılığı hesaplanırken $F(\tau_J - x\beta) = F(\infty - x\beta) = 1$ olduğundan birinci terim bir olacaktır. Böylece, dört adet gözlenen sonucun söz konusu olduğu bir sıralı probit modeli için ilgili olasılık değerleri,

$$P(y_i = 1|x_i) = \phi(\tau_1 - \alpha - \beta x_i) \quad (3.23)$$

$$P(y_i = 2|x_i) = \phi(\tau_2 - \alpha - \beta x_i) - \phi(\tau_1 - \alpha - \beta x_i) \quad (3.24)$$

$$P(y_i = 3|x_i) = \phi(\tau_3 - \alpha - \beta x_i) - \phi(\tau_2 - \alpha - \beta x_i) \quad (3.25)$$

$$P(y_i = 4|x_i) = 1 - \phi(\tau_3 - \alpha - \beta x_i) \quad (3.26)$$

şeklinde yazılmaktadır (Long, 1997: 121).

3.1.1.2.Sıralı Logit Modeli

Sıralı logit modelinin bahis oranları dikkate alınarak yorumlanması mümkün olmaktadır. Daha önce de bahsedildiği gibi bu yorum biyoistatistik alanında orantılı bahis modeli adı altında ilk olarak McCullagh (1980) tarafından yapılmıştır. Ortak değişkenlerin x değerini aldığı, k sayıda sıralı kategoriden oluşan ve sonuç değerlerine ait olasılıkların $\pi_1(x), \pi_2(x), \dots, \pi_k(x)$ değerleri ile temsil edildiği bir modelde, Y değişkeninin sözü edilen olasılık değerlerine göre 1, ... , k aralığında değer alan bir sonuç değişkeni, $K_j(x)$ değeri ise $Y \leq x$ olduğu durumda verilen x değerlerine göre elde edilen bahis değerleri olmak üzere, orantılı bahis modeline göre sözü edilen $K_j(x)$ bahis değeri aşağıdaki eşitlikteki gibi yazılmaktadır:

$$K_j(x) = K_j \exp(-\beta^T x) \quad (1 \leq j < k) \quad (3.27)$$

Bu eşitlikte β bilinmeyen parametrelerin bir vektörü olarak tanımlanmaktadır. Orantılı bahis oranları modelinde ardışık bahis oranları,

$$K_j(x_1)/K_j(x_2) = \exp\{\beta^T(x_2 - x_1)\} \quad (1 \leq j < k) \quad (3.28)$$

j değerinden bağımsız olmakta ve yalnızca ortak değişkenler arasındaki farka ($x_2 - x_1$) göre değişmektedir. Ayrıca, $\gamma_j(x) = \pi_1(x) + \dots + \pi_j(x)$ ifadesi sağlandığında, $Y \leq x$ durumunda elde edilen bahis değeri, $\gamma_j(x)/\{1 - \gamma_j(x)\}$ şeklinde yazılmakta ve $\theta_j = \log K_j$ olmak üzere nihai orantılı bahis modeli,

$$\log[\gamma_j(x)/\{1 - \gamma_j(x)\}] = \theta_j - \beta^T x \quad (1 \leq j < k) \quad (3.29)$$

şeklinde tanımlanmaktadır. (3.29) eşitliğindeki bu model, ardışık birikimli logit değerleri arasındaki farkın dâhil oldukları kategorilerden bağımsız olduğunu ifade etmektedir. Ayrıca, yukarıdaki eşitlik, iki durumlu veriler için doğrusal lojistik model ile eşdeğer olmakta ve ilerleyen bölümlerde açıklanacağı gibi log-doğrusal modeli vermektedir. Ancak, çok durumlu modeller için böyle bir çıkarım yapılamamaktadır (McCullagh, 1980: 110).

Orantılı bahis modelinde Y sonucuna ait kodlamanın ve kategorilerin değişmesi β değerini değiştirmeyecek, ancak işaret değişikliğine ve θ_j değerlerinin etkilenmesine sebep olabilecektir (Ananth ve Kleinbaum, 1997: 1325). Ek olarak, öngörücü terim x değerinin önündeki eksi işareti, parametre tanımlaması adına gerekli olmasa bile, β değerinin her bir bileşenine ait pozitif ya da negatif etkinin belirlenmesi adına modele eklenmiştir. Ayrıca θ_j değerleri kesim noktalarını ifade etmekte ve kurulan bu model, her bir birikimli olasılık için x değerlerine ait etkinin özdeş olduğu varsayılarak eş anlamlı olarak $k - 1$ sayıdaki tüm birikimli olasılıklara uygulanmaktadır (Liu ve Agresti, 2005: 3). Orantılı bahis modeli ve alternatif ilişki fonksiyonları ile geliştirilen ilgili diğer modellerde sonuç kategorileri için puanlama yapmaya gereksinim duyulmamakta ve bu durum araştırmacılara kolaylık sağlamaktadır (Agresti, 1999: 2192).

Bununla birlikte, McCullagh'ın önerdiği bu model yalnızca bir açıklayıcı değişken için uygulama alanı bulabilmektedir. Bu tür modeller; herhangi bir sayıda açıklayıcı değişken, farklı örneklem durumları ve sıralı ana kütleler arasındaki ayrımlar için genişletilebilmektedir (Anderson ve Philips, 1981: 22). Orantılı bahis varsayımı, $\log(\Pi_j) = \theta_j - x'\beta_j$ şeklinde genişletilen bir model sayesinde hafifletilebilmektedir. Bu yeni modelde, β parametresinin j değerine göre değişebilmesine imkân tanınmaktadır. Bunun yanında, sözü edilen modelde orantılı olma varsayımının log-bahis oranının (β) içerisinde sınanması araştırmacılara kolaylık sağlamak ve bu durum $H_0: \beta_1 = \beta_2 = \beta_k$

şeklinde ifade edilen bir sıfır hipotezinin test edilmesi ile açıklanmaktadır (Ananth ve Kleinbaum, 1997: 1324-1325).

Sıralı logit modeli, birikimli olasıklar için elde edilen bahis oranları vasıtasıyla da yorumlanabilmektedir. Birikimli olasılık sonuç değişkeni m değerine eşit ya da bu değerden küçük olmak üzere,

$$P(y \leq m|x) = \sum_{j=1}^m P(y = j|x) \quad m = 1 \text{ için } J - 1 \quad (3.30)$$

şeklinde tanımlanmaktadır. Sonuç değerinin m değerine eşit ya da bu değerden küçük olma olasılığının büyük olma olasılığına oranı,

$$\Omega_m(x) = \frac{P(y \leq m|x)}{1 - P(y \leq m|x)} = \frac{P(y \leq m|x)}{P(y > m|x)} \quad (3.31)$$

şeklinde sonuç değerine ait bahis oranını vermektedir. Sıralı logit modeli için yukarıdaki eşitlikte gösterilen bahis oranı

$$\Omega_m(x) = \frac{P(y \leq m|x)}{P(y > m|x)} = \exp(\tau_m - x\beta) \quad (3.32)$$

eşitliğindeki gibi yazılmaktadır (Long, 1997: 138).

3.1.2. Sıralı Olmayan Modeller

Doğası gereği sıralı kategorik değerlere sahip olmayan modeller sıralı olmayan modeller olarak tanımlanmaktadır. Sıralı olmayan modeller, çok durumlu doğrusal olasılık modeli, logit ve probit modellerini, ardışık ve yuvalanmış modellerini kapsamaktadır.

3.1.2.1. Çok Durumlu Doğrusal Olasılık Modeli

Çok durumlu doğrusal olasılık modeli, iki alternatifte sahip benzer modelden farklı olarak J ($j = 1, 2, \dots, J$) adet eşitlik ile değerlendirilmektedir. Böylece, çok durumlu doğrusal olasılık modeli, (2.4) eşitliğinde tanımlanan iki durumlu model ile aynı varsayımlar altında düşünülmekte, ancak sözü edilen J adet eşitlik yapılacak analize ayrı ayrı dâhil edilmektedir. İki durumlu modelde olduğu gibi, çok durumlu doğrusal olasılık modelinde de En Küçük Kareler yöntemi analizinin sonuçları tutarlı olsa bile, modeldeki hata terimi değişen varyansa sahip olmaktadır (Aldrich ve Nelson,

1984: 23). Tesadüfî fayda fonksiyonu yaklaşımı ışığında çok durumlu doğrusal olasılık modeli şu şekilde açıklanmaktadır: Bilinmeyen parametrelerde log-doğrusal yapı gösteren bir $V(x, s)$ fayda fonksiyonu,

$$V(x^i, s) = \log \left[\sum_{k=1}^K \beta_k Z^k(x^i, s) \right] \quad (3.33)$$

şeklinde yazılabilmektedir. Bu eşitlikte x^i gözlenen tutumları, s karar vericilere ait özellikleri, β bilinmeyen parametreleri temsil etmektedir. Ayrıca,

$$0 \leq \sum_{k=1}^K \beta_k Z^k(x^i, s) \leq 1, \quad i = 1, \dots, J \quad (3.34)$$

$$\sum_{i=1}^J \sum_{k=1}^K \beta_k Z^k(x^i, s) = 1 \quad (3.35)$$

eşitliklerinin sağlandığı varsayıldığında, olasılık fonksiyonları aşağıdaki şekilde ifade edilmektedir:

$$P_i = h(B, s, i) = \sum_{k=1}^K \beta_k Z^k(x^i, s) \quad (3.36)$$

Bu eşitlikte; B alternatifler kümesi, H tercih davranışları kümesi olmak üzere $h \in H$ olduğu kabul edilmektedir. Yukarıdaki eşitlik iki durumlu doğrusal olasılık modelin çok durumlu uzantısı olarak görülmektedir. Bu genişletilmiş modelin doğrusal yapısı deneysel uygulamalar için ilgi uyandırır da toplam değerinin 0 ile 1 arasında olduğunu öngören eşitsizlik, sürekli uygulamalarda sözü edilen doğrusal yapının bozulmasına sebep olmaktadır (Domencich ve McFadden, 1975: 79-80).

Çok durumlu doğrusal olasılık modeli, iki alternatif için söz konusu olan doğrusal olasılık modelinin ikiden fazla karşılıklı dışlayıcı alternatif için genelleştirilmesi olarak tanımlanmaktadır. Bu bağlamda, üç alternatife sahip bir durum incelendiğinde her bir alternatifin tercih edilme olasılığı aşağıdaki eşitliklerdeki gibi olmaktadır:

$$P_{1i} = \alpha_1 + \beta_1 X_1 \quad (3.37)$$

$$P_{2i} = \alpha_2 + \beta_2 X_2 \quad (3.38)$$

$$P_{3i} = \alpha_3 + \beta_3 X_3 \quad (3.39)$$

Bu eşitliklerde; P_{ij} değeri i 'nci bireyin j 'inci alternatifi tercih etme olasılığını, X_i değişkeni i 'nci bireyin X değerini, α ve β değerleri regresyon katsayılarını temsil etmektedir. Böylece, olasılıklar toplamı ve olasılıkların ortalamaların toplamı 1, söz konusu olasılıkların ortalamadan farkları toplamı ise 0 olmakta ve her bir i için bu durum aşağıdaki eşitliklerdeki gibi ifade edilmektedir:

$$P_{1i} + P_{2i} + P_{3i} = 1 \quad (3.40)$$

$$\bar{P}_1 + \bar{P}_2 + \bar{P}_3 = 1 \quad (3.41)$$

$$p_{1i} + p_{2i} + p_{3i} = 0 \quad (3.42)$$

Bu bağlamda, alternatiflere ait standart En Küçük Kareler tahmin değerleri toplamı aşağıdaki gibi yazılabilmektedir:

$$\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 = \frac{\sum (P_{1i} + P_{2i} + P_{3i})X_i}{\sum X_i^2} = 0 \quad (3.43)$$

$$\begin{aligned} \hat{\alpha}_1 + \hat{\alpha}_2 + \hat{\alpha}_3 &= (\bar{P}_1 - \hat{\beta}_1 \bar{X}) + (\bar{P}_2 - \hat{\beta}_2 \bar{X}) + (\bar{P}_3 - \hat{\beta}_3 \bar{X}) \\ &= (\bar{P}_1 + \bar{P}_2 + \bar{P}_3) - (\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3) \bar{X} = 1 \end{aligned} \quad (3.44)$$

$$\begin{aligned} \hat{P}_{1i} + \hat{P}_{2i} + \hat{P}_{3i} &= (\hat{\alpha}_1 + \hat{\beta}_1 X_i) + (\hat{\alpha}_2 + \hat{\beta}_2 X_i) + (\hat{\alpha}_3 + \hat{\beta}_3 X_i) \\ &= (\hat{\alpha}_1 + \hat{\alpha}_2 + \hat{\alpha}_3) + (\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3) X_i = 1 \end{aligned} \quad (3.45)$$

Bu eşitliklerde, alternatiflere ait iki eşitliğin çözülmesi ile üçüncü eşitliğin çözüm kümesi elde edilebilmektedir. Çok durumlu doğrusal olasılık modelinde, her bir denklemin bağımsız değişkenlerinin özdeş olmaması halinde, olasılıkların toplamı 1 olmayacak ve En Küçük Kareler yöntemi uygun bir yaklaşım olmaktan çıkacaktır. Bu noktada, En Küçük Kareler yönteminin yerine hata terimleri arasındaki karşılıklı ilişkiyi dikkate alan genelleştirilmiş En Küçük Kareler yöntemi önerilmektedir (Zellner ve Lee, 1965: 388-390; İşyar, 1999: 279-281; Özer, 2004: 83-85).

Alternatiflerin toplamının 1 olması koşulunu sağlamak için ya bir alternatif için z^{ji} şeklinde bir artık değişkenin tanımlanması ya da tüm z^{ji} değerlerinin w_i gibi bir ağırlık değeri yardımıyla normal hale getirilmesi gerekmektedir. Ancak, söz konusu iki yöntemin uygulanması da bir alternatifi temsil eden faydasının tüm alternatiflere ait tutumlara bağlı olduğunu öngörmekte ve bu durum beğenilerin ve fırsatların bağımsızlığı varsayımını ihlal etmektedir. Bununla birlikte, ağırlıklandırma yöntemi

modelin doğrusal yapısına zarar vermektedir. Bu sebeplerle, çok durumlu doğrusal olasılık yönteminin tatmin edici istatistiksel özellikler ile uygulanabilir tahminler sağlayabileceği düşünülmemekte, ancak bazı koşulların sağlanması durumunda küçük örneklerde göreceli olarak tatmin edici parametre tahminleri elde etmek amacıyla doğrusal yaklaşımın kullanılması mümkün olabilmektedir (Domencich ve McFadden, 1975: 119).

3.1.2.2.Çok Durumlu Logit Modeli

Üç alternatife sahip bir modelde verilen herhangi bir ikili kategorik sonuç değeri için,

$$P_1(P_1 + P_0)^{-1} = L(\beta'_1 x) \quad (3.46)$$

$$P_2(P_2 + P_0)^{-1} = L(\beta'_2 x) \quad (3.47)$$

$$P_1(P_1 + P_2)^{-1} = L(\beta'_3 x) \quad (3.48)$$

P_0 , P_1 ve P_3 olasılıkları ilk iki eşitlikten kolayca hesaplanabilmekte ve $\beta_3 = \beta_1 - \beta_2$ ifadesi elde edilmektedir. Böylece, $D = 1 + \exp(\beta'_1 x) + \exp(\beta'_2 x)$ olmak üzere, İlk olarak Cox (1966) tarafından öne sürülen sıralı olmayan bir logit modeli,

$$P_2 = D^{-1} \exp(\beta'_2 x) \quad (3.49)$$

$$P_1 = D^{-1} \exp(\beta'_1 x) \quad (3.50)$$

$$P_0 = D^{-1} \quad (3.51)$$

şeklinde tanımlanmaktadır (Amemiya, 1975: 365).

3.1.2.2.1. İlişkisiz Alternatiflerin Bağımsızlığı Varsayımı

Belirleme testleri, ekonometri ve istatistik bilimleri üzerine yapılan araştırmalarda en önemli alanlardan birini oluşturmaktadır. Klasik regresyon denklemi ($y = X\beta + \varepsilon$) çerçevesinde, iki tür tesadüfi belirlemeden söz edilmektedir. Bunlardan birincisi, verilen bir X değişkenine göre, ε hata teriminin koşullu beklenen değerinin sıfıra eşit olması (ya da sabit X değerleri için hata teriminin sıfır olması); ikincisi ise küresel bir ortak değişken matrisine sahip olmasıdır. Bu iki varsayım, aşağıdaki eşitliklerdeki gibi özetlenmektedir:

$$E(\varepsilon | X) = 0 \quad \text{ya da büyük örneklerde} \quad {}^5\text{plim} \frac{1}{T} X' \varepsilon = 0 \quad (3.52)$$

$$V(\varepsilon | X) = \sigma^2 I \quad (3.53)$$

Birinci varsayımın ihlal edilmesi taraflı tahminlere, ikinci varsayımın ihlal edilmesi merkezi eğilim tahmincisi doğru olsa bile etkinlik kaybına sebep olabilecektir. Bunun yanında, birçok problemde, birinci varsayımının başarısının tespiti genelde ikinci varsayımının tespitinden daha fazla yapılmaktadır. Bu sebeple, ekonometri literatüründe ikinci varsayımın sağlanıp sağlanmadığının tespiti için özel belirleme testleri geliştirilmeye çalışılmaktadır (Hausman, 1978: 1251).

Bireylerin tercih davranışlarını gözleyebilmek ve değerlendirebilmek iktisat, işletme, pazarlama, siyaset bilimi, psikoloji gibi birçok farklı disiplinin temel sorunları arasında yer almaktadır. Bu noktada, tercih davranışları ile ilişkili en önemli özelliklerden biri ilişkisiz alternatiflerin bağımsızlığı varsayımı olarak görülmektedir. Sözü edilen özellik en genel anlamı ile tercih kümesindeki herhangi iki alternatifin tercih olasılık oranlarının diğer alternatiflere göre değişmediğini varsaymaktadır. Bir başka ifade ile x ve y gibi iki alternatifin olduğu bir tercih kümesine, z gibi üçüncü bir alternatif dâhil edildiğinde, z alternatifinin x ve y alternatifleri arasından orantılı olarak tercih edilme olasılığı, bu alternatiflerin olasılıkları ile aynı oranda olacaktır (Gensch ve Ghose, 1997: 201).

İlişkisiz alternatiflerin bağımsızlığı varsayımının teorik altyapısı şu şekilde tanımlanmaktadır: $R_1, \dots, R_n$ ve $R'_1, \dots, R'_n$ gibi iki bireysel sıralamalar kümesinin ve S ile temsil edilen belirli bir çevrede $C(S)$ ve $C'(S)$ gibi iki toplumsal tercih fonksiyonunun olduğu varsayıldığında, tüm i bireyleri ile x ve y değerleri için $xR_i y \leftrightarrow xR'_i y$ koşulunun sağlanması durumunda $C(S)$ ve $C'(S)$ fonksiyonları aynı değeri almaktadır (Arrow, 1951: 27). Arrow'un bu genel tanımının ışığında bir U kümesi, olası alternatiflerin oluşturduğu bir evrensel küme; T kümesi sözü edilen evrensel kümenin sonlu bir alt kümesi; R ve S kümeleri, T alt kümesinin bir alt kümesi; x ve y değişkenleri T alt kümesinin bir elemanı olmak üzere, S alt kümesinin tercih olasılığı aksiyomu P_S ile gösterilmekte ve aşağıdaki gibi yazılabilmektedir:

⁵ plim = Limit değerinin olasılığı

- i. Eğer $P(x, y) \neq 0$, 1 Bütün $x, y \in T$ olmaktadır. Bu durumda, $R \subset S \subset T$ için $P_T(R) = P_S(R)P_T(S)$;
- ii. Eğer $P(x, y) = 0$, Bazı $x, y \in T$ olmaktadır. Bu durumda, her $S \subset T$ için $P_T(S) = P_{T-\{x\}}(S - \{x\})$.

Yukarıdaki *i* ve *ii* aksiyomlarından hareket ile herhangi $S \subset T$ ve $x, y \in S$ için, ilişkisiz alternatiflerin bağımsızlığı aksiyomu,

$$\frac{P(x, y)}{P(y, x)} = \frac{P_S(x)}{P_S(y)} \quad (3.54)$$

şeklinde tanımlanacaktır (Luce, 1959: 6, 9).

İlişkisiz alternatiflerin bağımsızlığı varsayımının sezgisel amacı ve sınırlarının kapsamı, bireylerin tercihleri ile ilgili klasik tesadüfi fayda eşitliği dikkate alınarak açıklanabilmektedir. Bu açıklamayı yapabilmek için, bir alternatifler kümesi düşünülmekte, her bir alternatifin U_i gibi bir fayda değeri ile temsil edildiği ve herhangi bir birey için bu fayda değerinin, $U_i = \bar{U}_i + \varepsilon_i$ şeklinde tanımlandığı varsayılmaktadır. Bu eşitlikte, $\bar{U}_i$ ilgili alternatife ait ortalama fayda değerini temsil etmekte ve bu ortalama değerin belirlenmesinde bütün karar vericiler göz önünde tutulmaktadır. Ayrıca, eşitlikteki ε_i değeri belirli bir bireyin değerlemesinin ortalamadan sapmasını gösteren tesadüfi hata terimini ifade etmektedir. Bu bağlamda, üç alternatif olduğu düşünüldüğünde, bu alternatifler ile ilgili fayda fonksiyonları sırasıyla; $U_1 = \bar{U}_1 + \varepsilon_1$, $U_2 = \bar{U}_2 + \varepsilon_2$ ve $U_3 = \bar{U}_3 + \varepsilon_3$ şeklinde tanımlanmakta ve tesadüfi fayda teorisinin tanımı gereği bireyin alternatifler arasından en yüksek faydayı tercih edeceği varsayılmaktadır. Ek olarak, her bir alternatifin Y ve Z gibi iki farklı davranış biçimi ile temsil edildiği varsayıldığında, $U_1 = U(Y_1, Z_1) = \bar{U}(Y_1, Z_1) + \varepsilon_1$ eşitliği elde edilecek ve bu eşitlik benzer şekilde U_2 ve U_3 için de yazılabilecektir. İki boyutlu bir grafikte (Y_1, Z_1) , (Y_2, Z_2) , (Y_3, Z_3) noktaları işaretlendiğinde, anılan grafikte birinci alternatifin ikinci alternatife yakın bir yerde konumlandığı iddia edilebilecek ve bu durumda, ε_1 pozitif olmak üzere, birinci alternatife göre ortalama değerden büyük daha yüksek değeri tercih edecek birey, aynı zamanda ikinci alternatife göre de en yüksek değeri tercih edecektir. Böylece, ε_i hata terimleri ilişkili olacak ve ilişkinin derecesi, söz konusu davranış biçimleri uzayında davranış biçimlerinin birbirine “yakınlıklarının” derecesi arttıkça artacaktır. İlişkisiz alternatiflerin bağımsızlığı varsayımı, sözü edilen

sapma terimlerinin birbirinden bağımsız olduğunu kabul etmektedir (Hausman ve Wise, 1978: 404).

İlişkisiz alternatiflerin bağımsızlığı varsayımı, herhangi iki alternatifin tercih edilmesinin üçüncü bir tercihten bağımsız olduğunu ifade ettiğinden kullanılan belirleme testi yöntemleri, alternatiflerin çok durumlu modeldeki tahmin katsayılarının üzerindeki etkisini tespit edebilmek adına tercih kümesinden bir ya da daha fazla alternatifin model dışında bırakılmasını temel alan yöntemler olacaktır. Sözü edilen yöntemlerin üstünlükleri, yeni bir alternatif model kurulmasını gerektirmemeleri ve bu modellerden en azından bir kaçının bilgisayar paket programları yardımı ile çözümlenebilmeleridir (Small ve Hsao, 1985: 619). Bununla birlikte, ilişkisiz alternatiflerin varsayımı nedeniyle çok durumlu modellerin birbirlerini ikame edebilecek yakın alternatif kümelerine uygulanması araştırmacılar için mantıklı bir tercih olmayacaktır. Bu sebeple çok durumlu modellerin uygulama alanı, alternatiflerin her bir karar verici tarafından birbirinden bağımsız olarak görüldüğü durumlar ile kısıtlanmaktadır (McFadden, 1974: 113).

İlişkisiz alternatiflerin bağımsızlığı varsayımının ihlal edilip edilmediği tercih kümesi bölümlene ve model esaslı testler olmak üzere iki temel test türü kullanılarak tespit edilebilmektedir. Tercih kümesi bölümlene testleri, çok durumlu logit modelindeki bütün tahmin sonuçları ile yalnızca bazı tahminlerin olduğu sınırlanmış tahmin sonuçlarını karşılaştırmaktadır. Bu bağlamda, bütün modeldeki tahmin edilmiş katsayılar, sınırlanmış modeldeki sözü edilen katsayılar ile istatistiki olarak benzerlik gösterdiğinde, ilişkisiz alternatiflerin bağımsızlığı varsayımı ihlal edilmemiş olacaktır. Ancak, test istatistiği istatistiki açıdan anlamlı olursa, ilişkisiz alternatiflerin bağımsızlığı varsayımı reddedilecek ve sonuç olarak çok durumlu logit modelin uygun olmadığı tespit edilecektir (Cheng ve Long, 2007: 584).

3.1.2.2.2. Çok Durumlu Logit Modelinin Kurulması

Her bir değişkenin en az aralıklı ölçek ile ölçüldüğü varsayılan ve $x' = (x_1, x_2, \dots, x_p)$ şeklinde bir vektör ile temsil edilen p adet bağımsız değişken düşünüldüğünde, ayrıca mevcut çıktı değerlerinin koşullu olasılığı $P(Y = 1 | x) = \pi(x)$ olarak ifade edildiğinde, çoklu lojistik regresyon modelinin logit değeri aşağıdaki eşitlikteki gibi yazılmaktadır:

$$g(x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p \quad (3.55)$$

Şüphesiz ki, anılan eşitlik logit modelin

$$\pi(x) = \frac{e^{g(x)}}{1 + e^{g(x)}} \quad (3.56)$$

şeklinde ifade edildiği durumda geçerli olmaktadır. Sözü edilen bağımsız değişkenlerin birkaçının kesikli değer alması durumunda, bu değişkenlerin modele dâhil edilmesi uygun olmamakta ve aralıklı ölçekler ile ifade edilmesi gerekmektedir. Bu dönüşümü gerçekleştirmenin en elverişli yöntemlerinden birisi, sınıflama ölçeği ile ölçülen değişkenlerin gölge değişken olarak tanımlanmasıdır. Sınıflama ölçeği ile ölçülmüş bir değişkene ait k adet olası değer söz konusu ise, model sabit terim içerdiği sürece, $k - 1$ adet gölge değişken tanımlanması gerekmektedir. Bu gereksinimin cebirsel ifadesi şu şekilde yapılabilmektedir. Bir x_j değişkeni k_j adet olası değere sahip j 'inci bağımsız değişken olmak üzere, daha önce sözü edilen varsayım gereği bu değişken için katsayıları β_{jl} olan ve D_{jl} şeklinde gösterilen k_{j-1} adet gölge değişken tanımlanabilmektedir ($l = 1, 2, \dots, k_{j-1}$). Böylece, j 'inci bağımsız değişkeni kesikli değer alan p değişkene sahip logit modeli,

$$g(x) = \beta_0 + \beta_1 X_1 + \cdots + \sum_{l=1}^{k_j-1} \beta_{jl} D_{jl} + \beta_p X_p \quad (3.57)$$

şeklinde yeniden yazılacaktır (Hosmer ve Lemeshow, 2000: 31-33).

Bir X kümesi tercih edilen nesnelerin, bir S kümesi karar vericilerin ölçülen tutumlarının oluşturan vektörlerin evrensel kümesini, B kümesi alternatifler kümesini temsil etmek, $s \in S$ ve $X \subset B$ olmak üzere, $P(x | s, B)$ bir bireyin tesadüfi olarak x alternatifini tesadüfi olarak tercih etme durumunda gözlenen koşullu olasılığı ifade etmektedir. Sözü edilen durumda, herhangi bir sınamadaki gözlenen tercih çok durumlu dağılım göstermektedir. Bununla birlikte, H kümesi bir bireyin tercih davranışlarını ifade eden kurallar kümesi olmak üzere, bir H modeli gerçek bir ana kütleyi temsil ediyorsa, bu durumda anılan ana kütle içerisinde davranış kurallarının dağılımı gösteren ve H kümesinin (ölçülebilen) alt kümelerinin birinde tanımlanmış bir π olasılığı bulunmaktadır. Bu varsayımlar altında, aşağıdaki eşitlik (45)'te görüldüğü üzere, bir

bireyin x alternatifini tercih etme olasılığı, bu tercihte bulunduğu durumda gözlenen karar kuralının olasılığına eşit olmaktadır:

$$P(x | s, B) = \pi[\{h \in H \mid h(s, B) = x\}] \quad (3.58)$$

Tüm olası B alternatif kümeleri, ölçülen tutumlar s , alternatif kümelerinin elemanları x ve y için, Luce'nin ilişkisiz alternatiflerin bağımsızlığı varsayımı aşağıdaki gibi yeniden yazılabilmektedir:

$$P(x|s, \{x, y\})P(y|s, B) = P(y|s, \{x, y\})P(x|s, B) \quad (3.59)$$

$P(x | s, B)$ olasılığı pozitif olduğunda, yukarıdaki eşitliğe göre $P(x|s, \{x, y\})$ olasılığı da pozitif olacaktır ve sözü edilen eşitlik aşağıdaki şekli alacaktır:

$$\frac{P(y|s, \{x, y\})}{P(x|s, \{x, y\})} = \frac{P(y|s, B)}{P(x|s, B)} \quad (3.60)$$

Bu eşitliğe göre, tüm olası B alternatif kümeleri için $P(x|s, B) > 0$ olmaktadır. Öte yandan, x , y ve z alternatiflerine sahip bir B tercih kümesi düşünüldüğünde, $p_{xy} = P(x|s, \{x, y\})$ ve $p_{xx} = 1/2$ eşitlikleri sağlandığında, (44) eşitliğinden hareket ile,

$$P(y|s, B) = \frac{p_{yx}}{p_{xy}} P(x|s, B) \quad (3.61)$$

$$1 = \sum_{y \in B} P(y|s, B) = \left(\sum_{y \in B} \frac{p_{yx}}{p_{xy}} \right) P(x|s, B) \quad (3.62)$$

eşitlikleri elde edilmektedir. Çoklu tercih olasılıkları, ikili bahis değerleri ile yazılabildiğinden,

$$P(x|s, B) = \frac{1}{\sum_{y \in B} (p_{yx}|p_{xy})} \quad (3.63)$$

eşitliği sağlanmaktadır ve $V(s, x, z)$ bir alternatif fayda fonksiyonunu temsil etmek üzere z alternatifi referans alınarak $V(s, x, z) = \log((p_{yx}|p_{xy}))$ eşitliği sağlandığında, yukarıdaki eşitlik yeniden aşağıdaki şekilde yazılabilmektedir:

$$P(x|s, B) = \frac{e^{V(s,x,z)}}{\sum_{y \in B} e^{V(s,y,z)}} \quad (3.64)$$

$V(s, x, z)$ fonksiyonunda; s değeri ölçülmüş “beğeni etkisi”, x değeri “tercih alternatifi etkisi” ve z değeri “alternatif küme etkisi” olarak tanımlanmaktadır. Yukarıdaki eşitlikteki tercih olasılıklarını belirleyen fayda fonksiyonu, $V(s, x, z) = v(s, x) - v(s, z)$ şeklinde yeniden tanımlanabilecek ve bu durumda sözü edilen eşitlik ise,

$$P(x|s, B) = \frac{e^{v(s,x)}}{\sum_{y \in B} e^{v(s,y)}} \quad (3.65)$$

şeklinde yazılabilecektir. Bu eşitlik çok durumlu logit modeli olarak bilinmekte ve v fonksiyonu “temsili” beğenilerin bir fayda indeksi olarak yorumlanmaktadır (McFadden, 1974: 107, 109-110).

En yüksek faydanın elde edilmesi varsayımından türetilen sözü edilen eşitlik daha genel bir ifade ile aşağıdaki şekilde yeniden düzenlenebilmektedir:

$$P_{ij} = \left[\sum_{k=0}^{m_i} \exp(x'_{ik}\beta) \right]^{-1} \exp(x'_{ij}\beta) \quad (3.66)$$

$$i = 1, 2, \dots, n \quad \text{ve} \quad j = 0, 1, \dots, m_i$$

Bu eşitlikte $x_{i0} = 0$ ifadesinin sağlandığı varsayılmakta ve log-olabilirlik fonksiyonu,

$$\log L = \sum_{i=1}^n \sum_{j=0}^{m_i} y_{ij} \log P_{ij} \quad (3.67)$$

şeklinde tanımlanmaktadır. Log-olabilirlik fonksiyonunun genel içbükeylik özelliğini açıklayabilmek için yukarıdaki eşitliğin β değerine göre türevi alındığında,

$$\frac{\partial \log L}{\partial \beta} = \sum_i \sum_j \frac{y_{ij}}{P_{ij}} \frac{\partial P_{ij}}{\partial \beta} \quad (3.68)$$

eşitliği elde edilmektedir. Bu eşitlikte $\sum_i$ ve $\sum_j$ değerleri sırasıyla $\sum_{i=0}^n$ ve $\sum_{j=0}^{m_i}$ toplam değerlerini temsil etmektedir. Yukarıdaki eşitliğin β değerine göre türevi alındığında,

$$\frac{\partial^2 \log L}{\partial \beta \partial \beta'} = \sum_i \sum_j \frac{y_{ij}}{P_{ij}} \left[\frac{\partial^2 P_{ij}}{\partial \beta \partial \beta'} - \frac{1}{P_{ij}} \frac{\partial P_{ij}}{\partial \beta} \frac{\partial P_{ij}}{\partial \beta'} \right] \quad (3.69)$$

eşitliği elde edilmektedir. Yukarıdaki eşitliğin β değerine göre türevi alınıp düzenlendiğinde,

$$\frac{\partial P_{ij}}{\partial \beta} = P_{ij}(x_{ij} - \bar{x}_i) \quad (3.70)$$

şeklinde yeniden yazılmaktadır. Bu eşitlikte $\bar{x}_i$ ve $\frac{\partial^2 P_{ij}}{\partial \beta \partial \beta'}$ değerleri,

$$\bar{x}_i = \left[\sum_k \exp(x'_{ik}\beta) \right]^{-1} \sum_k \exp(x'_{ik}\beta) x_{ik} \quad (3.71)$$

$$\frac{\partial^2 P_{ij}}{\partial \beta \partial \beta'} = P_{ij}(x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i) \quad (3.72)$$

$$-P_{ij} \left[\sum_k \exp(x'_{ik}\beta) \right]^{-1} \sum_k \exp(x'_{ik}\beta) x_{ik} x'_{ik} + P_{ij} \bar{x}_i \bar{x}_i'$$

eşitlikleri ile gösterilmektedir. Birinci ve ikinci eşitliği üçüncü eşitlikte yerine yazarak,

$$\frac{\partial^2 \log L}{\partial \beta \partial \beta'} = - \sum_i \sum_j P_{ij}(x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i') \quad (3.73)$$

nihai eşitliği elde edilmektedir. Görüldüğü gibi yukarıdaki eşitlik y_{ij} değerlerinden bağımsız bir eşitlik olacaktır. Bu modelde, $P_{ij} > 0$ olduğundan (3.71) matrisi $(x_{ij} - \bar{x}_i)' \alpha = 0$ olmadığı sürece her i, j ve bazı $\alpha \neq 0$ değerleri için negatif belirli bir matris olmaktadır. Bu tür bir durumun gerçekleşmesi son derece zor olduğundan, tüm uygulamalar için çok durumlu logit modeline ait log-olabilirlik fonksiyonunun genel içbükey yapıya sahip olduğu sonucuna varılmaktadır (Amemiya, 1985: 295-296).

Çok durumlu logit modelde, her bir sonuç kategorisinin araştırmacı tarafından keyfi olarak tespit edilebilen “referans” kategorisi ile karşılaştırılması sonucu elde edilen logit eşitlikleri aşağıdaki gibi olacaktır:

$$\ln \frac{P(y_i = 1)}{P(y_i = 0)} = x_i' \beta^{(1)} \quad (3.74)$$

$$\ln \frac{P(y_i = 2)}{P(y_i = 0)} = x_i' \beta^{(2)} \quad (3.75)$$

$$\ln \frac{P(y_i = m)}{P(y_i = 0)} = x_i' \beta^{(m)} \quad (3.76)$$

Mevcut eşitliklerde görüldüğü gibi, referans kategori olarak $y_i = 0$ değeri tercih edilmiştir (Montgomery vd., 2012: 443).

3.1.2.2.3. Çok Durumlu Modellerde Uyum İyiliği

3.1.2.2.3.1. Hosmer-Lemeshow Testi

Çok durumlu modellerde, uyum iyiliğinin değerlendirilmesi amacıyla birçok test istatistiği önerilmektedir. Bu test istatistikleri ki-kare testinin bir kontenjans tablosu için uygulanması ile elde edilmektedir. Sözü edilen kontenjans tablosunda beklenen frekanslar iki grupta ve dağılım varsayımı kullanılarak belirlenmektedir. Daha önce ifade edildiği gibi klasik bir lojistik regresyon modeli uygulamasında $i = 1, \dots, n$ olmak üzere bağımsız gözlemlerin (y_i, x_i) ile temsil edildiği n adet bağımsız gözlem bulunmakta; verilen örnekleme göre β_0 ve β bilinmeyen parametrelerinin tahmin edilmesi gerekmekte ve bu parametrelerin tahmin edilebilmesi için iki temel yöntem kullanılmaktadır. Bu noktada, tercih edilecek tahmin yöntemi (y, x) ikililerinin dağılımına bağlı olmakta ve iki farklı dağılım varsayımı önerilmektedir.

Birinci dağılım varsayımına göre, verilen bir Y değişkeni için X değişkeninin koşullu dağılımının aşağıdaki gibi çok değişkenli normal yapıya sahip olduğu varsayılmaktadır:

$$X | Y = 1 \sim N(\mu_1, \Sigma) \quad (3.77)$$

$$X | Y = 0 \sim N(\mu_0, \Sigma) \quad (3.78)$$

Bu varsayım kümesinin diskriminant fonksiyonu analizinin uygulamasına benzer bir varsayım olduğu görülmektedir. Ayrıca, β_0 ve β bilinmeyen parametrelerinin diskriminant fonksiyonu tahminleri sırasıyla $\tilde{\beta}_0$ ve $\tilde{\beta}$ ile temsil edildiği varsayıldığında söz konusu tahminler μ_0 ve μ_1 değerlerinin en yüksek olasılık tahmininden yola çıkılarak hesaplanabilmektedir. Birinci varsayımın ihlal edildiği durumda, daha az kısıtlayıcı olan ikinci varsayıma göre, $P(Y = 1 | X = x) = \pi(x)$ eşitliği sağlanmaktadır. Bu varsayım altında, X değeri $n \times p$ boyutunda bir matris ve $\pi_i = \pi(x = x_i)$, $i = 1, \dots, n$ olmak üzere, olasılık fonksiyonu,

$$l(X, \beta_0, \beta) = \prod_{i=1}^n \pi_i^{y_i} (1 - \pi_i)^{1-y_i} \quad (3.79)$$

şeklinde tanımlanmaktadır. Böylece, $\hat{\beta}_0$ ve $\hat{\beta}$ en yüksek olabilirlik tahmin değerleri, $\log l(X, \beta_0, \beta)$ ifadesinden gelecek $p + 1$ sayıda olabilirlik eşitliğinin çözümü ile elde edilmektedir. Bu olabilirlik eşitlikleri doğrusal olmayan yapı göstermekte ve yinelemeli bir yöntem vasıtasıyla çözülmektedir. Birinci varsayımda tanımlanan diskriminant fonksiyonu da bir olabilirlik fonksiyonu olmasına rağmen koşullu olmayan bir olabilirlik fonksiyonun en yüksek değerinin elde edilmesi söz konusu olmaktadır.

Uyum iyiliği istatistiğinin belirlenebilmesi adına, $w_i = j$ olmak üzere ve $c_{j-1} \leq \pi < c_j$ ($j = 1, \dots, g$ ve $i = 1, \dots, n$) olduğu durumda W gibi bir tesadüfi değişken vasıtasıyla bir kontenjans tablosu oluşturulduğu varsayılmaktadır. Bu kontenjans tablosunda sabit c_j değerleri $0 = c_0 < c_1 < \dots < c_{g-1} < c_g = 1$ ifadesini sağlamaktadır. Bununla birlikte, $k = 0, 1$ ve $j = 1, 2, \dots, g$ olmak üzere örneklemdeki $(y_i = k, w_i = j)$ ikililerin gerçekleşme sıklığı n_{kj} ile temsil edilmektedir ve (Y, W) ikilileri $2g$ şeklinde çok durumlu bir tesadüfi değişkeni ifade etmektedir. Bu varsayımlar altında, uyum iyiliği istatistiği, gözlenen ve beklenen frekans değerlerinin karşılaştırılması sonucu elde edilmektedir. Beklenen frekans değerlerinin hesaplanabilmesi için ise $p_{kj} = P(Y = k, W = j)$ için ilgili değerlerin belirlenmesi gerekmektedir. Bu noktada, $\pi(X)$ değişkenin tesadüfi bir değişken olduğu ve bu değişkene ait f olasılık fonksiyonunun teorik olarak X değişkeninin olasılık fonksiyonundan elde edildiği varsayılmaktadır.

$2 \times g$ boyutunda bir kontenjans tablosunu oluşturabilmenin alternatif bir yolu da W tesadüfi değişkene ait dağılımın tekdüze hale getirilmesi ile mümkün olmaktadır. Bu bağlamda, $\hat{\pi}$ olasılık tahminine ait n/g değerleri her bir aralığa düşecek ve kesim noktaları bu koşula göre belirlenecektir. Bu değerler elde edilen belirli örnekleme bağlı olacak ve böylece artık katsayıların sabit olması gerekmeyecektir. Bu durumu gerçekleştirebilme adına tahmin değerleri $\hat{\pi}_{(1)} \leq \hat{\pi}_{(2)} \leq \dots \leq \hat{\pi}_{(n)}$ şeklinde sıralı olarak tanımlanmaktadır. Eğer $c_{j-1}^* \leq \hat{\pi}_i < c_j^*$; $\hat{w}_i^* = j$ ve $\hat{J}_j = \{i: c_{j-1}^* \leq \hat{\pi}_i < c_j^*\}$ ise olasılık tahminleri elde edilebilecektir. Bu varsayımlar altında, Hosmer-Lemeshow test istatistiği,

$$\hat{C}_g^* = \sum_{j=1}^g \frac{(\hat{\pi}_{1j}^* - \sum_{r \in \hat{J}_j} \hat{\pi}_r)^2}{(\sum_{r \in \hat{J}_j} \hat{\pi}_r)(1 - g \sum_{r \in \hat{J}_j} \hat{\pi}_r/n)} \quad (3.80)$$

şeklinde tanımlanmaktadır (Hosmer ve Lemeshow, 1980: 1043-1050).

İlk olarak Hosmer ve Lemeshow (1980) tarafından önerilen sözü edilen test istatistiği, tahmin olasılıklarının değerlerinin belirlenmesine dayanan bir gruplama şekli olarak düşünülmektedir. Bu durumda, birinci sütunun en küçük değere ve n 'inci sütunun en yüksek değere karşılık geldiği n sütunlu bir tahmin olasılığı tablosunun olduğu varsayılmaktadır. Hosmer-Lemeshow test istatistiği belirlenirken tahmin edilen olasılıkların yüzdesel değerine ya da sabit değerlerine göre tablo oluşturulmasını öngören iki farklı gruplama stratejisi önerilmektedir. Herhangi bir gruplama stratejisi için $\hat{C}$ Hosmer-Lemeshow uyum iyiliği istatistiği, $g \times 2$ boyutlu bir kontenjans tablosundaki gözlenen ve beklenen değerlere Pearson ki-kare istatistiğinin hesaplanması ile elde edilebilmekte ve daha basit bir ifade ile,

$$\hat{C} = \sum_{k=1}^g \frac{(o_k - n'_k \bar{\pi}_k)^2}{n'_k \bar{\pi}_k (1 - \bar{\pi}_k)} \quad (3.81)$$

şeklinde ifade edilmektedir. Bu eşitlikte n'_k değeri k 'ıncı gruptaki toplam örnek sayısını, c_k değeri k 'ıncı ondalıktaki ortak değişken örüntüsü sayısını,

$$o_k = \sum_{j=1}^{c_k} y_i \quad (3.82)$$

eşitliği c_k ortak değişken örüntüleri arasındaki sonuç değerlerinin sayısını ve

$$\bar{\pi}_k = \sum_{j=1}^{c_k} \frac{m_j \hat{\pi}_k}{n'_k} \quad (3.83)$$

eşitliği ortalama tahmini olasılığı ifade etmektedir. Hosmer ve Lemeshow (1980) geniş kapsamlı bir simülasyon yöntemi kullanarak $J = n$ olduğunda ve doğru logit modeli kullanıldığında, $\hat{C}$ test istatistiğine ait dağılımın $g - 2$ serbestlik derecesinde ki-kare dağılımı ile doğru bir şekilde tahmin edilebildiğini ortaya koymuşlardır (Hosmer ve Lemeshow, 2000: 147-149).

Çok durumlu çerçevede çok sayıda gruplama stratejisi bulunmaktadır. Bu bağlamda önerilen yöntemlerden biri risklerin oransal değeri temelinde ve $\sum_{j=1}^{c-1} \bar{\pi}_{ij} = 1 - \bar{\pi}_{i0}$ şeklinde ifade edilen tahmin olasılıklarının toplamı vasıtasıyla kurulmaktadır.

Genel olarak g sayıda grup için birinci grup en düşük toplam değerine sahip n/g gözlemi, ikinci grup sonraki en düşük toplam değerine sahip n/g gözlemi, vb temsil etmektedir. Bununla birlikte, n/g oranının aralık değeri almadığı durumda tüm gruplar aynı sayıda gözlem sayısına sahip olmayacaktır. Ayrıca, çok sayıda eşit değere sahip gözlem olmadığı sürece, grup sayısındaki küçük farklılıklar test istatistiği değerini çok fazla etkilemeyecektir. O_{kj} ve E_{kj} değerleri her bir gruptaki her bir sonuç kategorisine ait sırasıyla gözlenen frekansları ve tahmin edilen olasılıkları temsil etmek üzere; $k = 1, \dots, g$ ve $j = 0, \dots, c-1$ için bu değerler aşağıdaki eşitliklerdeki gibi tanımlanacaktır:

$$O_{kj} = \sum_{l \in \Omega_k} \bar{y}_{lj} \quad (3.84)$$

$$E_{kj} = \sum_{l \in \Omega_k} \bar{\pi}_{lj} \quad (3.85)$$

Bu eşitliklerde, Ω_k parametresi k grubundaki n/g gözlemine ait indeks değerlerini temsil etmektedir. Böylece, gözlenen ve tahmin edilen frekanslar tablosundan elde edilen Pearson χ^2 istatistiği,

$$C_g = \sum_{k=1}^g \sum_{j=0}^{c-1} (O_{kj} - E_{kj})^2 / E_{kj} \quad (3.86)$$

şeklinde tanımlanmaktadır. İki durumlu lojistik regresyon için gözlenen C_g dağılımı dikkate alınarak ($c = 2$), C_g istatistiğinin $(g-2) \times (c-1)$ serbestlik derecesinde χ^2 dağılımına yaklaşması beklenmektedir (Fagerland, Hosmer ve Bofin, 2008: 4240-4241).

3.1.2.2.4. Çok Durumlu Logit Modelinin En Yüksek Olabilirlik Yöntemi ile Tahmini

Çok durumlu bir modelde, sözü edilen modeli türetmek için hangi yaklaşım kullanılsa kullanılsın elde edilen sonuç olasılığına ait eşitlik aynı olmakta ve bu eşitlik en yüksek olabilirlik tahmini esasına dayanmaktadır. Buna göre, β_2 'den β_j 'ye kadar parametre değerine sahip bir x_i değerine karşılık bir $y_i = m$ değerinin gözlenmesi olasılığı $P(y_i = m \mid x_i, \beta_2, \dots, \beta_j)$ olduğunda ve p_i değeri, i 'nci gözlemde hangi hangi y değerlerinin gözlemlendiğini gösteren olasılık değeri olmak üzere, eğer gözlemler bağımsız ise olabilirlik eşitliği,

$$L(\beta_2, \dots, \beta_j | y, X) = \prod_{i=1}^N p_i \quad (3.87)$$

şeklinde ifade edilmektedir. Bu eşitlik daha geniş bir ifade ile

$$L(\beta_2, \dots, \beta_j | y, X) = \prod_{m=1}^J \prod_{y_i=m} \frac{\exp(x_i \beta_m)}{\sum_{j=1}^J \exp(x_i \beta_j)} \quad (3.88)$$

olarak yazılmaktadır. Yukarıdaki eşitliğin logaritması alınarak log-olabilirlik fonksiyonu kurulmuş ve bu fonksiyonun sayısal yöntemler ile en yüksek değerinin elde edilmesi vasıtasıyla β parametrelerine ait tahminlere ulaşılmış olmaktadır. En yüksek olabilirlik yöntemi ile elde edilen tahminler tutarlı, asimtotik normal ve etkin tahminlerdir (Long, 1997: 156-157).

3.1.3. Koşullu Logit Model

Çok durumlu logit modeller, alternatifler arasındaki tercihlerin tercihte bulunan bireyin özelliklerinin yanı sıra tutumlarına da bağlı olduğu durumlarda koşullu olabilmektedir. Bu noktada, olası bir başka sorun ise, bireyin özellikleri ile tutumlarının etkileşim sorunudur (Borooah, 2002: 2). Koşullu logit analizinin uygulandığı tercih sorunları; üniversite tercihi, meslek tercihi, iş gücüne katılım tercihi, coğrafi bölge tercihi, çocuk sayısı tercihi, konut satın alma tercihi, sahip olunan otomobil sayısı ve markası tercihi, alışveriş güzergâhı ve yeri tercihi gibi örnekler ile sıralanabilmektedir (McFadden, 1974: 106).

Çok durumlu logit modelinde detaylı olarak açıklanan eşitlikteki $v(s, x)$ fonksiyonu ise aşağıdaki şekilde yazılabilmektedir:

$$v(s, x) = \theta_1 v^1(s, x) + \dots + \theta_K v^K(s, x) \quad (3.89)$$

Bu eşitlikte, $v^k(s, x)$ değerleri belirli sayısal fonksiyonları ve θ_k değerleri bilinmeyen parametreleri ifade etmektedir. Mevcut durumda, s_n değeri bir bireyin ölçülen tutumlarının bir vektörü ve B_n değeri bir alternatif kümesi olmak üzere, bir tercih deneyi N farklı sına (s_n, B_n) üzerindeki gözlemleri sağlamaktadır. Alternatif B_n kümesinin x_{jn} adet vektöre sahip J_n ($j = 1, 2, \dots, j_n$) adet alternatif içerdiği, $\theta = (\theta_1, \dots, \theta_K)$, $z_{jn}^k = v^k(s_n, x_{jn})$ ve $z_{jn} = (z_{jn}^1, \dots, z_{jn}^K)$ olduğu varsayıldığında, tercih olasılıkları aşağıdaki eşitliği sağlamaktadır:

$$P_{in} = P(x_{in} | s_n, B_n) = \frac{e^{z_{in}\theta}}{\sum_{j=1}^{J_n} e^{z_{jn}\theta}} \quad (3.90)$$

Sözü edilen tercih deneyi, tekrarsız bir olayın belirli bir durumda R_n kez tekrar edildiğinde ve S_{in} kez i 'nci alternatifin seçildiği gözlemlendiğinde,

$$R_n = \sum_{j=1}^{J_n} S_{jn} = 1 \quad (3.91)$$

olmaktadır ve sözü edilen bu duruma, koşullu logit modeli denilmektedir. Bu yapısı ile koşullu logit model, çok durumlu logit modelin eşit ve muhtemelen sıralı olmayan alternatifler için genelleştirilmiş hali olarak değerlendirilmektedir (McFadden, 1974: 113-114).

Koşullu logit modeli yalnızca tutumların etkilerini kapsamakta ve genelde tercihlerin ön planda olduğu verilere uygulanmaktadır. Bunun yanı sıra çok durumlu logit modeli ise yalnızca özelliklerin etkilerini değerlendirmekte ve çoğunlukla bireylerin özelliklerinin göz önünde tutulduğu veriler için tercih edilmektedir (Borooah, 2002: 47).

3.1.4. Genelleştirilmiş Uç Değer Modeli

İlk olarak McFadden (1978) tarafından önerilmiş olan genelleştirilmiş uç değer modeli, kapalı biçimdeki tesadüfi fayda modellerini geliştirmek için bir teorik altyapı sağlamaktadır. Bazı araştırmacılar için hala karmaşık bir model olarak görülse de, genelleştirilmiş uç değer modelleri yardımıyla yeniden ifade edilen modeller, araştırmacılara hem tahminleme hem de uygulama açısından hesaplama kolaylığı sunmakta ve tüketicilerin toplam faydasını özetleyen genelleştirilmiş uç değer fonksiyonu, ileriki modeller ya da değerlendirmeler için daha uygun bir ölçüm değeri olarak görülmektedir (Daly ve Bierlaire, 2006: 285-286). Genelleştirilmiş uç değer modelinin kapsamında çok durumlu logit modeli, yuvalanmış logit modeli, çapraz yuvalanmış logit modeli ve genelleştirilmiş yuvalanmış logit modeli gibi birçok model yer almaktadır (Abbe, Bierlaire ve Toledo, 2007: 796). Daha geniş bir bakış açısıyla, genelleştirilmiş uç değer modelleri iki temel sınıfa ayrılabilir. Birinci sınıftaki modeller alternatifler arasında çapraz esnekliklere imkân tanımakta; ikinci sınıftaki modeller ise alternatifler arasında çapraz esneklikler için önceden özel yapısal bir ilişki

tanınmaktadır. Birinci sınıflama içerisinde genelleştirilmiş çok durumlu logit modeli, yuvalanmış logit modeli, çapraz yuvalanmış logit modeli ve eşleştirilmiş birleşik logit modeli gibi modeller; ikinci sınıflama içerisinde ise sıralı genelleştirilmiş uç değer modeli, çapraz ilişkili logit modeli gibi modeller yer almaktadır. Böylece, genelleştirilmiş uç değer modelleri, alternatifler için eşit varyansa sahip uç değer hata dağılımını esas alan en yüksek faydayı elde etme amaçlı tüm kapalı-biçim eşitliklerini kapsamaktadır (Koppelman ve Sethi, 2000: 216-217).

Çok durumlu logit modelini bir tesadüfi fayda modelinden türetilmek için bağımsız olarak uç değerlerde dağılmış fayda düzeyleri kullanılmaktadır. Çok durumlu modeli genelleştirebilmek için önerilen yaklaşımlardan biri daha genel çok değişkenli bir uç değer dağılımından başlamak olacaktır. Bu bağlamda, i alternatifinin $q_i = -\beta'z_i$ gibi bir ölçek değerine sahip olduğu varsayıldığında aşağıdaki sonuçlar logit benzeri bir olasılık tercih sisteminin geniş bir sınıfını vermekte ve genelleştirilmiş uç değer modelleri olarak adlandırılmaktadır. Bir $\mathcal{B}$ kümesi sonlu ve boş olmayan tercih kümelerinin bir elemanı ve B bir tercih kümesi olmak üzere, $B = \{1, \dots, m\} \in \mathcal{B}$ için,

1. $H(y_B, z_B, s)$; $y \geq 0$ eşitsizliğinin negatif olmayan, doğrusal ve türdeş bir fonksiyonudur.
2. $\lim_{y_i \rightarrow +\infty} H(y_B, z_B, s) = +\infty$
3. H fonksiyonunun karışık parçalı türevleri mevcut olup sırasıyla pozitif ve negatif olmayan çift ve tek türevleri süreklidir.
4. $B = \{i_1, \dots, i_m\} \in \mathcal{B}$ ise ve $B' = \{i'_1, \dots, i'_m, \dots, i'_{m+n}\} \in \mathcal{B}$ ifadesi $k = 1, \dots, m$ için $z_{i_k} = z_{i'_k}$ eşitliğini sağlıyorsa, bu durumda $H(y_B, z_B, s) = H((y_B, 0, \dots, 0), z_{B'}, s)$ olmaktadır. Eğer H genel uç değer modelini sağlıyorsa ve $B = \{1, \dots, m\}$ ise bu durumda, $F(\varepsilon_B, z_B, s) = \exp\{-H(e^{-\varepsilon_1}, \dots, e^{-\varepsilon_m}, z_B, s)\}$ eşitliği çok değişkenli bir uç değer dağılımını ifade etmektedir.

Genelleştirilmiş uç değer modeli, $0 \leq \sigma < 1$ için

$$H(y_B, z_B, s) = \left[\sum_{i \in B} y_i^{\frac{1}{1-\sigma}} \right]^{1-\sigma} \quad (3.92)$$

olduğu durumda Luce (1959) modeline indirgenmektedir (McFadden, 1981: 227-228).

3.1.5. Yuvalanmış (Nested) Logit Modeli

İktisadi tercihlerde, çapraz alternatiflerin ikame edilebilmesi hususunun algılanan benzerlik ile ilgili olduğu söylenebilmektedir. Bu durumda, birbirine en uzak alternatif kolları en az benzer alternatifler olarak düşünülmektedir. Böylece, düğümlerin katı bir hiyerarşik sıraya sahip olduğu ve tercih edilecek tek bir alternatife ulaşılmıncaya kadar istenilmeyen alternatiflerin sözü edilen hiyerarşik sıraya göre elendiği karar ağacına benzer bir tercih modeli tasarlanabilmektedir. Ağaç uç değer modeli çok durumlu modelin fonksiyonel biçimini genelleştirmektedir ve bu yolla çok durumlu modelin hesaplama üstünlüğünü koruyarak tercih ağacına ait benzerlik yapılarında gözlenen alternatifler arası ikame örneklerini daha uygun hale getirmektedir. Ağaç uç değer modeli, çok durumlu logit modelin yuvalanmış bir sonucu olarak yazılabilmektedir (McFadden, 1981: 231).

Tercih alternatifleri arasındaki ilişkileri tespit etmek adına çok durumlu logit modelin bir uzantısı olarak geliştirilen yuvalanmış logit modeli ilk kez Ben-Akiva (1973) tarafından ulaşım tercihlerinin çözümlenmesi amacıyla kullanılmıştır. Yuvalanmış logit modeli çok durumlu logit modeldeki ilişkisiz alternatiflerin bağımsızlığı yaklaşımını hafiflete amacıyla en yaygın kullanılan model olup alternatif çiftlerine ya da gruplarına ait hata terimlerinin ilişkili olmasına imkân tanımaktadır. Bununla birlikte, bazı önemli durumlarda alternatif çiftleri arasındaki çapraz esnekliğin eşitliği varsayımı, yuvalanmış modeli gerçekçilikten uzaklaştırabilmektedir (Wen ve Koppelman, 2001: 628). Yuvalanmış logit modelin doğrudan genişletilmiş şekli tüm yuvaların alt yuvalara bölünmesi ile gerçekleştirilmekte, sözü edilen bu model her bir yuva için geçerli olmakta ve böylece tüm model tekrarlı olarak kurulmaktadır. Ancak anılan modelin karmaşıklığı nedeniyle model yapısı bir karar ağacı şeklinde gösterilmektedir (Ben-Akiva ve Bierlaire, 1999: 12, 15).

H ($h = 1, \dots, H$) düzeyde düğümlere sahip bir karar ağacı olduğu düşünüldüğünde, sözü edilen karar ağacında h düzeydeki bir düğüm, (i_h, σ_h) şeklinde ikili bir ifade ile temsil edilmektedir. Anılan ifadede $\sigma_h = (i_{h+1}, \dots, i_H)$ şeklinde olup bu eşitlik $h + 1$ düzeyindeki komşu düğümü tanımlamaktadır. Ayrıca, bir B tercih kümesi için, B_{σ_h} değeri B kümesinde σ_h düğümünün altındaki ağaç dalında yer alan temel alternatifler kümesini temsil etmekte ve $B_{\sigma_h} \equiv B$ denkliği sağlanmaktadır. Bu

varsayımlar altında, $B_{\sigma_0}, \sigma_0 \in B$ olmak üzere bu düğümün alabileceği tek değer olmakta ve $C = B_{\sigma_h}$ değeri, alternatifleri σ_h düğümünün altındaki ağaç dalı ile sınırlandırılmış bir tercih kümesi olarak değerlendirilmektedir. Benzer şekilde, $k \leq h$ ve $k > h$ için sırasıyla $C_{\sigma_k} = B_{\sigma_k}$ ve $C = C_{\sigma_k} = C_{\sigma_H}$ eşitlikleri sağlanmaktadır. $x_{i_h\sigma_h}^h$ değerinin $i_h\sigma_h$ düğümünün altındaki alternatiflere ait gözlenen değerlerin ve γ^h değerinin beğeni ağırlıklarının uygun birer vektörü oldukları varsayıldığında, σ_h alternatifi ile ilişkili ölçek değeri, $\sigma_0 \in B_{\sigma_{h-1}}$ olmak üzere,

$$v_{\sigma_0} = \exp \left[\sum_{h=1}^H \gamma^h x_{\sigma_{h-1}}^h \right] \quad (3.93)$$

şeklinde ifade edilmektedir. Öte yandan, σ_0 düğümüne giden yolda $\theta_{\sigma_1}, \dots, \theta_{\sigma_H}$ farklılık parametreleri olduğunda, $0 < \theta_{\sigma_1} \leq \theta_{\sigma_2} \leq \dots \leq \theta_{\sigma_H} \equiv 1$ koşulunu sağlamak üzere ve $h = 2, \dots, H$ için

$$Q(B_{i_h\sigma_h} | B_{\sigma_h}, s) = \frac{\exp \left[\frac{x_{i_h\sigma_h}^h \gamma^h}{\theta_{\sigma_h}} + \frac{y_{i_h\sigma_h} \theta_{i_h\sigma_h}}{\theta_{\sigma_h}} \right]}{\sum_{i \in B_{\sigma_h}} \exp \left[\frac{x_{i_h\sigma_h}^h \gamma^h}{\theta_{\sigma_h}} + \frac{y_{i_h\sigma_h} \theta_{i_h\sigma_h}}{\theta_{\sigma_h}} \right]} \quad (3.94)$$

$$y_{i_{h+1}\sigma_{h+1}} = \ln \sum_{i \in B_{\sigma_h}} \exp \left[\frac{x_{i_h\sigma_h}^h \gamma^h}{\theta_{\sigma_h}} + \frac{y_{i_h\sigma_h} \theta_{i_h\sigma_h}}{\theta_{\sigma_h}} \right] \quad (3.95)$$

eşitlikleri elde edilmektedir. Bu eşitliklerdeki $y_{i_h\sigma_h}$ ifadesi $i_h\sigma_h$ düğümünün altındaki dallara ait “kapsayıcı” değer olarak tanımlanmaktadır. Bu modelin genelleştirilmiş uç değer modeli ile tutarlı olabilmesi için gerekli ve yeterli koşul, her bir kapsayıcı değer birim aralık üzerinde yer alması olacaktır. Bazı tahmin yöntemleri için, $B_{\sigma_h} = \{1, \dots, m\}$ olmak üzere,

$$\beta_{\sigma_h}^h = \left[\frac{\gamma^h}{\theta_{\sigma_h}}, \frac{\theta_{1\sigma_h}}{\theta_{\sigma_h}}, \dots, \frac{\theta_{m\sigma_h}}{\theta_{\sigma_h}} \right] \quad (3.96)$$

$$z_{i_h\sigma_h}^h = [x_{i_h\sigma_h}^h, 0, \dots, 0, y_{i_h\sigma_h}, 0, \dots, 0]$$

şeklinde bir gösterim ile yukarıdaki eşitlikteki yuvalanmış logit modeli,

$$Q(i_h|\sigma_h, z_{i_h\sigma_h}^h, \beta_{\sigma_h}^h) = \frac{\exp[z_{i_h\sigma_h}^h \beta_{\sigma_h}^h]}{\sum_{i \in B_{\sigma_h}} \exp[z_{i_h\sigma_h}^h \beta_{\sigma_h}^h]} \quad (3.97)$$

$$y_{\sigma_h} = \ln \sum_{i \in B_{\sigma_h}} \exp[z_{i_h\sigma_h}^h \beta_{\sigma_h}^h]$$

eşitliklerindeki gibi yeniden yazılabilmektedir (McFadden, 1981: 238-241).

Yuvalanmış logit modelin karmaşık yapısını gidermek için karar ağacı şeklinde gösterilmesi önerilmişse de çok düzeyli sınamalarda anılan karar ağacı yapısı da karmaşıklık sorununu giderememektedir. Yuvalanmış logit modelindeki parametre sayısının bazı yollar ile sınırlandırılması da modelde karmaşıklığa yol açmaktadır. Üç alternatife (a , b , c) sahip iki düzeyli bir yuvalanmış logit modelinde farklılık parametreleri için $\theta_a = \theta_c \equiv \theta$ şeklinde yeni bir sınırlama getirilmesinin model tahminini kolaylaştıracağı düşünülmektedir (Brownstone ve Small, 1989: 68).

3.1.6. Çapraz Yuvalanmış Logit Modeli

Bazı uygulamalarda, genelleştirilmiş uç değer sınıfından türetilen ve yuvalanmış logit modelin genelleştirilmiş hali olarak kabul edilen çapraz yuvalanmış logit modeli daha uygun bir yaklaşım olarak görülmektedir. Klasik yuvalanmış modelinde olduğu gibi gruplar arası benzerlikler varsayımının yerine çapraz yuvalanmış modelde farklı ve birleşik yollar arasındaki çapraz benzerlikler göz önünde tutulmaktadır. Klasik yuvalanmış modelde esnek olmayan grup benzerlikleri söz konusu iken çapraz yuvalanmış yapı, yollar arasında ikili benzerlerin farklı yöntemler ile ölçülmesine imkân tanımaktadır (Vovsha, 1997: 6).

Çapraz yuvalanmış logit modeli, ilk olarak McFadden (1978) tarafından genelleştirilmiş uç değer modelin özel bir durumu olarak gösterilmiş, Small (1987) tarafından kalkış zamanı tercihi ve Vovsha (1997) tarafından bir ulaşım yolu seçimini modellemek amacıyla uygulanmıştır. Çapraz yuvalanmış logit modeli, klasik yuvalanmış logit modelin genişletilmiş bir şekli olup bu yeni modelde her bir alternatif birden fazla bir yuvaya ait olabilmektedir. Klasik yuvalanmış logit modele benzer olarak C_n tercih kümesi, C_{mn} ile temsil edilen M ($1, \dots, m$) sayıda alt yuvaya ayrılmaktadır. Diğer taraftan, her bir alternatif ve ilgili yuva için m yuvasındaki i

alternatifinin aitlik düzeyini gösteren $\alpha_{im} (0 \leq \alpha_{im} \leq 1)$ parametreleri tanımlanmaktadır. Bu durumda, i alternatifinin faydası,

$$U_{imn} = \tilde{V}_{in} + \tilde{\epsilon}_{in} + \tilde{V}_{C_{mn}} + \tilde{\epsilon}_{C_{mn}} + \ln \alpha_{im} \quad (3.98)$$

şeklinde ifade edilmektedir. Bu eşitlikte $\tilde{\epsilon}_{in}$ ve $\tilde{\epsilon}_{C_{mn}}$ hata terimlerinin bağımsız olduğu varsayılmaktadır. Ayrıca $\tilde{\epsilon}_{in}$ hata teriminin bağımsız ve özdeş olarak Gumbel dağılımına ve sabit ölçek parametresine; $\tilde{\epsilon}_{C_{mn}}$ hata teriminin dağılımı ile ilişkili olarak ise $\max_{j \in C_{mn}} U_{jmn}$ tesadüfi değişkeninin Gumbel dağılımına ve μ gibi bir ölçek parametresine sahip olduğu varsayılmaktadır. Bu durumda, n bireyinin i alternatifini tercih etme olasılığı,

$$P(i|C_n) = \sum_{m=1}^M P(C_{mn}|C_n) P_n(i|C_{mn}) \quad (3.99)$$

şeklinde tanımlanmaktadır. Bu tercih olasılığı,

$$P(C_{mn}|C_n) = \frac{e^{\mu V_{C_{in}}}}{\sum_{i=1}^M e^{\mu V_{C_{in}}}} \quad (3.100)$$

$$P(i|C_{mn}) = \frac{\alpha_{im} e^{\tilde{V}_{in}}}{\sum_{j \in C_{mn}} \alpha_{im} e^{\tilde{V}_{in}}} \quad (3.101)$$

$$V_{C_{mn}} = \tilde{V}_{C_{mn}} + \ln \sum_{j \in C_{mn}} \alpha_{jm} e^{\tilde{V}_{jn}} \quad (3.102)$$

eşitliklerinin sağlandığı durumda geçerli olmaktadır (Ben-Akiva ve Bierlaire, 1999: 16-17).

3.1.7. Çok Durumlu Probit Modeli

Çok durumlu probit modeli; sıfır ortalamaya, keyfi varyans-ortak değişken matrisine ve ortak çok değişkenli normal dağılıma sahip hata terimleri içeren bir tesadüfi fayda fonksiyonu modeli olarak tanımlanmaktadır. Bu durumda çok durumlu probit modelinde hata terimlerinin varyansları farklı ve ilişkili olabilmektedir (Daganzo, 1979: 17). Bununla birlikte, çok durumlu probit modeli ile çoklu zaman dönemleri boyunca iki durumlu tercih kararının gözlendiği çok dönemli probit modelleri arasında yakın bir benzerliğin olduğu bilinmektedir (McCulloch ve Rossi, 1994: 207-208). Çok durumlu probit modeli, alternatiflerin gizli faydaları arasından koşullu ortak değişken

özelliğine esneklik tanınması sayesinde tercih davranışı modelleri arasında ilgi çekici modeller arasında olmasına rağmen uygulamada en yüksek olabilirliği tahmincisinin hesaplama zorluğu sorunu ile karşılaşıldığından literatürde yaygın biçimde kullanılmamaktadır. Bununla birlikte, çok durumlu probit modelin ilişkisiz alternatiflerin bağımsızlığı varsayımını hafifletmesi kural olarak çok durumlu logit modele karşı bir üstünlük olarak kabul edilmektedir (Geweke, Keane ve Runkle, 1994: 694).

Bir olayın gerçekleşmesi için $s + 1$ sayıda aşamanın olduğu, bir alternatifin bu aşamalardan en az birinde gözlemlendiği, bilinmeyen parametrelerin tahmin edilebileceği ve x_α ($\alpha = 1, 2, \dots, m$) ile temsil edilen gözlemlerin m farklı zamanda gerçekleştiği varsayıldığında, i 'nci aşamada ($i = 1, 2, \dots, s$) bir alternatifin gözlenebilmesi geçen zaman negatif olmayan ξ_i gibi bir tesadüfi değişken ile gösterildiğinde, bu noktada araştırmacının temel sorunu her bir ξ_i için ortalama λ_i değerlerini tahmin etmek olacaktır. Ayrıca bir gözlem için $1, 2, \dots, r$ aşamalarında geçen toplam zaman $\eta_r = \sum_{i=1}^r \xi_i$ gibi bir tesadüfi değişken ile gösterilmektedir. Eğer $\mu_r = E(\eta_r)$ ($r = 1, 2, \dots, s$) ise λ_i ifadesinin tahmin değeri μ_r değeri tahmin edilerek elde edilmektedir ve bu tahminlerin farkları alınmaktadır. Her bir r için η_r değerinin dağılım fonksiyonu G_r sürekli ise,

$$P(\eta_r < x) = P(\eta_r \leq x) = G_r(x) \quad (3.103)$$

olacak, i 'nci aşamada ve x zamanında bir alternatifin tesadüfi olarak seçilme olasılığı $\pi_i(x)$ değeri aşağıdaki eşitliklerdeki yazılacaktır:

$$\pi_1(x) = 1 - G_1(x) \quad (3.104)$$

$$\pi_i(x) = G_{i-1}(x) - G_i(x) \quad (i = 2, 3, \dots, s) \quad (3.105)$$

$$\pi_{s+1}(x) = G_s(x) \quad (3.106)$$

Öte yandan, η_i değerlerinin dağılım fonksiyonları sürekli olduğu ve $\eta_i < x$ ifadesi $\eta_{i-1} < x$ ifadesine eşit olduğu sürece $\pi_i(x)$ değeri,

$$\pi_i(x) = P(\eta_{i-1} < x, \eta_i > x) \quad (3.107)$$

$$= P(\eta_{i-1} < x) - P(\eta_{i-1} < x, \eta_i \leq x)$$

$$= P(\eta_{i-1} < x) - P(\eta_i < x)$$

şeklinde yeniden yazılabilmektedir. Şüphesiz ki, söz konusu eşitlikler, gözlenen alternatiflerin olabilirliğini η_r değerlerinin dağılım fonksiyonları cinsinden açıklamaya imkân tanımaktadır. İlaveten, ξ_i ve η_r değişkenlerinin normal dağıldığı ve tesadüfi değişkenin σ_i gibi bir standart sapmaya sahip olduğu varsayıldığında,

$$G_r(x) = \phi(z_r) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{z_r} e^{-\frac{1}{2}t^2} dt \quad (3.108)$$

şeklinde elde edilen dağılım fonksiyonu eşitliği ile çok durumlu probit modeli tamamlanmış olacaktır. Sözü edilen yukarıdaki eşitlikte,

$$z_r = \frac{x - \mu_r}{\theta_r} \quad (3.109)$$

$$\theta_r^2 = \text{var}(\eta_r) = \text{var}\left(\sum_{i=1}^r \xi_i\right) \quad (3.110)$$

eşitliklerinin geçerli olduğu varsayılmaktadır (Aitchison ve Silvey, 1957: 131-132).

Çok durumlu bir probit modelinin yapısal eşitlikleri şu şekilde ifade edilmektedir:

$$U_{ij} = x'_{ij}\beta + \varepsilon_{ij}, \quad j = 1, \dots, J, [\varepsilon_{i1}, \varepsilon_{i2}, \dots, \varepsilon_{ij}] \sim N[0, \Sigma] \quad (3.111)$$

Benzer şekilde, log-olabilirlik içerisindeki bir q alternatifinin tercih edilme olasılığını ifade eden terim,

$$P[\text{tercih}_{iq}] = P[U_{iq} > U_{ij}, \quad j = 1, \dots, J, \quad j \neq q] \quad (3.112)$$

şeklinde yazılabilecek ve $J - 1$ adet diğer tercih arasından bu alternatifin seçilme olasılığı aşağıdaki eşitlikteki gibi özetlenecektir:

$$P[\text{tercih}_{iq}] = P[\varepsilon_{iq} - \varepsilon_{iq} < (x_{iq} - x_{i1})'\beta, \dots, \varepsilon_{ij} - \varepsilon_{iq} < (x_{iq} - x_{ij})'\beta] \quad (3.113)$$

Bu eşitlikteki olasılık, $J - 1$ değişkenli bir normal dağılım içerisindeki birikimli bir olasılık fonksiyonu şeklinde tanımlanmaktadır. Çok durumlu probit modeli sınırlı olmayan $(J - 1) \times (J - 1)$ adet korelasyon yapısına ve $J - 2$ adet bağımsız standart sapma değeri elde edilmesine imkân tanımaktadır (Greene, 2008: 850).

3.1.8. Ardışık Modeller

Sonuç değişkenine ait kategorilerin aşama aşama ve birbirini izleyen sırada seyrettiği bir durumda, $r = 1, \dots, k - 1$ olmak üzere u_r ile temsil edilen gizli

değişkenlerin $u_r = x'\beta + \epsilon_r$ ifadesinde olduğu gibi doğrusal bir yapıda olduğu varsayılmaktadır. Sözü edilen ifadede ϵ_r değeri D gibi bir dağılım fonksiyonuna ait tesadüfi bir değişkeni göstermektedir. Bu varsayımlar altında ve θ_r değeri eşik noktasını temsil etmek üzere sonuç değerlerine ait aşağıdaki şekilde bir döngüden bahsedilmektedir:

$$y = r \quad \text{verilen } y \geq r \text{ için eğer } u_r \leq \theta_r \quad (3.114)$$

ya da

$$y > r \quad \text{verilen } y \geq r \text{ için eğer } u_r > \theta_r \quad (3.115)$$

Bu eşitlikte ardışık döngü verilen bir r kategorisine ulaşıldığında bu kategoriden $r + 1$ kategorisine geçişi modellemektedir. Söz konusu modelde bir geçiş süreci geçişi belirleyen gizli değişkenin yalnızca eşik noktasının üzerinde bir değer alması durumunda gerçekleşmektedir. Bu durumda ise kategorik özellikler dikkate alınmaktadır. Ardışık modelin benzer modellerden temel farkının geçiş sürecinin koşullu modellenmesi olduğu kabul edilmekte ve ardışık döngüye ait her bir aşamada iki durumlu bir kararın verildiği varsayılmaktadır. Ayrıca, ardışık döngüde verilen bir r kategorisine ulaşıldığında, sürecin son kategoride durması ya da daha yüksek kategoriler ile devam etmesi kararı verilmekte ve yalnızca nihai sonuç kategorisi gözlenebilmektedir. Her bir aşamadaki iki durumlu kararlar, birikimli dağılım fonksiyonu kabul edilerek ve klasik regresyon yaklaşımı kullanılarak modellenmektedir. Bu varsayımlar ışığında basit ardışık model,

$$p(y = r | y \geq r, x) = D(\theta_r - x'\beta) \quad (3.116)$$

ve bu modelin olasılıkları,

$$p(y = r | x) = D(\theta_r - x'\beta) \prod_{i=1}^{r-1} (1 - D(\theta_i - x'\beta)) \quad (3.117)$$

şeklinde ifade edilmektedir. Bu eşitlikte tanımlanan ardışık modele literatürde sürdürme oranı modeli de denilmektedir. Bu eşitlikte, aşamalı geçişlerin her bir aşamada aynı değeri alan β parametresi tarafından belirlendiği varsayılmaktadır. Ancak $p(y \leq r | x)$ ile ifade edilen birikimli olasılıklar yerine $p(y = r | y \geq r, x)$ şeklinde ifade edilen kategorik değişiklikler, doğrusal terim $\theta_r - x'\beta$ ve D dönüşümü vasıtasıyla tespit edilmektedir (Tutz, 1991: 278-27).

3.1.9. Çok Değişkenli Modeller

Çok değişkenli bir kategorik yanıt modeli, iki ya da daha fazla kesikli bağımlı değişkenin ortalık olasılık dağılımını açıklamaktadır. Örneğin, y_1 ve y_2 iki durumlu bağımlı değişkenleri varsayıldığında, $P_{jk} = P(y_1 = j, y_2 = k)$ olmak üzere ortak olasılık dağılımına ait değerler P_{00}, P_{01}, P_{10} ve P_{11} şeklinde tanımlanacak ve çok değişkenli model P_{11}, P_{10} ve P_{01} olasılık değerlerinin bağımsız değişkenler ve parametrelerin bir fonksiyonu olarak tanımlanmasıyla kurulmuş olacaktır. Şüphesiz ki, P_{00} olasılığı, diğer üç olasılık toplamının 1'den çıkarılması ile elde edilebilmektedir. Bu noktada, çok değişkenli kategorik yanıt modeli çok durumlu kategorik modelin özel bir durumu olarak görülebilmektedir. Öyle ki, ortak olasılık dağılımı ışığında açıklanan çok değişkenli model, bir kesikli bağımlı değişken için P_{00}, P_{01}, P_{10} ve P_{11} olasılık değerlerini alan çok durumlu modele eş değer görülmektedir. Böylece, çok durumlu modeller için daha önce açıklanan istatistiksel çıkarımlar herhangi bir değişikliğe ihtiyaç duyulmadan çok değişkenli modeller için de geçerli olmaktadır (Amemiya, 1985: 311).

3.1.9.1. Çok Değişkenli Logit Modeli

Çok boyutlu bir kontenjans tablosu düşünüldüğünde, $y_1, \dots, y_q$ şeklinde ifade edilen ve $I_1, \dots, I_q$ sıralı olmayan alternatif değerlerinden herhangi birini alabilen q sayıda değişken olduğu varsayıldığında, toplam ihtimallerin sayısı,

$$Q = \prod_{k=1}^q I_k \quad (3.118)$$

olacaktır. Ayrıca, $i_1 = 1, \dots, I_1$ ve $i_q = 1, \dots, I_q$ için y_1 değişkeninin i_1 'nci değeri alma olasılığı, y_2 değişkeninin i_2 'uncu değerini alma olasılığı, vb gibi ortak olaylara ait olasılıklar $P_{i_1}, \dots, P_{i_q}$ şeklinde gösterildiğinde, genel olarak sözü edilen olasılık değerleri,

$$P_{i_1, \dots, i_q} = \frac{e^{\theta_{i_1, \dots, i_q}}}{\sum_{i_1=1}^{I_1} \dots \sum_{i_q=1}^{I_q} e^{\theta_{i_1, \dots, i_q}}} \quad (3.119)$$

şeklinde gösterilmektedir. Bu eşitlik çok durumlu logit modelin simetrik şekli olarak tanımlanmaktadır (Nerlove ve Press, 1973: 22-23).

3.1.9.2. Çok Değişkenli Probit Modeli

İki durumlu gözlemler sürekli bir değişkenin işaretine göre sıklıkla modellenebilmektedir. Her bir i gözlemi için d_{it} gibi gözlem sonuçları mevcut ise tek değişkenli probit model çok değişkenli probit söz konusu olmaktadır (Kiefer, 1982: 161). Çok değişkenli probit yaklaşımı, ilk olarak Ashford ve Sowden (1970) tarafından bazı risk faktörleri üzerindeki model bağımlılığını modellemek amacıyla öne sürülmüştür. Çok değişkenli probit modeli, açıklayıcı değişkenlerin etkisini göz önüne aldığı için bir korelasyon uzantısı olarak davranmakta, öte yandan birden fazla olayı eş anlamlı olarak değerlendirdiği için klasik probit modelinin bir uzantısı olarak da görülebilmektedir. Bununla birlikte, marjinal hata olasılıklarının ve bu olasılıklar arasındaki ilişkiyi normal bir dağılım vasıtasıyla modellenerek ortak olasılık değerleri elde edildiğinden çok değişkenli probit modeli marjinal yaklaşıma örnek olarak da gösterilebilmektedir (Lesaffre ve Kaufmann, 1992: 805).

Klasik probit modelinin alternatif bir gösterimi olarak S gibi belirli bir sonuç sistemine ait y gibi bazı gerçek ya da varsayıma dayanan nicel sonuçlar içeren yeni bir model önerilmektedir. Sözü edilen model,

$$y = x(z) - \psi \quad (3.120)$$

şeklinde ifade edilebilmektedir. Bu eşitlikte ψ değeri S sistemindeki hata terimini temsil etmekte ve $\phi(\psi)$ dağılım fonksiyonuna ait bir tesadüfi değişken olarak görülmektedir. Benzer şekilde, yukarıdaki eşitlikteki model $S_1, \dots, S_k$ şeklinde k sayıda sonuç sistemi için ve $i = 1, \dots, k$ olmak üzere,

$$y_i = x_i(z) - \psi_i \quad (3.121)$$

şeklinde yazılabilmektedir. Böylece, hem $y = (y_1, \dots, y_k)$ yanıt değişkenleri hem de $\psi = (\psi_1, \dots, \psi_k)$ hata terimleri çok değişkenli bir yapıya sahip olmaktadır. Çok değişkenli model yapısında ψ hata teriminin $T(\psi)$ gibi bir ortak dağılım fonksiyonuna sahip olduğu kabul edilmektedir. Birden fazla bir sistemin ortak yanıtları göz önüne alındığında, ilgili hata terimlerinin ortak dağılımının incelenmesi gerekmektedir. Bu durumu gerçekleştirebilmek için de klasik probit analizinin çok değişkenli yapıda

genelleştirilmesi gerekmektedir. Standartlaştırılmış iki değişkenli normal dağılıma ve ρ korelasyon katsayısına sahip bir dağılım fonksiyonu $F(w_1, w_2; \rho)$ ile gösterilirse, ikili yanıt değerlerinin olasılığı,

$$P_{ij}^{11}(z) = F[x_i(z), x_j(z); \rho_{ij}] \quad (3.122)$$

şeklinde yazılabilmektedir. Ortak yanıt değişkenine sahip ikiden fazla sistem için de benzer eşitlikleri yazmak mümkün olmakta ve yalnızca $\{\rho_{ij}\}$ korelasyon değerleri değiştirilebilmektedir (Ashford ve Sowden, 1970: 537-538).

Ashford ve Sowden (1970)'in ortaya koyduğu yukarıdaki eşitlikteki çok değişkenli probit modeli en yüksek olabilirlik yöntemi ile tahmin edilebilmektedir. Daha açık bir ifade ile Y_t ve Z_t yalnızca 0 ya da 1 değerlerini alan ve ortak olasılık dağılımına sahip iki durumlu değişkenler olmak üzere,

$$P_Y(t) = P[Y_t = 1] = \phi(\beta'x_t) \quad (3.123)$$

$$P_Z(t) = P[Z_t = 1] = \phi(\gamma'x_t) \quad (3.124)$$

$$P_{YZ}(t) = P[Y_t = 1, Z_t = 1] = \phi(\beta'x_t, \gamma'x_t, \rho) \quad (3.125)$$

şeklinde ifade edilmektedir. Bu eşitliklerde; x_t parametresi K boyutlu bir vektörün bilinen sabit değerlerini, β ve γ değerleri K boyutlu bir vektörün bilinmeyen parametrelerini, ϕ değeri standart normal bir değişkenin dağılım fonksiyonunu, $t = 1, \dots, T$ olmak üzere T sabit bir sayıyı ve F fonksiyonu standart normal değişkenlerin ρ korelasyon katsayısına sahip iki değişkenli dağılım fonksiyonunu temsil etmektedir. Ayrıca, (Y_t, Z_t) değerlerinin geçici bağımsız olduğu varsayılmaktadır. Bu noktada, temel sorun, n_t sayıda bağımsız gözlem için $i = 1, \dots, n_t$ olmak üzere (Y_{ti}, Z_{ti}) ve (Y_t, Z_t) varsayımlarının ışığında β , γ ve ρ değerlerinin nasıl tahmin edileceğini belirleyebilmektedir. En yüksek olabilirlik tahmincisini elde edebilmek adına,

$$r_Y(t) = \sum_{i=1}^{n_t} Y_{ti} \quad (3.126)$$

$$r_Z(t) = \sum_{i=1}^{n_t} Z_{ti} \quad (3.127)$$

$$r_{YZ}(t) = \sum_{i=1}^{n_t} Y_{ti} Z_{ti} \quad (3.128)$$

eşitlikleri tanımlandığında, $F_t = F(\beta'x_t, \gamma'x_t, \rho)$ olmak üzere log-olabilirlik fonksiyonu,

$$\begin{aligned} \log L = \sum_{t=1}^T \{ & r_{YZ}(t) \log F_t + [r_Y(t) - r_{YZ}(t)] \cdot \log[\phi(\beta'x_t) - F_t] \\ & + [r_Z(t) - r_{YZ}(t)] \cdot \log[\phi(\gamma'x_t) - F_t] \\ & + [n_t - r_Y(t) - r_Z(t) + r_{YZ}(t)] \cdot \log[1 - \phi(\beta'x_t) - \phi(\gamma'x_t) \\ & + F_t] \} \end{aligned} \quad (3.129)$$

şeklinde yazılmaktadır. Bu bağlamda, en yüksek olabilirlik tahmincilerinin tutarlı ve asimtotik normal olduğu ($n_t \rightarrow \infty$) görülmektedir (Amemiya, 1972: 940-941).

Literatürde çok değişkenli probit modeli, bireylerin belirli bir eylemi gerçekleştirme ya da gerçekleştirme tercihinin karşılık gelmektedir. Bu durumda, sözü edilen tercihler bireylerin bir takım özelliğine bağlı olarak değişmektedir. Çok durumlu probit modeli, alternatiflere ait tutumları temsil eden X değişkenlerinin standart normal dağılan bir I indeksi ile özetlendiğini varsaymakta ve X değişkeni ile I indeksi arasındaki ilişkinin doğrusal olduğunu varsaymaktadır. Bununla birlikte, her bir alternatif için yanıt değerinin I_i gibi bir eşik değerine sahip olduğu kabul edilmektedir. Bu bağlamda, $\phi(u)$ fonksiyonu birim normal değişkenin olasılık yoğunluk fonksiyonunu temsil etmek üzere, i 'nci alternatifin pozitif bir yanıt değeri içermesi olasılığı,

$$P_i = P(y_i = 1) = P(I < I_i) = \int_{-\infty}^{I_i} \phi(u) du \quad (3.130)$$

şeklinde tanımlanmaktadır. Ayrıca, m sayıda gözlem için istatistiksel olarak bağımsız yanıtlar verildiği varsayıldığında, yanıtlardan $(y_1, y_2, \dots, y_m)$ şeklinde bir grupta elde edebilmek için gerekli log-olabilirlik değeri,

$$L = \prod_{i=1}^m [P_i^{y_i} (1 - P_i)^{1-y_i}] \quad (3.131)$$

şeklinde ifade edilmektedir. Her bir P_i değeri, β parametrelerinin bir fonksiyonu olduğundan log-olabilirlik değeri de β parametrelerinin bir fonksiyonu olacak ve sözü edilen parametreler en yüksek olabilirlik yöntemi ile tahmin edilecektir. Log-olabilirlik değerinin en yüksek değerini elde edebilmek için çözülmesi gereken eşitlikler,

$$\sum_{i=1}^m \left[\frac{y_i}{P_i} - \frac{1 - y_i}{1 - P_i} \right] \phi(I_i) = 0 \quad (3.132)$$

$$\sum_{i=1}^m \left[\frac{y_i}{P_i} - \frac{1 - y_i}{1 - P_i} \right] \phi(I_i) x_{ij} = 0 \quad (3.133)$$

şeklinde yazılmaktadır. Gözlem değeri m yeterince büyük olduğunda, $\hat{\beta}$ tahmin değerleri de tatmin edici olabilmektedir (Rao ve Winter, 1978: 362-363).

3.1.9.3. Log-Doğrusal Model

Tesadüfi mutlak sıralamayı açıklayan genel doğrusal model dışında modellerin de olduğu bilinmektedir. Genel doğrusal modele alternatif mevcut yaygın yöntemlerden birisi de birleşim istatistiğini iki bileşene ayırmak amacıyla log-doğrusal modelin kullanılmasıdır. Sözü edilen iki bileşenlerden biri mutlak tesadüfi sıralamayı, diğeri de birleşimdeki en yüksek bileşeni açıklamaktadır (McCullagh, 1980: 121). Lojistik regresyon modelleri yapısal olarak daha çok sıralı regresyon modeline benzemektedir, çünkü anılan her iki modelde de açıklayıcı değişkenler kullanılarak sonuç değişkeninin tahmini söz konusu olmaktadır. Bu durumun aksine, log-doğrusal modellerin, her bir sınıflamanın bir sonuç değişkeni olarak görüldüğü olasılık tabloları için daha uygun oldukları düşünülmektedir. Bu yönüyle log-doğrusal analiz, korelasyon analizine daha çok benzemektedir. Log-doğrusal yaklaşım sonuç değişkeninin diğer değişkenlere göre modellenmesinin yerine değişken çiftleri arasındaki ilişkiyi araştırmaktadır (Agresti ve Finlay, 2009: 503).

Bir log-doğrusal model, çok değişkenli bir modelin özel bir durumu olarak düşünülmektedir. Çok değişkenli modelde açıklanan ortak olasılık dağılımı ve dört olasılık değeri düşünüldüğünde, bunun yanında sözü edilen modelde olasılıklar üzerinde herhangi bir bağımsız değişken ve sabit terim olmadığı varsayıldığında model bu dört olasılık değerinin herhangi üçü vasıtasıyla açıklanabilecektir (Amemiya, 1985: 314). Çok değişkenli modele benzer şekilde, 2 x 2 boyutunda bir kontenjans tablosunda A_1 ve A_2 gibi aynı kategoriye sahip iki değişken olduğu düşünüldüğünde ve olasılık değerlerinin $p_{11}, p_{12}, p_{21}, p_{22}$ şeklinde olduğu kabul edildiğinde, sözü edilen olasılık değerlerine ait çapraz çarpım oranları,

$$\alpha_1 = \frac{p_{11}p_{12}}{p_{21}p_{22}} \quad (3.134)$$

$$\alpha_2 = \frac{p_{11}p_{21}}{p_{12}p_{22}} \quad (3.135)$$

$$\alpha_3 = \frac{p_{11}p_{22}}{p_{12}p_{21}} \quad (3.136)$$

şeklinde yazılmaktadır. Bu çapraz çarpım oranlarının logaritmaları alındığında,

$$\log \alpha_1 = \log p_{11} - \log p_{12} - \log p_{21} - \log p_{22} \quad (3.137)$$

$$\log \alpha_2 = \log p_{11} - \log p_{12} + \log p_{21} - \log p_{22} \quad (3.138)$$

$$\log \alpha_3 = \log p_{11} - \log p_{12} - \log p_{21} + \log p_{22} \quad (3.139)$$

eşitlikleri elde edilmektedir. Bu üç eşitlik, $\sum p_{ij} = 1$ olmak üzere çözüldüğünde dört olasılık değeri tanımlanmış olacaktır. Bu bağlamda, hücre olasılıklarının doğal logaritmaları ışığında kurulan log-doğrusal model, varyans analizi modelleri ile benzerlik göstermekte ve u olasılıklara ait logaritmaların genel ortalaması olmak üzere,

$$\log p_{ij} = u + u_{1(i)} + u_{2(j)} + u_{12(ij)} \quad i = 1, 2; j = 1, 2 \quad (3.140)$$

$$u = \frac{1}{4}(\log p_{11} + \log p_{12} + \log p_{21} + \log p_{22}) \quad (3.141)$$

şeklinde yazılmaktadır. Birinci değişkenin i düzeyindeki ve ikinci değişkenin j düzeyindeki olasılıklara ait logaritmaların ortalamaları,

$$u + u_{1(i)} = \frac{1}{2}(\log p_{i1} + \log p_{i2}) \quad i = 1, 2 \quad (3.142)$$

$$u + u_{2(j)} = \frac{1}{2}(\log p_{1j} + \log p_{2j}) \quad j = 1, 2 \quad (3.143)$$

şeklinde tanımlanmaktadır. Ayrıca, $l_{ij} = \log p_{ij}$ eşitliği tanımlandığında; genel ortalama değeri,

$$u = \frac{l_{++}}{4} = \sum_{i,j} \frac{l_{ij}}{4} \quad (3.144)$$

temel etki değerleri,

$$u_{1(i)} = \frac{l_{i+}}{2} - \frac{l_{++}}{4} \quad (3.145)$$

$$u_{2(j)} = \frac{l_{+j}}{2} - \frac{l_{++}}{4} \quad (3.146)$$

ve etkileşim terimi,

$$u_{12(ij)} = l_{ij} - \frac{l_{i+}}{2} - \frac{l_{+j}}{2} + \frac{l_{++}}{4} \quad (3.147)$$

şekillerinde ifade edilmektedir (Bishop, Fienberg ve Holland, 1975: 15-17).

DÖRDÜNCÜ BÖLÜM

ERZURUM VE KARS İLLERİNDE 2008-2012 YILLARINDA MEYDANA

GELEN KARAYOLU TRAFİK KAZALARINA ETKİ EDEN RİSK

FAKTÖRLERİNİN ÇOK DURUMLU LOGİT MODELİ İLE ANALİZİ

4.1. DÜNYA’DA KARAYOLU TRAFİK KAZALARI İSTATİSTİKLERİ

Karayolu trafik kazalarının neden olduğu yaralanmalar, günümüzde bireylerin maruz kaldığı kalıcı sakatlıkların ve ölümlerin en önemli sebeplerinden biri olarak görülmektedir. Dünya Sağlık Örgütü’nün 2013 yılında yayımladığı Karayolu Güvenliği Küresel Durum Raporu’na göre, dünyada 1 milyon 240 bin kişi karayolları trafik kazalarının bir sonucu olarak yaşamını yitirmekte ve 20 ila 50 milyon kişi ise ölümlü olmayan trafik kazalarına karışmaktadır. Küresel karayolu ölümlerine ait sayısal verilerin motorlu taşıt sayısındaki % 15’lik artış dikkate alınarak değerlendirilmesi gerekmektedir. Bunun yanında, gerekli önlemler alınmadığı takdirde 2020 yılına kadar ölümlü karayolu trafik kazası sayısının 1,9 milyona çıkması tahmin edilmektedir. Dünyada, karayolu trafik kazalarındaki yaralanmalar çoğunlukla 15-29 arası genç nüfusun ölümüne sebep olmakta ve söz konusu kazaların yaklaşık % 91’i düşük ve orta düzeyde gayrisafi yurtiçi hasılaya sahip ülkelerde meydana gelmektedir. Bahsi geçen ülkelerin dünya üzerindeki araç sayısının yaklaşık olarak yarısına sahip olduğu göz önüne alındığında karayolu trafik kazalarının düşük ve orta düzey gelire sahip ülkeler için ne ölçüde büyük bir risk olduğu daha iyi anlaşılabilmektedir. Öte yandan, karayolu trafik kazalarında yaşamını yitiren kişilerin büyük çoğunluğunun yaya, bisiklet ve motosiklet sürücüleri gibi “savunmasız karayolu kullanıcıları” olduğu görülmektedir. Buna rağmen, dünyada yalnızca 416 milyon kişiyi kapsayan 28 ülkenin karayolu trafik kazaları ile ilgili uygun yasalara sahip olduğu bilinmektedir.

Dünyada karayolu trafik kazalarında ölüm oranı, 100 bin kişide 18 kişi olarak tespit edilmiştir. Bununla birlikte, daha önce ifade edildiği gibi orta düzeyde gayri safi milli gelire sahip ülkeler, 100 bin kişide 20,1 ölüm oranı ile en yüksek ölüm oranına sahip ülkeler olarak dikkat çekmektedir. Sözü edilen oran, yüksek gelire sahip ülkelerde 100 bin kişide 8,7 kişi olmuştur. Karayolu trafik kazaları ölümlerinin % 80’i orta düzeyde gelire sahip ülkelere meydana gelmiştir. Bu ülkeler, toplam dünya nüfusunun % 72’sine tekabül etmektedir. Dikkat çekici bir diğer istatistik ise sözü edilen karayolu

trafik kazalarına karışan araçların yalnızca % 52'sinin trafik kaydının bulunmasıdır. Bu durum, orta düzeyde gelire sahip ülkelerin makineleşme düzeyleri ile karayolu trafik kazası ölümleri arasındaki uyumsuzluğu ortaya koymaktadır. Karayolu trafik kazası ölüm oranları dikkate alındığında bölgeler arasında da büyük eşitlikler göze çarpmaktadır. Özellikle Afrika Bölgesi, 100 bin kişide 24,1 ölüm oranı ile karayolu trafik kazası yaralanmaları sonucu ölüm riskinin en fazla olduğu bölge olmuştur. Bu oran Avrupa Bölgesi'nde 100 bin kişide 10,3 ölüm oranı ile en düşük değerindedir. Öte yandan, aynı bölgedeki ülkeler arasında da benzer eşitsizlikler tespit edilmiştir.

Dünyada meydana gelen karayolu trafik kazası ölümlerinin yarısını motosiklet sürücüleri (% 23), yayalar (% 22) ve bisiklet sürücüleri (% 5) oluşturmaktadır. Ölen kişilerin % 31'inin araç sürücüleri, % 19'unun ise diğer karayolu kullanıcılarının olduğu belirlenmiştir. Bahsi geçen yüzdeler bölgelere ve ülkelere göre değişiklik göstermektedir. Savunmasız karayolu kullanıcılarına ait en yüksek ölüm oranı % 57 ile düşük gelire sahip ülkelerde görülmekte, bu oran orta düzeyde gelire sahip ülkelerde % 51 ve yüksek gelire sahip ülkelerde % 39'dur. Karayolu trafik kazası ölümlerinin yaş ve cinsiyet gruplarına göre değerlendirilmesi yapıldığında, 15-44 yaş aralığındaki kişilerin dünyadaki karayolu trafik kazası ölümlerinin % 59'unu oluşturduğu görülmektedir. Bununla birlikte, karayolu trafik kazalarında ölen kişilerin üçte dördünden fazlası (% 77) erkek iken, bu oran en yüksek Batı Pasifik Bölgesi'nde ölçülmüştür.

Dünyadaki her bir ölümlü trafik kazasında, en az 20 kişinin ölümlü olmayan yaralanmalar yaşadığı tespit edilmiştir. Birçok ülkede, bahsi geçen yaralanmaların ciddiyet derecesine göre sınıflandırılması ve resmi kayıt altına alınması hususunda trafik polislerinin yeterli eğitime sahip olmadığı görülmektedir. Yaralanmaların ciddiyet derecesi hususunda farklı tanımlamalara gidilmesi ise bu sakatlıkların raporlanmasını daha da karmaşık hale getirmektedir. Ölümlü olmayan yaralanmalar ile ilgili bilgilerin değerlendirilmesi, yaralanma derecesine göre alması gereken sağlık hizmetinin belirlenmesi için de büyük önem arz etmektedir. Birçok ülke, bu bilgileri coğrafi bir örnekleme göre toplayan ve tüm ülkeye göre genelleştiren sistemler kullanmaktadır. Buna rağmen, yaralanmalı trafik kazalarının gerçek sayısının tespit edilmesi birçok ülke için halen önemli bir sorun olarak görülmektedir (Dünya Sağlık Örgütü, 2013: 4-8).

4.2. TÜRKİYE'DEKİ KARAYOLU TRAFİK KAZALARI İSTATİSTİKLERİ

Karayolları Trafik Kanunu (1983)'na göre "karayolu üzerinde hareket halinde olan bir ya da birden fazla aracın karıştığı ölüm, yaralanma ve zararlı sonuçlanmış olay" trafik kazası olarak tanımlanmaktadır. Trafik kazasının meydana gelmesi için kaza, genel trafiğe açık cadde, sokak ya da yolda meydana gelmeli; kazadan bir ya da birden fazla kişi ölmüş ya da yaralanmış olmalı; kazada hareket halinde en az bir taşıt bulunmalıdır (Türkiye İstatistik Kurumu, 2011). Türkiye İstatistik Kurumu'nun Haziran 2013'te yayımladığı Motorlu Kara Taşıtları Bülteni'ne göre Nisan ayı sonu itibariyle trafiğe kayıtlı araç sayısı 17.350.448'tir. Ocak-Nisan döneminde trafikteki toplam taşıt sayısı 317.035 adet artmıştır. Bu taşıtların % 51'ini otomobil, % 16,4'ünü kamyonet, % 15,5'ini kamyonet, % 8,8'ini traktör, % 4,3'ünü kamyon, % 2,3'ünü minibüs, % 1,4'ünü otobüs, % 0,2'sini ise özel amaçlı taşıtlar oluşturmaktadır. 2003-2013 yılları arasında Türkiye'deki motorlu kara taşıtları sayısı ise Tablo 4.1'de gösterilmiştir. Tablo 4.1.'de görüldüğü gibi son on yılda yaklaşık % 95'lik bir artış olmuştur. Karayolu trafiğinde en çok seyreden araç türleri sırasıyla otomobil, kamyonet, motosiklet ve traktör olmuştur.

Tablo 4. 1. Türkiye'deki Motorlu Kara Taşıtları Sayısı (2003-2013 Yılları Arası)

Yıllar	Toplam	Otomobil	Minibüs	Otobüs	Kamyonet	Kamyon	Motosiklet	Özel Amaçlı	Traktör
2003	8.903.843	4.700.343	245.394	123.500	973.457	579.010	1.073.415	24.468	1.184.256
2004	10.236.357	5.400.440	318.954	152.712	1.259.867	647.420	1.218.677	28.004	1.210.283
2005	11.145.826	5.772.745	338.539	163.390	1.475.057	676.929	1.441.066	30.333	1.247.767
2006	12.227.393	6.140.992	357.523	175.949	1.695.624	709.535	1.822.831	34.260	1.290.679
2007	13.022.945	6.472.156	372.601	189.128	1.890.459	729.202	2.003.492	38.573	1.327.334
2008	13.765.395	6.796.629	383.548	199.934	2.066.007	744.217	2.181.383	35.100	1.358.577
2009	14.316.700	7.093.964	384.053	201.033	2.204.951	727.302	2.303.261	34.104	1.368.032
2010	15.095.603	7.544.871	386.973	208.510	2.399.038	726.359	2.389.488	35.492	1.404.872
2011	16.089.528	8.113.111	389.435	219.906	2.611.104	728.458	2.527.190	34.116	1.466.208
2012	17.033.413	8.648.875	396.119	235.949	2.794.606	751.650	2.657.722	33.071	1.515.421
2013*	17.350.448	8.855.801	397.747	241.017	2.852.645	753.940	2.686.157	34.480	1.528.661

Kaynak: Türkiye İstatistik Kurumu Yıllara Göre Motorlu Kara Taşıtları Sayısı

*Nisan ayı sonu itibariyle

Tablo 4.2’de görüldüğü üzere Türkiye’de son on yılda yaklaşık 9 milyon karayolu trafik kazası meydana gelmiştir. Bu kazaların 1 milyondan fazlası ölümlü veya yaralanmalı, yaklaşık 8 milyonu ise maddi hasarlı kaza olarak kayıtlara geçmiştir. Bu kazalarda 43 bin beş yüzden fazla kişi yaşamını yitirirken, yaklaşık 2 milyon kişi yaralanmıştır.

Tablo 4. 2. Türkiye’de Meydana Gelen Karayolu Trafik Kazaları İstatistikleri (2003-2013 Yılları Arası)¹

Yıllar	Kaza Sayısı	Ölümlü, Yaralanmalı Kaza Sayısı	Maddi Hasarlı Kaza Sayısı	Ölü Sayısı	Yaralı Sayısı
2003	455.637	67.031	388.606	3.959	117.551
2004	537.352	77.008	460.344	4.427	136.437
2005	620.789	87.273	533.516	4.505	154.086
2006	728.755	96.128	632.627	4.633	169.080
2007	825.561	106.994	718.567	5.007	189.057
2008 ²	950.120	104.212	845.908	4.236	184.468
2009 ²	1.053.346	111.121	942.225	4.324	201.380
2010 ²	1.104.388	116.804	987.584	4.045	211.496
2011 ²	1.228.928	144.559	1.084.369	3.835	238.074
2012 ²	1.296.636	130.360 ³	1.166.276	3.750	268.102
2013 ^{2,4}	129.634	49.855	79.779	835	79.946
TOPLAM	8.931.146	1.091.345	7.839.801	43.556	1.949.677

Kaynak: Emniyet Genel Müdürlüğü Trafik Hizmetleri Başkanlığı Genel Kaza İstatistikleri ve Türkiye İstatistik Kurumu Karayolu Trafik Kazaları İstatistikleri 2011

¹Emniyet Genel Müdürlüğü ve Jandarma Genel Komutanlığı sorumluluk bölgesinde meydana gelen trafik kaza bilgilerini kapsamaktadır.

²1 Nisan 2008 tarihinde uygulamaya konulan tarafların anlaşarak kendi aralarında tutanak tanzim ettiği maddi hasarlı trafik kaza sayıları da dâhil edilmiştir.

³Emniyet Genel Müdürlüğü sorumluluk bölgesinde meydana gelen trafik kaza bilgilerini kapsamaktadır.

⁴Mayıs ayı sonu itibariyle

Tablo 4.3.’te karayolu trafik kazalarında taşıt ve nüfus başına düşen ortalama ölü ve yaralı sayısı özetlenmiştir. Buna göre 100,000 taşıta düşen ölü sayısı son yıllarda azalma eğilimi gösterse de yaralı sayısı artmaya devam etmektedir. Diğer taraftan, yine 100,000 nüfusa düşen ölü sayısı azalma eğilimindeyken, yaralı sayısı artmaktadır.

Tablo 4. 3. Karayolu Trafik Kazalarında Taşıt ve Nüfus Başına Düşen Ortalama Ölü ve Yaralı Sayısı (2002-2012 Yılları Arası)

Yıllar	100,000 taşıta düşen		100,000 nüfusa düşen	
	Ölü Sayısı	Yaralı Sayısı	Ölü Sayısı	Yaralı Sayısı
2002	47.3	1345.0	5.9	167.2
2003	44.3	1327.7	5.6	168.3
2004	43.2	1332.9	6.2	190.0
2005	40.4	1382.5	6.3	213.8
2006	37.9	1382.8	6.3	231.7
2007	38.4	1451.7	7.1	267.8
2008	30.8	1340.1	5.9	257.9
2009	30.2	1406.6	6.0	277.5
2010	26.8	1401.0	5.5	286.9
2011	23.8	1479.7	5.1	318.6
2012	22.0 ¹	1574.0 ¹	5.0 ²	354.5 ²

Kaynak: Türkiye İstatistik Kurumu Trafik Kazaları İstatistikleri 2011

¹İlgili yıla ait ölü, yaralı ve taşıt sayısından hareketle yazar tarafından hesaplanmıştır.

²31 Aralık 2012 tarihi itibarıyla Türkiye nüfusunun 75.627.384 kişi olduğu göz önünde tutularak yazar tarafından hesaplanmıştır.

Tablo 4.4.'te 2002-2012 yıllar arasında Türkiye nüfusunun karayolu trafik kazalarından etkilenme oranı sunulmuştur. Buna göre nüfusun ve kazadan etkilenenlerin trafik kazalarından etkilenme oranı yıllara göre artmaya devam etmektedir.

Tablo 4. 4. Nüfusun Karayolu Trafik Kazalarından Etkilenme Oranı (2002-2012 Yılları Arası)

Yıllar	Karayolu Trafik Kazasında Ölen ve Yaralanan Sayısı	Nüfus (Bin)	Karayolu Trafik Kaza Sayısı	Nüfusun Trafik Kazasından Etkilenme Oranı (‰)	Kazadan Etkilenenlerin Oranı (‰)
2002	120,505	69,626	65,748	1.73	1832.83
2003	122,160	70,231	67,031	1.74	1822.44
2004	140,864	71,794	77,008	1.96	1829.21
2005	158,591	72,065	87,273	2.20	1817.18
2006	173,713	72,974	96,128	2.38	1807.10
2007	194,064	70,586	106,994	2.75	1813.78
2008	188,704	71,517	104,212	2.64	1810.77
2009	205,704	72,561	111,121	2.83	1851.17
2010	215,541	73,723	116,804	2.92	1845.32
2011	241,909	74,724	131,845	3.24	1834.80
2012	272,852	75,627 ¹	130,360 ²	3.61 ³	2093.07 ³

¹31 Aralık 2012 itibariyle, Kaynak: Türkiye İstatistik Kurumu Adrese Dayalı Nüfu Kayıt Sistemi Sonuçları Haber Bülteni, 2013

²Emniyet Genel Müdürlüğü sorumluluk bölgesinde meydana gelen trafik kaza bilgilerini kapsamaktadır.

³İlgili yıla ait verilerden hareketle yazar tarafından hesaplanmıştır.

4.3.ERZURUM VE KARS İLLERİNDEKİ KARAYOLU TRAFİK AĞLARINA AİT GENEL BİLGİLER

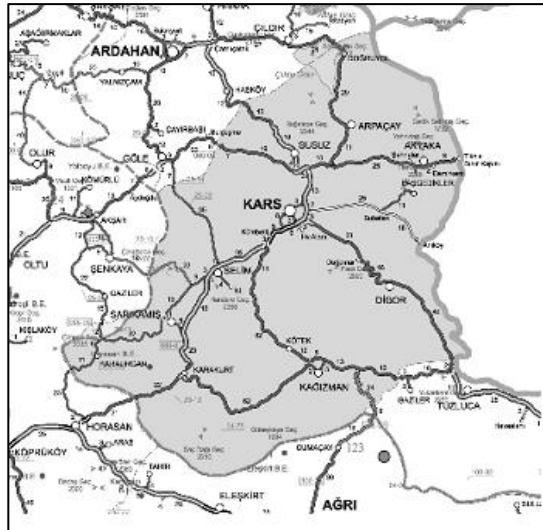
Erzurum ili, Karayolları Genel Müdürlüğü 12. Bölge Müdürlüğü'ne bağlı bulunmaktadır. Karayolları Genel Müdürlüğü verilerine göre, Karayolları 12. Bölge Müdürlüğü 38.000 km²'lik bir alanı kapsamakta ve 1514 km'lik devlet yolu ve 689 km'lik il yolu olmak üzere toplam 2203 km'lik bir karayolu ağında hizmet vermektedir. Erzurum iline ait 1061 km devlet yolu, 475 km il yolu bulunmaktadır. Erzurum ilindeki devlet yollarının 222 km'si, il yollarının 27 km'si asfalt kaplama yollardır. Karayolları 12. Bölge Müdürlüğü'nde trafiğe kayıtlı araç sayısı 109.419'dur. Şekil 4.1.'de Erzurum iline ait karayolu haritası sunulmuştur.



Kaynak: Karayolları Genel Müdürlüğü

Şekil 4. 1. Erzurum İline ait Karayolu Ağı Haritası

Kars ili, Karayolları Genel Müdürlüğü 18. Bölge Müdürlüğü'ne bağlı bulunmaktadır. Karayolları Genel Müdürlüğü verilerine göre, Karayolları 18. Bölge Müdürlüğü yaklaşık 19.000 km²'lik bir alanı kapsamakta ve 1014 km'lik devlet yolu ve 310 km'lik il yolu olmak üzere toplam 1324 km'lik bir karayolu ağında hizmet vermektedir. Kars iline ait 499 km devlet yolu, 176 km il yolu bulunmaktadır. Kars ilindeki devlet yollarının 14 km'si, il yollarının 1 km'si asfalt kaplama yollardır. Şekil 4.1.'de Kars iline ait karayolu haritası sunulmuştur.



Kaynak: Karayolları Genel Müdürlüğü

Şekil 4. 2. Kars İline ait Karayolu Ağı Haritası

4.4. ERZURUM VE KARS İLLERİNDE MEYDANA GELEN KARAYOLU TRAFİK KAZALARI İSTATİSTİKLERİ (2007-2011)

Tablo 4.5’te görüldüğü gibi Erzurum ilinde 2007-2011 yılları arasında meydana gelen karayolu trafik kazası sayısı 34,018’dir. Sözü edilen dönemdeki ölümlü kazaların sayısı 98, yaralanmalı kazaların sayısı 3703, maddi hasarlı kazaların sayısı 30,177’dir. Bu kazalarda 135 kişi yaşamını yitirmiş, 7982 kişi yaralanmıştır.

Tablo 4. 5. Erzurum İlinde 2007-2011 Yılları Arasında Meydana Gelen Karayolu Trafik Kaza Sayıları

Yıllar	Toplam Kaza	Ölümlü Kaza	Yaralanmalı Kaza	Maddi Hasarlı Kaza	Ölü Sayısı	Yaralı Sayısı
2007	4848	24	628	4196	31	1352
2008	5761	11	594	5156	13	1223
2009	6901	19	716	6166	28	1616
2010	7675	26	829	6780	35	1800
2011	8833	18	936	7879	28	1991
TOPLAM	34,018	98	3703	30,177	135	7982

Kaynak: Türkiye İstatistik Kurumu Veri Tabanı

Tablo 4.6’da görüldüğü gibi Kars ilinde 2007-2011 yılları arasında 5426 karayolu trafik kazası meydana gelmiştir. Sözü edilen dönemdeki ölümlü kazaların sayısı 29, yaralanmalı kazaların sayısı 1043, maddi hasarlı kazaların sayısı 4354’tür. Bu kazalarda 33 kişi yaşamını yitirmiş, 2536 kişi yaralanmıştır.

Tablo 4. 6. Kars İlinde 2007-2011 Yılları Arasında Meydana Gelen Karayolu Trafik Kaza Sayıları

Yıllar	Toplam Kaza	Ölümlü Kaza	Yaralanmalı Kaza	Maddi Hasarlı Kaza	Ölü Sayısı	Yaralı Sayısı
2007	849	8	183	658	9	490
2008	963	3	166	794	3	358
2009	1057	6	182	869	7	375
2010	1080	6	229	845	6	600
2011	1477	6	283	1188	8	723
TOPLAM	5426	29	1043	4354	33	2536

Kaynak: Türkiye İstatistik Kurumu Veri Tabanı

4.5. KARAYOLU TRAFİK KAZALARINA AİT VERİLERİN KATEGORİK MODELLER İLE ANALİZİ

4.5.1. Literatür Taraması

Chipman (1996), 1990 yılında Ontario Sağlık Araştırması yapılan bir anket çalışmasından elde edilen verileri kullanarak karayolları trafik kazalarını etkileyen risk faktörlerini belirlemeye çalışmış ve araştırma sonuçlarına göre ilaç kullanımın, alkol alma düzeyinin ve kronik hastalıkların yaralanmalara sebep olan risk faktörleri olduğunu ortaya koymuştur.

Shankar, Mannering ve Barfield (1996), Washington şehrinde meydana gelen 5 yıllık kaza verilerini maddi hasarlı, olası yaralanmalı, yaralanmalı ve sakatlanmaya neden olan yaralanma ya da ölümlü olmak üzere dört farklı kaza türü ile kategorilere ayırarak yuvalanmış logit modeli ile tahmin etmişler ve sonuç olarak çevresel koşulların, yol yüzeyinin, kaza türünün, sürücü ve araç özelliklerinin kaza sonucuna etki ettiklerini ortaya koymuşlardır.

Zhang vd. (1998), yaş gruplarına göre Kanadalı sürücülerin ölümlü karayolu trafik kazalarına karışma durumunu etkileyen faktörleri bahis oranları yardımıyla inceledikleri çalışmalarında genç sürücülerin daha çok risk ve alkol alma, hız yapma, emniyet kemeri kullanmama, yorgunken araç kullanma eğiliminde olduklarını ortaya koymuşlardır.

Al-Ghamdi (2002), Riyad şehrinde 1997-1998 yılları arasında meydana gelen karayolu trafik kazasına sebep olan faktörleri lojistik regresyon modeli ile tahmin etmiş ve analiz sonuçlarına göre kaza yeri ile sürücü kusurlarının kazaları etkileyen risk faktörleri olduğunu tespit etmiştir.

Jones ve Jorgensen (2003), Norveç'te 1985-1996 yılları arasında meydana gelen 16000 ölümlü ve ağır yaralanmalı kazaları analiz ettikleri çalışmalarında çok düzeyli modelleri ve lojistik regresyon modelini kullanarak ölümlü kazaları etkileyen risk faktörlerinin yaş, cinsiyet, kazaya karışan araç türü ve yol yüzeyindeki özellikleri olduğunu ortaya koymuşlardır.

Yau (2004), Hong Kong'da 1999-2000 yılları arasında meydana gelen karayolu trafik kazalarını analiz ettiği çalışmasında; risk faktörlerini insan, araç, güvenlik ve

 evresel fakt rlerin kazalar  zerindeki etkisini analiz etmiř ve s z  edilen gruplara ait farklı risk fakt rlerini ortaya koymuřtur.

Islam ve Mannering (2006) s r c  yařının tek ara lı kazalara etkisini  ok durumlu logit modeli ile analiz ettikleri  alıřmalarında yolcularla trafięe  ıkma ve ara  yařına g re s r c n n yařının kazalara etkisi ortaya konulmuřtur.

Sze ve Wong (2007), Hong Kong’da 1991-2004 yılları arasında meydana gelen ve yayaların d hil olduęu 73746 karayolu trafik kazasını iki durumlu lojistik regresyon modeli ile tahmin ederek ve demografik,  evresel ve dięer risk fakt rlerini kontrol grubu olarak deęerlendirerek yayaların yaralanmalı y zdesinde belirtilen yıllar arasında azalma olduęunu ortaya koymuřlardır.

Awadzi vd. (2008), gen  (35-54 yař arası) ve yařlı (65 yař ve  st ) s r c ler i in karayolu trafik kazası sonu larının arařtırıldıęı  alıřmalarında  ok durumlu logit modelini kullanmıřlar ve elde edilen arařtırma sonu larına g re saat, tek bir s r c n n olması ve g nd z ara  kullanılması gibi fakt rlerin kazalarda etkili olduęunu tespit etmiřlerdir.

Milton, Shankar ve Mannering (2008), Washington řehrinde 1990-1994 yılları arasında meydana gelen 22568 karayolu trafik kazasını karıřık logit modeli ile analiz ederek ortalama trafik yoęunluęunun, kamyonlara ait g nl k trafik yoęunluęunun, trafikte seyreden kamyonlarının y zdesinin ve kar yaęıřının s z  edilen kazaları etkileyen risk fakt rleri olduęunu tespit etmiřlerdir.

Eluru, Bhat ve Hensher (2008), Birleřik Devletler’de 2004 yılında meydana gelen motorlu olmayan yaralanmalı kazaların karıřık genelleřtirilmiř sıralı yanıt deęiřkeni modeli kullanılarak analiz edildięi  alıřmalarında kazaya karıřan kiřilerin yařlarının, karayolundaki hız limitinin, kaza yerlerinin ve kazanın meydana geldięi saatin belirtilen  zellięe sahip kazaları etkileyen fakt rler olduęunu ortaya koymuřlardır.

Tay, Rifaat ve Chin (2008), Singapur’da 1992-2002 yılları arasında meydana gelen kazaya karıřan aracın  arpıp ka tıęı kazaları iki durumlu lojistik regresyon model ile analiz ettikleri  alıřmalarında s r c lerin gece, k pr   zerindeyken, d z yoldayken ve alıřveriř merkezlerinin yanındayken, iki ara lı kaza olduęunda, iki tekerlekli ara lı olduęunda, s r c  erkek olduęunda ve 45-69 yařları arasında olduklarında  arpıp ka ma olasılıklarının daha y ksek olduęu g r lm řt r.

Kartal, Kutlar ve Beğen (2011), Sivas, Kayseri ve Yozgat illerinde 2000-2006 yıllarında meydana gelen karayolu trafik kazalarını etkileyen risk faktörlerini iki durumlu lojistik regresyon yöntemi ile analiz etmişler ve ölümlü kazalarda karayolu kullanıcıların daha fazla yaşamını yitirdiğini, bölünmemiş yollarda ölümlü trafik kazaların arttığını ortaya koymuşlardır.

Drucker vd. (2013), sivil sürücülerin cankurtaran araçları ile karışıklıkları kazaları etkileyen faktörleri çok değişkenli logit modeli kullanarak araştırdıkları çalışmalarında, kazanın meydana geldiği yolda görüşe engel cismin bulunmasının, kazanın gece olmasının, zıt yönden gelen araçların çarpışmasının yaralanmalı kazaların olma olasılığına etki ettiğini ortaya koymuşlardır.

Elvik, Fridstrøm, Kaminska ve Meyer (2013), çivili araç lastiklerinin Norveç'teki trafik kazaları üzerindeki etkisini 1991-2000 yılları ve 2002-2009 yılları arasında gerçekleştirilen iki araştırmadan elde edilen verilerle negatif binom regresyon yardımıyla analiz etmişler ve yüzdesel değişimleri ortaya koymuşlardır.

Mamčič ve Sivilevičius (2013) Litvanya'daki bölgesel çakıl yollarda meydana gelen trafik kazalarının istatistiksel analizini yaptıkları çalışmalarında, mevsim değişkeninin önemli bir faktör olduğunu ve kazalarının ilkbahardan sonbahara doğru artma eğilimi gösterdiğini tespit etmişlerdir.

4.5.2. Uygulama

4.5.2.1. Araştırmanın Amacı, Kapsamı ve Önemi

Bu tez çalışmasının temel amacı, karayolu trafik kazalarına etki eden risk faktörlerinin istatistiksel analizinin yapılabilmesidir. Bu bağlamda, çalışma kapsamında yapılan araştırma, Erzurum ve Kars illerinde 2008-2012 yılları arasında meydana gelen trafik kazalarını etkileyen risk faktörlerini belirleyebilme amacıyla gerçekleştirilmiştir. Bilindiği kadarı ile Erzurum ve Kars illerinde bu amaçla yapılan benzer bir çalışma olmadığından karayolu trafik kazalarını etkileyen risk faktörlerinin belirlenmesi vasıtasıyla söz konusu illerde ileride meydana gelebilecek trafik kazalarının olası olumsuz sonuçlarının azaltılması ve elde edilen araştırma bulguları ışığında karayolu trafik güvenliği adına daha etkin politikaların uygulanabilmesi mümkün olabilecektir.

4.5.2.2. Araştırmanın Yöntemi

Bu tez çalışması kapsamında gerçekleştirilen araştırmada bağımlı değişken olarak karayolu trafik kazasının türü seçilmiştir. Bağımlı değişkene ait ölümlü, yaralanmalı ve maddi hasarlı kazalar olmak üzere ikiden fazla durumlu bir sınıflandırma söz konusu olduğundan, bu sınıflandırma herhangi bir doğal sıralamaya tabi olmadığından ve araştırmada kategorik veriler kullanıldığından araştırma yöntemi olarak çok durumlu logit modelinin uygulanacak en uygun yöntem olduğu düşünülmüştür. Stata 11.2 bilgisayar paket programı kullanılarak elde edilen bulgular sonucu kurulan çok durumlu modelin ilişkisiz alternatiflerin bağımsızlığı varsayımını sağlayıp sağlayamadığı, uyum iyiliği, çoklu bağlantı durumu, model katsayılarının ve risk faktörlerine ait bahis oranlarının anlamlılığı ilgili yöntemler yardımıyla test edilmiştir.

4.5.2.3. Araştırmanın Veri Kümesi ve Kısıtları

Bu tez çalışması kapsamında gerçekleştirilen araştırmanın veri kümesi, Erzurum ve Kars illerindeki Trafik Tescil ve Denetleme ve Bölge Trafik Denetleme Şube Müdürlükleri'nin istatistik birimlerinde kaydı tutulan ölümlü, yaralanmalı ve maddi hasarlı karayolu trafik kaza tespit tutanaklarından elde edilen 9162 kaza raporunu kapsamaktadır. Araştırmanın örnekleminin oluşturulmasında tabakalı örneklem yöntemi kullanılmıştır. Sözü edilen kaza raporları 2008-2012 yıllarını kapsamaktadır. Elde edilen veriler bilgisayar ortamına aktarılarak yeniden düzenlenmiş ve birtakım faktörler veri eksikliği nedeniyle araştırma dışında tutulmuştur.

Araştırmanın veri kümesinin temin edilmesi aşamasında görevli polis memurları ile yapılan birebir görüşmelerde araç sürücüsünün emniyet kemeri takıp takmadığına ait tespitin yapılamadığı ve ilgili trafik kazası tespit tutanağında “belirsiz” olarak kayıt altına alındığı belirtilmiştir. Bu sebeple araç sürücüsüne ait emniyet kemeri faktörü araştırma dışında tutulmuştur. Öte yandan, araç sürücülerinin çok büyük bölümü erkek ve Türkiye Cumhuriyeti vatandaşı olduğundan cinsiyet ve uyruk değişkenleri de araştırmaya dâhil edilmemiştir. Ek olarak, bulgular ve değerlendirme bölümünde daha detaylı açıklanacağı üzere frekans değeri düşük ve birbiri ile yakın ilişkili olan kategorilerin birleştirilmesi yoluna gidilmiştir. Trafik tespit tutanaklarında kayıt altına alınan demografik değişkenlerin değerlendirilmesinde yalnızca araç sürücüsüne ait

veriler kullanılmış olup diğer karayolu kullanıcılarına ait verilerde eksiklik olduğundan bütünlüğün sağlanması açısından araştırmaya dâhil edilmemiştir.

4.5.2.4. Bulgular ve Değerlendirme

4.5.2.4.1. Değişkenlerin Tanımlanması ve Tanımlayıcı İstatistikler

Çoklu bağlantı sorunu çözüldükten ve ilişkisiz alternatiflerin bağımsızlığı varsayımı sağlandıktan sonra araştırma kapsamına alınan bağımsız değişkenlere ait tanımlamalar, tanımlayıcı istatistikler ve alt kategoriler Tablo 4.7’de gösterilmiştir. Araştırmanın başlangıç aşamasında modele dâhil edilmek istenen araç sürücüsünün alkol durumu, kazanın haftanın hangi günü olduğu, trafik işaretleri ve yol durumu ile ilgili daha farklı değişkenler çoklu bağlantı sorununa sebep olduklarından, ilişkisiz alternatiflerin bağımsızlığı varsayımını ihlal ettiklerinden ya da modelde bu değişkenlerin bazılarını kapsayan alternatif değişkenler olduğundan model dışında tutulmuştur. Bağımlı değişkene ve bağımsız değişkenlere ait kategorilerin belirlenmesi aşamasında literatürdeki benzer çalışmalar (Abdel-Aty vd., 1998; Al-Ghamdi, 2002; Jones ve Jorgensen, 2003; Yau, 2004; Sze ve Wong, 2007; Depaire vd., 2008; Tay vd., 2008; De Boni vd., 2013) referans olarak alınmıştır.

Tablo 4. 7. Modele Dâhil Edilen Bağımsız Değişkenlere ait Tanımlamalar

Değişken	Tanımı	Ortalama	Standart Sapma
Kaza Türü (kazatür)	Bağımlı Değişken (Ölümlü kaza = 1; Yaralanmalı kaza = 2; *Maddi hasarlı kaza = 3)	2.667	0.485

(1) Mevsim (mevsim)	Karayolu trafik kazasının meydana geldiği mevsim Kış = 1; Diğer = 0 İlkbahar = 1; Diğer = 0 Yaz = 1; Diğer = 0 *Sonbahar = 1; Diğer = 0	0.286 0.221 0.240 0.253	0.452 0.415 0.427 0.435
(2) Saat (saat)	Karayolu trafik kazasının meydana geldiği saat *00:00-05:59 = 1; Diğer = 0 06:00-11:59 = 1; Diğer = 0 12:00-17:59 = 1; Diğer = 0 18:00-23:59 = 1; Diğer = 0	0.055 0.270 0.425 0.250	0.228 0.444 0.494 0.433
(3)Haftaiçi/Haftasonu (hafta)	Karayolu trafik kazasının meydana geldiği gün Hafta içi = 1; Hafta sonu = 0	0.743	0.437
(4) Hava durumu (hava)	Karayolu trafik kazasının meydana geldiği sıradaki hava durumu Açık = 1; Diğer = 0 Yağmurlu / Bulutlu = 1; Diğer = 0 *Sisli / Karlı / Fırtınalı / Tipili = 1; Diğer = 0	0.684 0.244 0.072	0.465 0.430 0.258
(5) Gün durumu (gün)	Karayolu trafik kazasının meydana geldiği sıradaki gün durumu Gündüz = 1; Gece = 0	0.664	0.472
(6) Yaya kaldırımı (yayakal)	Karayolu trafik kazasının meydana geldiği yolda yaya kaldırımının olup olmadığı Var =1; Yok = 0	0.623	0.485
(7) Yolun yüzeyi(yyüzey)	Karayolu trafik kazasının meydana geldiği yolun yüzeyi Kuru/Tozlu = 1; Diğer = 0 Islak/Çamurlu/Su ya da yağ birikintili = 1; Diğer = 0 *Karlı/Buzlu = 1; Diğer = 0	0.610 0.218 0.172	0.488 0.413 0.377
(8) Aydınlatma (aydınlatma)	Karayolu trafik kazasının meydana geldiği yolda aydınlatma olup olmadığı Var =1; Yok = 0	0.564	0.496
(9) Araç sayısı (asay)	Karayolu trafik kazasına karışan araç sayısı Tek araçlı kaza = 1; Diğer = 0 *İki araçlı kaza = 1; Diğer = 0 Zincirleme kaza = 1; Diğer = 0	0.410 0.537 0.054	0.492 0.499 0.226
(10) Araç cinsi (acinsi)	Karayolu trafik kazasına karışan araç cinsi *Otomobil =1; Diğer = 0 Kamyonet /Kamyon = 1; Diğer = 0 Otobüs/ Minibüs = 1; Arazi taşıtı/Traktör/Çekici, vb = 1; Diğer = 0	0.593 0.269 0.087 0.025	0.491 0.444 0.282 0.156
(11) Yaş grubu (yas)	Karayolu trafik kazasına karışan araç sürücüsünün yaşı 16 yaşından küçük = 1; Diğer = 0 16-24 yaş arası = 1; Diğer = 0 25-34 yaş arası = 1; Diğer = 0	0.045 0.115 0.400	0.207 0.319 0.490

	35-44 yaş arası = 1; Diğer = 0	0.240	0.427
	45-54 yaş arası = 1; Diğer = 0	0.122	0.327
	*55 yaş ve üstü = 1; Diğer = 0	0.078	0.269
(12) Öğrenim durumu (ogrenim)	Karayolu trafik kazasına karışan araç sürücüsünün öğrenim durumu		
	İlköğretim = 1; Diğer = 0	0.428	0.495
	Ortaöğretim = 1; Diğer = 0	0.348	0.476
	*Yükseköğretim = 1; Diğer = 0	0.222	0.416
(13) Sürücü belgesi (ehliyet)	Karayolu trafik kazasına karışan araç sürücüsüne ait sürücü belgesinin sınıfı		
	E sınıfı = 1; Diğer = 0	0.592	0.491
	B sınıfı = 1; Diğer = 0	0.328	0.470
	*Diğer sınıflar (örneğin; A1, A2, C, D) = 1; Diğer = 0	0.079	0.270
(14) Kusur (kusur)	Karayolu trafik kazasına karışan araç sürücüsü ve diğer karayolu kullanıcılarına ait kusurlar		
	Hız = 1; Diğer = 0	0.387	0.487
	Geçiş üstünlüğüne riayet etmemek = 1; Diğer = 0	0.396	0.489
	*Arkadan çarpma = 1; Diğer = 0	0.112	0.316
	Diğer kusurlar (örneğin; dikkatsizlik, park ihlali, araç takip mesafesi ihlali) = 1; Diğer = 0	0.098	0.298
*Referans kategori			

Kurulacak modele dâhil edilen bütün değişkenlere ait kategorilerin etkilerini görebilmek adına bütün bağımsız değişkenler gölge değişken olarak tanımlanmıştır. Dünya Sağlık örgütünün karayolu trafik kazaları ve trafik güvenliği ile ilgili birçok raporunda daha önce ifade edildiği gibi karayolu kullanıcılarının yaş faktörü gruplara ayrılarak değerlendirilmektedir. Bu sebeple, sürekli değişken olarak modele dâhil edilebilecek yaş değişkeni, yaş grupları arasındaki etkiyi daha açık görebilme adına kategorik olarak değerlendirilmiştir. Ayrıca, yine Dünya Sağlık Örgütü'nün ilgili raporlarında dikkat çekildiği üzere karayolu trafik kazalarından en çok etkilenen kişiler, savunmasız karayolu kullanıcıları olarak adlandırılan motosiklet, bisiklet kullanıcıları ve yayalardır. Bu sebeple, yaya kaldırımının olup olmamasının kaza türü ile ilişkisini görebilmek adına yaya kaldırımı değişkeni özellikle modele dâhil edilmiştir. Analizin başlangıç aşamasında modele dâhil edilmesi planlanan kaza yeri, yol durumu ve yıl değişkenleri sonraki bölümde açıklandığı üzere çok yüksek göreceli risk oranlarına sahip olduğundan model dışında tutulmuştur.

4.5.2.4.2. Çok Durumlu Logit Modelinin Kurulması

Bağımsız değişken ve bağımlı değişkenler belirlendikten sonra, maddi hasarlı kaza türü referans kategorisi kabul edilerek, $\bar{O}$ ölümlü, Y yaralanmalı ve M maddi hasarlı kaza türlerini temsil etmek üzere tahmin edilecek genel çok durumlu logit modelleri aşağıdaki gibi yazılmaktadır:

$$\begin{aligned} \ln \Omega_{\bar{O}|M}(x_i) = & \beta_{0,\bar{O}|M} + \beta_{1,\bar{O}|M}D_{mevsim1} + \beta_{2,\bar{O}|M}D_{mevsim2} + \beta_{3,\bar{O}|M}D_{mevsim3} + \beta_{4,\bar{O}|M}D_{saat2} \\ & + \beta_{5,\bar{O}|M}D_{saat3} + \beta_{6,\bar{O}|M}D_{saat4} + \beta_{7,\bar{O}|M}D_{hafta1} + \beta_{8,\bar{O}|M}D_{hava1} \\ & + \beta_{9,\bar{O}|M}D_{hava2} + \beta_{10,\bar{O}|M}D_{gün1} + \beta_{11,\bar{O}|M}D_{yayakal} + \beta_{12,\bar{O}|M}D_{yyuzey1} \\ & + \beta_{13,\bar{O}|M}D_{yyuzey2} + \beta_{14,\bar{O}|M}D_{aydınlatma} + \beta_{15,\bar{O}|M}D_{asay1} + \beta_{16,\bar{O}|M}D_{asay3} \\ & + \beta_{17,\bar{O}|M}D_{acinsi2} + \beta_{18,\bar{O}|M}D_{acinsi3} + \beta_{19,\bar{O}|M}D_{acinsi4} + \beta_{20,\bar{O}|M}D_{yas1} \\ & + \beta_{21,\bar{O}|M}D_{yas2} + \beta_{22,\bar{O}|M}D_{yas3} + \beta_{23,\bar{O}|M}D_{yas4} + \beta_{24,\bar{O}|M}D_{yas5} \\ & + \beta_{25,\bar{O}|M}D_{ogrenim1} + \beta_{26,\bar{O}|M}D_{ogrenim2} + \beta_{27,\bar{O}|M}D_{ehliyet1} + \beta_{28,\bar{O}|M}D_{ehliyet2} \\ & + \beta_{29,\bar{O}|M}D_{kusur1} + \beta_{30,\bar{O}|M}D_{kusur2} + \beta_{31,\bar{O}|M}D_{kusur4} \end{aligned}$$

$$\begin{aligned} \ln \Omega_{Y|M}(x_i) = & \beta_{0,Y|M} + \beta_{1,Y|M}D_{mevsim1} + \beta_{2,Y|M}D_{mevsim2} + \beta_{3,Y|M}D_{mevsim3} + \beta_{4,Y|M}D_{saat2} \\ & + \beta_{5,Y|M}D_{saat3} + \beta_{6,Y|M}D_{saat4} + \beta_{7,Y|M}D_{hafta1} + \beta_{8,Y|M}D_{hava1} \\ & + \beta_{9,Y|M}D_{hava2} + \beta_{10,Y|M}D_{gün1} + \beta_{11,Y|M}D_{yayakal} + \beta_{12,Y|M}D_{yyuzey1} \\ & + \beta_{13,Y|M}D_{yyuzey2} + \beta_{14,Y|M}D_{aydınlatma} + \beta_{15,Y|M}D_{asay1} + \beta_{16,Y|M}D_{asay3} \\ & + \beta_{17,Y|M}D_{acinsi2} + \beta_{18,Y|M}D_{acinsi3} + \beta_{19,Y|M}D_{acinsi4} + \beta_{20,Y|M}D_{yas1} \\ & + \beta_{21,Y|M}D_{yas2} + \beta_{22,Y|M}D_{yas3} + \beta_{23,Y|M}D_{yas4} + \beta_{24,Y|M}D_{yas5} \\ & + \beta_{25,Y|M}D_{ogrenim1} + \beta_{26,Y|M}D_{ogrenim2} + \beta_{27,Y|M}D_{ehliyet1} + \beta_{28,Y|M}D_{ehliyet2} \\ & + \beta_{29,Y|M}D_{kusur1} + \beta_{30,Y|M}D_{kusur2} + \beta_{31,Y|M}D_{kusur4} \end{aligned}$$

4.5.2.4.3. Kurulan Modelin İlişkisiz Alternatiflerin Bağımsızlığı Varsayımını Sağlayıp Sağlamadığının Testi

Bir önceki bölümde kurulan modelin ilişkisiz alternatiflerin bağımsızlığı varsayımını sağlayıp sağlamadığını test etmek adına Stata 11.2. paket programında *mlogit*, *hausman base* komutu çalıştırılmıştır. İlişkisiz alternatiflerin bağımsızlığı varsayımının testinde kullanılan sıfır hipotezi,

$H_0: J$ ve K sonuç değişkenlerine ait bahis oranları diğer alternatiflerden bağımsızdır.

şeklinde. Elde edilen sonuçlar Tablo 4.8’de gösterilmiştir.

Tablo 4. 8. İlişkisiz Alternatiflerin Bağımsızlığı Varsayımı Testi

Bağımlı Değişken	χ^2	Serbestlik Derecesi	$P > \chi^2$	Sonuç
Ölümlü kaza	-0.901	32	1.000	H_0 kabul
Yaralanmalı kaza	8.974	32	1.000	H_0 kabul
Maddi hasarlı kaza	-229.631	32	1.000	H_0 kabul

Tabloda elde edilen sonuçlara göre, karayolu trafik kazalarını etkileyen risk faktörlerinin belirlenmesi adına kurulan çok durumlu logit modelinin ilişkisiz alternatiflerin bağımsızlığı varsayımını sağladığını göstermektedir. Hausman ve McFadden (1984: 1226)’ın ortaya koyduğu gibi ilişkisiz alternatiflerin bağımsızlığı varsayımı testinde χ^2 değerleri, sonlu örnek uygulamalardaki pozitif yarı belirlilik eksikliği sebebiyle sıklıkla negatif değer almaktadır.

4.5.2.4.4. Kurulan Modeldeki Değişkenlere Ait Çoklu Bağlantı Testi

Araştırma kapsamında kurulan model, ilişkisiz alternatiflerin bağımsızlığı varsayımını sağladıktan sonra, modele ait değişkenler çoklu bağlantı testine tabi tutulmuştur. Modelde çoklu bağlantı sorunu olup olmadığını test edebilmek adına Stata 11.2 paket programında *vif* komutu çalıştırılarak bağımsız değişkenlerin varyans şişkinlik faktörlerine (variance inflation factor) bakılmıştır. Bu testte, varyans şişkinlik faktör değerleri 10 ve üzerinde olan değişkenlerin çoklu bağlantıya sebep olduğu düşünülmektedir. Modeldeki risk değişkenlerine ait varyans şişkinlik faktörü değerleri Tablo 4.9.’da sunulmuştur.

Tablo 4. 9. Modeldeki Bağımsız Değişkenlere ait Varyans Şişkinlik Faktörleri

Bağımsız Değişken	VIF	1/VIF
saat3	7.12	0.141
saat2	6.47	0.155
hava1	4.27	0.234
saat4	4.25	0.235
yas3	3.95	0.253
hava2	3.81	0.262
ehliyet2	3.67	0.273
ehliyet1	3.65	0.274
kusur1	3.24	0.309
yolyüzeyi1	3.24	0.309
yas4	3.24	0.309
gündüz1	2.84	0.352
kusur2	2.71	0.369
yolyüzeyi2	2.38	0.421
yas2	2.35	0.426
yas5	2.30	0.434
ogrenim1	2.16	0.464
mevsim1	2.11	0.473
yas1	1.89	0.529
ogrenim2	1.84	0.544
kusur4	1.70	0.587
mevsim3	1.53	0.654
mevsim2	1.51	0.662
asay1	1.47	0.680
aydınlatma	1.17	0.851
acinsi3	1.17	0.854
yayakal	1.15	0.868
acinsi2	1.15	0.873
asay3	1.06	0.942
acinsi4	1.05	0.954
hafta1	1.01	0.988
Ortalama VIF	2.63	

Tablo 4.9.'da görüldüğü gibi modele dâhil edilen hiçbir bağımsız değişken 10 ya da daha fazla varyans şişkinlik faktörüne sahip değildir. Buna göre, modeldeki değişkenler arasında çoklu bağlantı sorununa sebep olan herhangi bir değişken bulunmamaktadır.

4.5.2.4.5. Kurulan Modelin Uyum İyiliğinin Test Edilmesi

Araştırmada kullanılan modelin uyum iyiliğinin test etmek için Stata 11.2 paket programında *fitstat* komutu çalıştırılmıştır. Tablo 4.10.'da modelin uygunluk değerleri sunulmuştur.

Tablo 4. 10. Kurulan Modelin Uygunluk Değerleri

Uygunluk Kriteri	Değer
Log-olabilirlik oranı (eksen)	-6079.365
Log-olabilirlik değeri (tüm model)	-5035.374
Sapma(9066)	10070.748
LR(62)	2087.981
$P > LR$	0.000
McFadden'ın R^2 değeri	0.172
Düzeltilmiş McFadden R^2	0.156
Akaike bilgi kriteri (AIC)	1.120
Bayes bilgi kriteri (BIC)	-72636.736
BIC'	-1522.367

Tablo 4.10.'da görüldüğü üzere Akaike bilgi kriterinin düşük değerde olması ve Bayes bilgi değerinin aldığı çok yüksek negatif değer modelin uygun olduğunu göstermektedir. Bununla birlikte, *mlogitgof* komutu çalıştırılarak kurulan model Pearson χ^2 istatistiğine göre uyum iyiliği testine tabi tutulduğunda elde edilen sonuçlar Tablo 4.11'de gösterilmiştir.

Tablo 4. 11. Kurulan Modelin Uyum İyiliği Testi

Çok Durumlu Logit Modelin Uyum İyiliği Testi	
Bağımlı değişken	Kaza türü
Gözlem sayısı	9162
Referans değeri	Maddi hasarlı kaza
Grup sayısı	10
Pearson χ^2 test istatistiği	10.844
Serbestlik derecesi	16
$P > \chi^2$	0.819

Tabloda görüldüğü üzere, Pearson χ^2 istatistiğine ait olasılık değeri $p > 0.05$ olduğundan kurulan modelin uyumunun iyi olduğu söylenebilmektedir.

4.5.2.4.6. Bağımsız Değişkenler için Olabilirlik Oranı Testi

Kurulan modeldeki bağımsız değişkenler için olabilirlik-oranı testi yapabilmek için Stata 11.2 paket programında *mlogtest*, *all* komutu çalıştırılmıştır. Bağımsız değişkenlere ait olabilirlik oranı testi için kurulan sıfır hipotezi,

H_0 : Verilen bağımsız değişken(ler)e ait bütün katsayılar sıfırdır.

şeklindedir. Modeldeki bağımsız değişkenler için olabilirlik-oranı testinden elde edilen sonuçlar Tablo 4.12’de sunulmuştur. Buna göre yıldız(lar)la işaretlenmiş bağımsız değişkenler için H_0 hipotezinin reddedileceğini ve kurulan model içerisinde bağımlı değişken üzerindeki etkilerinin istatistiksel olarak anlamlı olduğunu söylemek mümkündür.

Tablo 4. 12. Bağımsız Değişkenler için Olabilirlik Oranı Testi

Bağımsız Değişken	χ^2	Serbestlik Derecesi	$P > \chi^2$
Mevsim			
Kış	65.863	2	0.000*
İlkbahar	13.937	2	0.001*
Yaz	1.788	2	0.409
Saat			
06:00-11:59	0.576	2	0.750
12:00-17:59	2.648	2	0.266
18:00-23:59	1.772	2	0.412
Haftaiçi/haftasonu			
Haftaiçi	0.897	2	0.638
Hava			
Açık	10.418	2	0.005*
Yağmurlu/Bulutlu	4.497	2	0.106
Gün durumu			
Gündüz	1.012	2	0.603
Yaya kaldırımı	87.190	2	0.000*
Yol Yüzeyi			
Kuru/Tozlu	78.388	2	0.000*
Islak/Çamurlu/ Su ya da yağ birikintili	48.959	2	0.000*
Aydınlatma	20.179	2	0.000*
Araç sayısı			
Tek araçlı	19.950	2	0.000*
Zincirleme	10.609	2	0.005*
Araç cinsi			
Kamyonet/Kamyon	4.440	2	0.109

Otobüs/Minibüs	16.205	2	0.000*
Diğer	3.087	2	0.214
Yaş grubu			
16 yaşından küçük	333.851	2	0.000*
16-24 yaş arası	1.479	2	0.477
25-34 yaş arası	3.216	2	0.200
35-44 yaş arası	0.008	2	0.996
45-54 yaş arası	2.357	2	0.308
Öğrenim durumu			
İlköğretim	102.750	2	0.000*
Ortaöğretim	36.613	2	0.000*
Sürücü belgesi			
E Sınıfı	14.761	2	0.001*
B Sınıfı	18.284	2	0.000*
Kusur			
Hız	189.302	2	0.000*
Geçiş üstünlüğü ihlali	29.905	2	0.000*
Diğer kusurlar	87.240	2	0.000*

*0.01 anlam düzeyinde istatistiksel olarak anlamlı

4.5.2.4.7. Bağımsız Değişkenler için Wald Testi

Kurulan modeldeki bağımsız değişkenler için Wald testi yapabilmek için Stata 11.2 paket programında *mlogtest*, *all* komutu çalıştırılmıştır. Bağımsız değişkenlere ait Wald testi için kurulan sıfır hipotezi olabilirlik oranı testinde olduğu gibi,

H_0 : Verilen bağımsız değişken(ler)e ait bütün katsayılar sıfırdır.

şeklinde. Modeldeki bağımsız değişkenler için olabilirlik-oranı testinden elde edilen sonuçlar Tablo 4.13.'te sunulmuştur. Buna göre yıldız(lar)la işaretlenmiş bağımsız değişkenler için H_0 hipotezinin reddedileceğini ve kurulan model içerisinde bağımlı değişken üzerindeki etkilerinin istatistiksel olarak anlamlı olduğunu söylemek mümkündür. Bağımsız değişkenler için olabilirlik oranı ve Wald testleri sonuçlarının benzerlik gösterdiği görülmektedir.

Tablo 4. 13. Bağımsız Değişkenler için Wald Testi

Bağımsız Değişken	χ^2	Serbestlik Derecesi	P > χ^2
Mevsim			
Kış	64.676	2	0.000*
İlkbahar	13.876	2	0.001*
Yaz	1.769	2	0.413
Saat			
06:00-11:59	0.574	2	0.750
12:00-17:59	2.754	2	0.252
18:00-23:59	1.850	2	0.397
Haftaici/haftasonu			
Haftaici	0.918	2	0.632
Hava			
Açık	10.584	2	0.005*
Yağmurlu/Bulutlu	4.574	2	0.102
Gün durumu			
Gündüz	1.004	2	0.605
Yaya kaldırımı	84.768	2	0.000*
Yol Yüzeyi			
Kuru/Tozlu	74.819	2	0.000*
Islak/Çamurlu/ Su ya da yağ birikintili	48.084	2	0.000*
Aydınlatma	16.462	2	0.000*
Araç sayısı			
Tek araçlı	19.764	2	0.000*
Zincirleme	12.603	2	0.002*
Araç cinsi			
Kamyonet/Kamyon	4.544	2	0.103
Otobüs/Minibüs	17.576	2	0.000*
Diğer	3.487	2	0.175
Yaş grubu			
16 yaşından küçük	188.782	2	0.000*
16-24 yaş arası	1.407	2	0.495
25-34 yaş arası	3.394	2	0.183
35-44 yaş arası	0.008	2	0.996
45-54 yaş arası	2.243	2	0.326
Öğrenim durumu			
İlköğretim	99.846	2	0.000*
Ortaöğretim	36.063	2	0.000*
Sürücü belgesi			
E Sınıfı	14.313	2	0.001*
B Sınıfı	17.162	2	0.000*
Kusur			
Hız	164.936	2	0.000*
Geçiş üstünlüğü ihlali	27.722	2	0.000*
Diğer kusurlar	82.390	2	0.000*

*0.01 anlam düzeyinde istatistiksel olarak anlamlı

**0.05 anlam düzeyinde istatistiksel olarak anlamlı

4.5.2.4.8. Kurulan Modelin Tahmini ve Katsayıların Yorumlanması

Stata 11.2 paket programındaki *mlogit* komutunun tanımlanan bağımsız değişken ve bağımsız değişkenler için çalıştırılması ile çok durumlu logit modeline ait elde edilen sonuçlar Tablo 4.14.'te gösterilmiştir. Tablo 4.14'te görüldüğü üzere modelde kullanılan gözlem sayısı 9162 karayolu trafik kazası olup modele ait χ^2 değeri istatistiksel olarak anlamlıdır ($p < 0.01$). Modele ait pseudo- χ^2 değeri 0.172 ve log-olabilirlik değeri -5035.3739 bulunmuştur. Tablo 4.14.'ün ilk sütununda göreceli risk oranları (relative risk ratios) iki durumlu lojistik regresyon modellerinin yorumlanmasında kullanılan bahis oranı değerlerinin çok durumlu modellerdeki karşılığı olarak değerlendirilmektedir. Sözü edilen göreceli risk oranlarının % 95 güven aralığında (G.A.) alabilecekleri alt ve üst sınır değerleri ise Tablo 4.14.'ün son sütununda gösterilmiştir. Buna göre, maddi hasarlı kaza türü referans kategori olarak kabul edildiğinde, ölümlü kazalar için karayolu trafik kazasının meydana geldiği gün havanın açık olması, karayolu trafik kazasının meydana geldiği yolda aydınlatma ve yaya kaldırımının olup olmaması, zincirleme kazanın ve araç cinsinin otobüs/minibüs olmasının 0.01 önem seviyesinde; karayolu trafik kazasının meydana geldiği yolun yüzeyinin kuru/tozlu olmasının, tek araçlı kaza olmasının ve kazaya karışan aracın kamyon/kamyonet olmasının 0.05 önem seviyesinde istatistiksel olarak anlamlı olduğu tespit edilmiştir.

Tablo 4. 14. Karayolu Trafik Kazalarını Etkileyen Risk Faktörlerinin Çok Durumlu Logit Modeli ile Analizi

Bağımsız Değişken	RRR	Standart Hata	z	P > z	[%95 G. A.]
Ölümlü kaza					
Mevsim					
Kış	0.45	0.227	-1.58	0.114	0.17-1.21
İlkbahar	0.98	0.399	-0.05	0.962	0.44-2.17
Yaz	1.49	0.549	1.08	0.279	0.72-3.07
Saat					
06:00-11:59	0.94	0.632	-0.09	0.926	0.25-3.51
12:00-17:59	0.35	0.226	-1.63	0.104	0.10-1.24
18:00-23:59	0.58	0.292	-1.08	0.282	0.22-1.55
Hafta içi/Haftasonu					
Hafta içi	0.76	0.230	-0.90	0.367	0.42-1.38
Hava durumu					
Açık	0.18	0.112	-2.76	0.006*	0.05-0.61
Yağmurlu/Bulutlu	0.32	0.195	-1.87	0.061	0.10-1.06
Gün durumu					

Gündüz	0.61	0.310	-0.97	0.331	0.22-1.65
Yaya kaldırımı	0.19	0.076	-4.17	0.000*	0.09-0.42
Yol yüzeyi					
Kuru/Tozlu	4.12	2.555	2.28	0.022**	1.22-13.89
Islak/Çamurlu/ Su ya da yağ birikintili	2.84	1.753	1.69	0.090	0.85-9.52
Aydınlatma	0.22	0.084	-3.96	0.000*	0.10-0.46
Araç sayısı					
Tek Araçlı	2.63	1.037	2.46	0.014**	1.22-5.70
Zincirleme	6.32	3.483	3.34	0.001*	2.15-18.61
Araç cinsi					
Kamyonet/Kamyon	1.93	0.646	1.97	0.049**	1.00-3.72
Otobüs/Minibüs	4.53	2.117	3.23	0.001*	1.81-11.32
Diğer	2.53	1.268	1.85	0.064	0.95-6.76
Yaş grubu					
16 yaşından küçük	18.69	17.733	8.09	0.002*	2.91-120
16-24 yaş arası	0.46	0.301	-1.19	0.236	0.13-1.66
25-34 yaş arası	0.40	0.205	-1.79	0.074	0.15-1.09
35-44 yaş arası	0.96	0.460	-0.08	0.935	0.38-2.46
45-54 yaş arası	2.01	0.953	1.49	0.137	0.80-5.09
Öğrenim durumu					
İlköğretim	2.20	0.902	1.92	0.055	0.98-4.91
Ortaöğretim	0.93	0.428	-0.15	0.878	0.38-2.30
Sürücü belgesi					
E Sınıfı	1.98	0.996	1.37	0.172	0.74-5.31
B Sınıfı	0.06	0.054	-0.08	0.002*	0.01-0.35
Kusur					
Hız	5.52	5.953	1.58	0.114	0.67-45.74
Geçiş üstünlüğü ihlali	3.69	3.943	1.22	0.222	0.45-29.99
Diğer kusurlar	7.55	8.182	1.86	0.062	0.90-63.20
Sabit terim	-	1.481	-3.33	0.001*	-
Yaralanmalı kaza					
Mevsim					
Kış	0.52	0.043	- 7.98	0.000*	0.44-0.61
İlkbahar	0.77	0.055	- 3.71	0.000*	0.67-0.88
Yaz	0.96	0.065	- 0.65	0.513	0.84-1.10
Saat					
06:00-11:59	1.11	0.159	0.71	0.462	0.84-1.47
12:00-17:59	1.02	0.137	0.13	0.895	0.78-1.32
18:00-23:59	1.08	0.124	0.69	0.491	0.86-1.35
Hafta içi/Haftasonu					
Hafta içi	0.98	0.056	- 0.41	0.684	0.87-1.09
Hava durumu					
Açık	0.18	0.112	-2.76	0.006*	0.05-0.61
Yağmurlu/Bulutlu	0.32	0.195	-1.87	0.061	0.10-1.06
Gün durumu					
Gündüz	0.61	0.310	-0.97	0.331	0.22-1.65
Yaya kaldırımı	0.19	0.076	-4.17	0.000*	0.09-0.42
Yol yüzeyi					
Kuru/Tozlu	4.12	2.555	2.28	0.022**	1.22-13.89
Islak/Çamurlu/ Su ya da yağ birikintili	2.84	1.753	1.69	0.090	0.85-9.52
Aydınlatma	0.22	0.084	-3.96	0.000*	0.10-0.46
Araç sayısı					
Tek Araçlı	2.63	1.037	2.46	0.014**	1.22-5.70
Zincirleme	6.32	3.483	3.34	0.001*	2.15-18.61
Araç cinsi					

Kamyonet/Kamyon	1.93	0.646	1.97	0.049**	1.00–3.72
Otobüs/Minibüs	4.53	2.117	3.23	0.001*	1.81–11.32
Diğer	2.53	1.268	1.85	0.064	0.95–6.76
Yaş grubu					
16 yaşından küçük	18.69	17.733	8.09	0.002*	2.91–120
16-24 yaş arası	0.46	0.301	-1.19	0.236	0.13–1.66
25-34 yaş arası	0.40	0.205	-1.79	0.074	0.15–1.09
35-44 yaş arası	0.96	0.460	-0.08	0.935	0.38–2.46
45-54 yaş arası	2.01	0.953	1.49	0.137	0.80–5.09
Öğrenim durumu					
İlköğretim	2.20	0.902	1.92	0.055	0.98–4.91
Ortaöğretim	0.93	0.428	-0.15	0.878	0.38–2.30
Sürücü belgesi					
E Sınıfı	1.98	0.996	1.37	0.172	0.74–5.31
B Sınıfı	0.06	0.054	-0.08	0.002*	0.01–0.35
Kusur					
Hız	5.52	5.953	1.58	0.114	0.67–45.74
Geçiş üstünlüğü ihlali	3.69	3.943	1.22	0.222	0.45–29.99
Diğer kusurlar	7.55	8.182	1.86	0.062	0.90–63.20
Sabit terim	-	0.244	-10.18	0.000*	-
Maddi Hasarlı Kaza (Referans Kategori)					

*0.01 anlam düzeyinde istatistiksel olarak anlamlı

**0.05 anlam düzeyinde istatistiksel olarak anlamlı

Görel risk oranlarına ait değerlerin yorumlanmasına geri dönüldüğünde, karayolu trafik kazasının meydana geldiği günde havanın açık olmasının (RRR = 0.18, % 95 G.A. = 0.05-0.61, $p < 0.01$) ölümlü kaza olma olasılığını azalttığı görülmektedir. Başka bir ifade ile hava değişkenindeki bir birimlik değişim ölümlü kaza meydana gelme olasılığını 0.18 birim azaltmaktadır. Benzer şekilde karayolu trafik kazasının meydana geldiği yolda yaya kaldırımının (RRR = 0.19, % 95 G.A. = 0.09-0.42, $p < 0.01$) ile aydınlatmanın olmasının (RRR = 0.22, % 95 G.A. = 0.10-0.46, $p < 0.01$) ve kazaya karışan araç sürücüsünün B sınıfı ehliyeteye sahip olmasının (RRR = 0.06, % 95 G.A. = 0.01-0.035, $p < 0.01$) ölümlü trafik kazasının meydana gelme olasılığını azaltan diğer risk faktörleri olduğu tespit edilmiştir. Öte yandan, karayolu trafik kazasının meydana geldiği yolun yüzeyinin kuru ya da tozlu olması (RRR = 4.12, % 95 G.A. = 1.22-13.89, $p < 0.05$), kazanın tek araçlı (RRR = 2.63, % 95 G.A. = 1.22-5.70, $p < 0.01$) ya da zincirleme (RRR = 6.32, % 95 G.A. = 2.15-18.61, $p < 0.01$) kaza olması, araç cinsinin kamyonet/kamyon (RRR = 1.93, % 95 G.A. = 1.00-3.72, $p < 0.05$) ya da otobüs/minibüs (RRR = 4.53, % 95 G.A. = 1.81- 11.32, $p < 0.01$) olması, kazaya karışan araç sürücüsünün 16 yaşından küçük (RRR = 18.69, % 95 G.A. = 2.91-120, $p < 0.01$) olması ölümlü karayolu trafik kazasının olma olasılığını artırmaktadır.

Elde edilen sonuçların genel bir değerlendirmesi yapıldığında ölümlü karayolu trafik kazalarının olma olasılığını en çok artıran risk faktörü araç sürücüsünün 16 yaşından küçük olmasıdır. Hatırlanacağı üzere, Dünya Sağlık Örgütü'nün trafik güvenliği ile ilgili raporlarında da yaş gruplarına dikkat çekilerek 15-29 yaş aralığındaki karayolu kullanıcılarının toplam ölümlerin çok büyük bölümünü oluşturduğu belirtilmektedir. Araştırma bulgularına göre elde edilen yaş aralığının bir bölümü Dünya Sağlık Örgütü'nün ortaya koyduğu yaş aralığında yer almaktadır. Göreceli risk oranlarının değerine göre yaş grubu değişkeni takip eden ve ölümlü karayolu trafik kazalarını en çok etkileyen risk faktörleri sırasıyla karayolu trafik kazasının zincirleme kaza olarak gerçekleşmiş olması, kazaya karışan araç cinsinin otobüs ya da minibüs olması ve kazanın meydana geldiği yol yüzeyinin kuru ya da tozlu olmasıdır. Ölümlü karayolu trafik kazasının meydana gelme olasılığını en çok düşüren risk faktörleri ise havanın kazanın meydana geldiği yolda aydınlatmanın ile yaya kaldırımının olması ve kazanın meydana geldiği gün havanın açık olmasıdır.

Göreceli risk oranları ve istatistiksel anlamlılık değerleri üzerinden yaralanmalı karayolu trafik kazalarının olma olasılığı etkileyen faktörlere bakıldığında ise kazanın meydana geldiği geldiği mevsimin kış ($RRR = 0.52$, % 95 G.A. = 0.44-0.61, $p < 0.01$) ya da ilkbahar ($RRR = 0.77$, % 95 G.A. = 0.67-0.88, $p < 0.01$) olması, kazanın meydana geldiği günde havanın açık ($RRR = 0.79$, % 95 G.A. = 0.62-1.00, $p < 0.05$) olması, kazaya karışan araç sürücüsünün B sınıfı sürücü belgesine ($RRR = 0.75$, % 95 G.A. = 0.61-0.91, $p < 0.01$) sahip olması yaralanmalı karayolu trafik kazalarının meydana gelme olasılığını azaltmaktadır. Yaralanmalı trafik kazalarının meydana gelme olasılığını artıran risk faktörleri ise yol yüzeyinin kuru ya da tozlu ($RRR = 2.41$, % 95 G.A. = 1.97-2.95, $p < 0.01$) ıslak, çamurlu, su ya da yağ birikintili ($RRR = 2.04$, % 95 G.A. = 1.66-2.50, $p < 0.01$) olması, kazaya karışan araç cinsinin otobüs ya da minibüs ($RRR = 1.34$, % 95 G.A. = 1.11-1.61, $p < 0.01$) olması, kazaya karışan araç sürücüsünün 16 yaşından küçük ($RRR = 41.50$, % 95 G.A. = 24.38-70.63, $p < 0.01$), öğrenim durumunun ilköğretim ($RRR = 2.09$, % 95 G.A. = 1.81-2.42, $p < 0.01$) ve ortaöğretim ($RRR = 1.54$, % 95 G.A. = 1.34-1.77, $p < 0.01$), sürücü belgesinin E sınıfı ($RRR = 1.42$, % 95 G.A. = 1.18-1.72, $p < 0.01$) olması, kazanın meydana gelme sebebinin aşırı hız ($RRR = 3.90$, % 95 G.A. = 3.17-4.80, $p < 0.01$), geçiş üstünlüğü

ihlali (RRR = 1.69, % 95 G.A. = 1.38-4.06, $p < 0.01$) ve diğer kusurlar (RRR = 2.90, % 95 G.A. = 2.30-3.37) olmasıdır.

Yaralanmalı karayolu trafik kazalarının meydana gelme olasılığını etkileyen faktörlerin genel bir değerlendirmesi yapıldığında ise bu kazaların meydana gelme olasılığını en çok artıran faktör, araç sürücüsünün 16 yaşından küçük olmasıdır. Daha sonra karayolu trafik kazasına karışan aracın aşırı hızlı olması, araç sürücüsünün Ek 1’de detaylı olarak gösterilen diğer trafik kurallarından birini ihlal etmiş olması, yol yüzeyi ve araç sürücüsünün öğrenim durumu değişkenleri yaralanmalı kazaların meydana gelmesini en çok etkileyen faktörler olarak tespit edilmiştir. Araştırmanın veri tasnifi aşamasında araştırmacının gözlemine göre araç sürücüleri daha çok dikkatsizlik sebebiyle karayolu trafik kazalarına karışmaktadır. Araştırma bulgularına göre, yaralanmalı trafik kazalarının meydana gelme olasılığını en çok azaltan risk faktörleri ise mevsimin ilkbahar olması, araç sürücüsünün B sınıfı sürücü belgesine sahip olması, mevsimin kış olması gibi faktörlerdir.

4.5.2.4.8.1. Referans Kategorisinin Değiştirilmesi ile Elde Edilen Bulguların Yorumlanması

Çok durumlu logit modelinde, bahis oranı katsayılarının yorumlanması hususundaki en büyük güçlük, bir değişkenin etkisini görebilmek için bağımlı değişkene ait kategoriler açısından katsayı karşılaştırması yapabilmektir. Çok durumlu logit modelin klasik sonuç tablosu referans kategori için yalnızca $J - 1$ adet karşılaştırma yapmaktadır (Long ve Freese, 2001: 204). Stata 11.2 paket programında çok durumlu model analiz edildikten sonra *listcoef* komutunun çalıştırılmasıyla bağımlı değişkene ait kategorilerin ikili bahis oranları karşılaştırmalarını gösteren tablolar elde edilmektedir. Karşılaştırılması yapılan yeni modeller de ilişkisiz alternatiflerin bağımsızlığı varsayımını sağlamış olup araştırmanın bu bölümünde, elde edilen bu katsayılardan istatistiksel olarak anlamlı bulunan değerler değerlendirilmiştir.

Tablo 4.15’te mevsim değişkeninin bağımlı değişken üzerindeki etkisi sunulmaktadır. Buna göre, kış mevsiminde diğer risk faktörleri sabit kaldığında maddi hasarlı kazaların yaralanmalı kazalara göre meydana gelme olasılığının (OR = 1.93, $p < 0.01$) yaklaşık iki kat daha fazla olduğu görülmektedir. İlkbahar mevsiminde maddi hasarlı kazaların yaralanmalı kazalara göre meydana gelme olasılığının (OR = 1.31, $p <$

0.01) biraz daha olduğu tespit edilmiştir. Yaz mevsiminde, bağımlı değişkene ait referans kategorisinin değiştirilmesi ile elde edilen bahis oranları değerlerinin hiçbirisi istatistiksel olarak anlamlı bulunamamıştır.

Tablo 4. 15. Karayolu Trafik Kazasının Meydana Geldiği Mevsime Göre Bahis Oranlarının İkili Karşılaştırılması

Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = Kış			
	B	Z	P > z	e ^b
Ölümlü kaza – yaralanmalı kaza	-0.14	-0.27	0.786	0.88
Ölümlü kaza – maddi hasarlı kaza	-0.79	-1.58	0.114	0.45
Yaralanmalı kaza – ölümlü kaza	0.14	0.27	0.786	1.16
Yaralanmalı kaza – maddi hasarlı kaza	-0.66	-7.99	0.000*	0.52
Maddi hasarlı kaza – ölümlü kaza	0.79	1.59	0.114	2.21
Maddi hasarlı kaza – yaralanmalı kaza	0.66	7.99	0.000*	1.93
Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = İlkbahar			
	B	Z	P > z	e ^b
Ölümlü kaza – yaralanmalı kaza	0.25	0.61	0.543	1.28
Ölümlü kaza – maddi hasarlı kaza	-0.02	-0.05	0.962	0.98
Yaralanmalı kaza – ölümlü kaza	-0.25	-0.61	0.543	0.78
Yaralanmalı kaza – maddi hasarlı kaza	-0.27	-3.71	0.000*	0.90
Maddi hasarlı kaza – ölümlü kaza	0.02	0.05	0.962	1.1
Maddi hasarlı kaza – yaralanmalı kaza	0.27	3.71	0.000*	1.31
Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = Yaz			
	B	Z	P > z	e ^b
Ölümlü kaza – yaralanmalı kaza	0.44	1.21	0.227	1.56
Ölümlü kaza – maddi hasarlı kaza	0.40	1.09	0.279	1.49
Yaralanmalı kaza – ölümlü kaza	-0.44	-1.21	0.227	0.64
Yaralanmalı kaza – maddi hasarlı kaza	-0.04	-0.66	0.513	0.96
Maddi hasarlı kaza – ölümlü kaza	-0.40	-1.09	0.279	0.67
Maddi hasarlı kaza – yaralanmalı kaza	0.04	0.66	0.513	1.05

*0.01 anlam düzeyinde istatistiksel olarak anlamlı

**0.05 anlam düzeyinde istatistiksel olarak anlamlı

Tablo 4.16.'da karayolu trafik kazasının meydana geldiği saate göre bahis oranlarının ikili karşılaştırılması gösterilmektedir. Karayolu trafik kazasının meydana geldiği saate göre bağımlı değişkene ait referans kategorisinin değiştirilmesi ile elde edilen bahis oranları değerlerinin hiçbirisi istatistiksel olarak anlamlı bulunamamıştır.

Tablo 4. 16. Karayolu Trafik Kazasının Meydana Geldiği Saate Göre Bahis Oranlarının İkili Karşılaştırılması

Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = 05:59-11:59			
	b	z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	-0.17	-0.25	0.803	0.85
Ölümlü kaza – maddi hasarlı kaza	-0.06	-0.09	0.926	0.94
Yaralanmalı kaza – ölümlü kaza	0.17	0.25	0.803	1.18
Yaralanmalı kaza – maddi hasarlı kaza	-0.11	-0.74	0.462	0.90
Maddi hasarlı kaza – ölümlü kaza	0.06	0.09	0.926	1.06
Maddi hasarlı kaza – yaralanmalı kaza	-0.11	0.74	0.462	1.11
Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = 12:00-17:59			
	b	z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	-1.05	-1.66	0.097	0.35
Ölümlü kaza – maddi hasarlı kaza	-1.04	-1.63	0.104	0.35
Yaralanmalı kaza – ölümlü kaza	1.05	1.66	0.097	2.87
Yaralanmalı kaza – maddi hasarlı kaza	0.02	0.13	0.895	1.02
Maddi hasarlı kaza – ölümlü kaza	1.04	1.63	0.104	2.82
Maddi hasarlı kaza – yaralanmalı kaza	-0.02	-0.13	0.895	0.98
Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = 18:00-23:59			
	b	z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	-0.62	-1.24	0.216	0.54
Ölümlü kaza – maddi hasarlı kaza	-0.54	-1.08	0.282	0.58
Yaralanmalı kaza – ölümlü kaza	0.62	1.24	0.216	1.85
Yaralanmalı kaza – maddi hasarlı kaza	0.08	0.69	0.491	1.08
Maddi hasarlı kaza – ölümlü kaza	0.54	1.08	0.282	1.71
Maddi hasarlı kaza – yaralanmalı kaza	-0.08	-0.69	0.491	0.92

Tablo 4.17’de karayolu trafik kazasının hafta içi ya da hafta sonu meydana gelişine göre bahis oranlarının ikili karşılaştırılması sunulmuştur. Bağımsız değişkenin hafta içi ya da haftasonu olması durumunda bağımlı değişkene ait referans kategorilerinin değiştirilmesi sonucu elde edilen bahis oranlarının hiçbirisi istatistiksel olarak anlamlı bulunamamıştır.

Tablo 4. 17. Karayolu Trafik Kazasının Meydana Geldiği Haftaya Göre Bahis Oranlarının İkili Karşılaştırılması

Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = Hafta			
	b	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	-0.25	-0.83	0.409	0.78
Ölümlü kaza – maddi hasarlı kaza	-0.27	-0.90	0.367	0.76
Yaralanmalı kaza – ölümlü kaza	0.25	0.83	0.409	1.28
Yaralanmalı kaza – maddi hasarlı kaza	0.02	0.408	0.684	1.02
Maddi hasarlı kaza – ölümlü kaza	0.27	0.90	0.367	1.31
Maddi hasarlı kaza – yaralanmalı kaza	-0.02	-0.408	0.684	0.98

Tablo 4.18.'de karayolu trafik kazasının meydana geldiği gündeki hava durumuna göre bahis oranlarının karşılaştırılması sunulmuştur. Buna göre, karayolu trafik kazasının meydana geldiği günde diğer risk faktörleri sabit tutulduğunda yaralanmalı kazaların ölümlü kazalara göre meydana gelme olasılığının (OR = 4.34, $p < 0.05$) yaklaşık 4 kat, maddi hasarlı kazaların ölümlü kazalara göre meydana gelme olasılığının (OR = 5.52, $p < 0.01$) yaklaşık 5.5 kat daha fazla olduğu görülmektedir. Benzer şekilde, maddi hasarlı kazaların yaralanmalı kazalara göre meydana gelme olasılığının (OR = 1.27, $p < 0.05$) biraz daha fazla olduğu tespit edilmiştir. Karayolu trafik kazasının meydana geldiği günde havanın yağmurlu ya da bulutlu olması durumunda elde edilen bahis oranları değerleri 0.05 ve 0.01 anlam seviyesinde anlamlı bulunamamıştır.

Tablo 4. 18. Karayolu Trafik Kazasının Meydana Geldiği Gündeki Hava Durumuna Göre Bahis Oranlarının İkili Karşılaştırılması

Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = Açık			
	B	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	-1.47	-2.38	0.018**	0.23
Ölümlü kaza – maddi hasarlı kaza	-1.71	-2.76	0.006*	0.18
Yaralanmalı kaza – ölümlü kaza	1.47	2.38	0.018**	4.34

Yaralanmalı kaza – maddi hasarlı kaza	-0.24	-1.96	0.049**	0.79
Maddi hasarlı kaza – ölümlü kaza	1.71	2.76	0.006*	5.52
Maddi hasarlı kaza – yaralanmalı kaza	0.24	1.96	0.049**	1.27
Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = Yağmurlu			
	B	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	-0.98	-1.62	0.105	0.38
Ölümlü kaza – maddi hasarlı kaza	-1.13	-1.87	0.061	0.32
Yaralanmalı kaza – ölümlü kaza	0.98	1.62	0.105	2.68
Yaralanmalı kaza – maddi hasarlı kaza	-0.15	1.20	0.229	0.86
Maddi hasarlı kaza – ölümlü kaza	1.13	1.87	0.061	3.10
Maddi hasarlı kaza – yaralanmalı kaza	0.15	1.20	0.229	1.16

*0.01 anlam düzeyinde istatistiksel olarak anlamlı

**0.05 anlam düzeyinde istatistiksel olarak anlamlı

Tablo 4.19.'da karayolu trafik kazasının gündüz meydana gelmesi durumunda bağımlı değişkene ait bahis oranlarının ikili karşılaştırmaları sunulmuştur. Buna göre, karayolu trafik kazasının gündüz meydana gelmesi durumunda elde edilen bahis oranları değerleri istatistiksel olarak anlamlı bulunamamıştır.

Tablo 4. 19. Karayolu Trafik Kazasının Meydana Geldiği Güne Göre Bahis Oranlarının İkili Karşılaştırılması

Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = Gündüz			
	b	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	-0.46	-0.91	0.362	0.63
Ölümlü kaza – maddi hasarlı kaza	-0.49	-0.97	0.331	0.61
Yaralanmalı kaza – ölümlü kaza	0.46	0.91	0.362	1.59
Yaralanmalı kaza – maddi hasarlı kaza	-0.03	-0.34	0.737	0.97
Maddi hasarlı kaza – ölümlü kaza	0.49	0.97	0.331	1.64
Maddi hasarlı kaza – yaralanmalı kaza	0.03	0.34	0.737	1.03

Tablo 4.20.'de karayolu trafik kazasının meydana geldiği yolda yaya kaldırımının olup olmamasına göre bahis oranlarının ikili karşılaştırılması gösterilmiştir. Buna göre, karayolu trafik kazasının meydana geldiği yolda yaya kaldırımının olması durumunda, ölümlü trafik kazalarının meydana gelme olasılığının (OR = 3.28, $p < 0.01$) ölümlü trafik kazalarına göre yaklaşık 3 kat daha fazla olduğu görülmektedir. Benzer şekilde, karayolu trafik kazasının meydana geldiği yolda yaya kaldırımının olması durumunda, yaralanmalı trafik kazasının meydana gelme olasılığının (OR = 0.64, $p < 0.01$) maddi hasarlı trafik kazasına göre biraz daha fazla olduğu tespit edilmiştir.

Tablo 4. 20. Karayolu Trafik Kazasının Meydana Geldiği Yolda Yaya Kaldırımının Olup Olmamasına Göre Bahis Oranlarının İkili Karşılaştırılması

Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = Yaya Kaldırımı			
	b	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	-1.19	-3.02	0.003*	0.31
Ölümlü kaza – maddi hasarlı kaza	-1.64	-4.17	0.060	5.16
Yaralanmalı kaza – ölümlü kaza	1.19	3.02	0.003*	3.28
Yaralanmalı kaza – maddi hasarlı kaza	-0.45	-8.44	0.000*	0.64
Maddi hasarlı kaza – ölümlü kaza	1.64	4.17	0.000*	0.19
Maddi hasarlı kaza – yaralanmalı kaza	0.45	8.44	0.060	1.57

*0.01 anlam düzeyinde istatistiksel olarak anlamlı

**0.05 anlam düzeyinde istatistiksel olarak anlamlı

Tablo 4.21.'de karayolu trafik kazasının meydana geldiği yolun yüzeyine göre bağımlı değişken üzerindeki bahis oranlarının ikili karşılaştırılması gösterilmiştir. Buna göre, yol yüzeyinin kuru ya da tozlu ve diğer risk faktörlerinin sabit olması durumunda, ölümlü kazaların maddi hasarlı kazalara göre meydana gelme olasılığı (OR = 4.12, $p < 0.05$) yaklaşık 4 kat, yaralanmalı kazaların maddi hasarlı kazalara göre meydana gelme olasılığının yaklaşık 2.5 kat (OR = 2.41, $p < 0.01$) fazla olduğu görülmüştür. Benzer şekilde, yolun yüzeyinin ıslak/çamurlu/su ya da yağ birikintili olması durumunda, yaralanmalı kazaların maddi hasarlı kazalara göre gerçekleşme olasılığının yaklaşık 2 kat daha fazla olduğu tespit edilmektedir.

Tablo 4. 21. Karayolu Trafik Kazasının Meydana Geldiği Yolu Yüzeyine Göre Bahis Oranlarının İkili Karşılaştırılması

Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = Kuru/tozlu			
	B	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	0.54	0.86	0.388	1.71
Ölümlü kaza – maddi hasarlı kaza	1.42	2.28	0.022*	4.12
Yaralanmalı kaza – ölümlü kaza	-0.54	-0.86	0.388	0.58
Yaralanmalı kaza – maddi hasarlı kaza	0.88	8.49	0.000*	2.41
Maddi hasarlı kaza – ölümlü kaza	-1.42	-2.28	0.022*	0.24
Maddi hasarlı kaza – yaralanmalı kaza	-0.88	-8.49	0.000*	0.65
Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = Islak/Çamurlu/Su ya da yağ birikintili			
	b	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	0.33	0.54	0.591	1.39
Ölümlü kaza – maddi hasarlı kaza	1.04	1.69	0.090	2.84
Yaralanmalı kaza – ölümlü kaza	-0.33	-0.54	0.591	0.72
Yaralanmalı kaza – maddi hasarlı kaza	0.71	6.83	0.000*	2.04
Maddi hasarlı kaza – ölümlü kaza	-1.044	-1.69	0.090	0.35
Maddi hasarlı kaza – yaralanmalı kaza	-0.71	-6.83	0.000*	0.49

*0.01 anlam düzeyinde istatistiksel olarak anlamlı

**0.05 anlam düzeyinde istatistiksel olarak anlamlı

Tablo 4.22.'de karayolu trafik kazasının meydana geldiği yolda aydınlatma olup olmamasına göre bahis oranlarının ikili karşılaştırılması sunulmuştur. Buna göre, karayolu trafik kazasının meydana geldiği yolda aydınlatma olduğunda maddi hasarlı trafik kazalarının meydana gelme olasılığı (OR = 4.64, $p < 0.01$) ölümlü kazalara göre yaklaşık 5 kat, yaralanmalı trafik kazalarının ölümlü trafik kazalarına göre meydana gelme olasılığının (OR = 4.35, $p < 0.01$) yaklaşık 4 kat daha fazla olduğu görülmektedir.

Tablo 4. 22. Karayolu Trafik Kazasının Meydana Geldiği Yolda Aydınlatma Olup Olmamasına Göre Bahis Oranlarının İkili Karşılaştırılması

Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = Aydınlatma			
	B	Z	P > z	e ^b
Ölümlü kaza – yaralanmalı kaza	-1.47	-3.8	0.000*	0.23
Ölümlü kaza – maddi hasarlı kaza	-1.53	-3.96	0.000*	0.22
Yaralanmalı kaza – ölümlü kaza	1.47	3.8	0.000*	4.35
Yaralanmalı kaza – maddi hasarlı kaza	-0.06	-1.2	0.248	0.94
Maddi hasarlı kaza – ölümlü kaza	1.53	3.96	0.000*	4.64
Maddi hasarlı kaza – yaralanmalı kaza	0.06	1.2	0.248	1.07

*0.01 anlam düzeyinde istatistiksel olarak anlamlı

**0.05 anlam düzeyinde istatistiksel olarak anlamlı

Tablo 4.23.'te bağımlı değişkene ait kategorilerin ikili değiştirilmesi ile kurulan modeldeki araç sayısı değişkeninin alt kategorilerine ait bahis oranları değişimleri gösterilmiştir. Buna göre, diğer risk faktörleri sabit kaldığında tek araçlı kazalar için ölümlü kazaların meydana gelme olasılığının (OR = 2.63, $p < 0.05$) maddi hasarlı kazalara göre 2.63 kat daha fazla olduğu görülmektedir. Benzer şekilde, yaralanmalı kazaların meydana gelme olasılığının (OR = 0.79, $p < 0.01$) maddi hasarlı kazalara göre biraz daha fazla olduğu tespit edilmiştir. Zincirleme kazaların bağımlı değişkene etkisine bakıldığında, diğer bütün risk faktörleri sabit kaldığında, ölümlü kazaların yaralanmalı kazalara göre meydana gelme olasılığının (OR = 5.28, $p < 0.01$) 5.28 kat, ölümlü kazaların maddi hasarlı kazalara göre meydana gelme olasılığının (OR = 6.32, $p < 0.01$) kat daha fazla olduğu görülmektedir. Ayrıca, maddi hasarlı kazaların yaralanmalı kazalara göre meydana gelme olasılığının (OR = 1.26, $p < 0.01$) 1.26 kat daha fazla olduğu tespit edilmiştir.

Tablo 4. 23. Karayolu Trafik Kazasına Karışan Araç sayısına Göre Bahis Oranlarının İkili Karşılaştırılması

Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = Tek araçlı kaza			
	B	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	0.73	1.87	0.062	2.09
Ölümlü kaza – maddi hasarlı kaza	0.97	2.46	0.014**	2.63
Yaralanmalı kaza – ölümlü kaza	-0.73	-1.87	0.062	0.48
Yaralanmalı kaza – maddi hasarlı kaza	0.23	3.87	0.000*	0.79
Maddi hasarlı kaza – ölümlü kaza	-0.97	-2.46	0.014**	0.38
Maddi hasarlı kaza – yaralanmalı kaza	-0.23	-3.87	0.000*	1.26
Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = Zincirleme kaza			
	B	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	1.66	3.01	0.003*	5.28
Ölümlü kaza – maddi hasarlı kaza	1.84	3.35	0.001*	6.32
Yaralanmalı kaza – ölümlü kaza	-1.66	-3.01	0.003*	0.19
Yaralanmalı kaza – maddi hasarlı kaza	0.18	1.52	0.000*	0.79
Maddi hasarlı kaza – ölümlü kaza	-1.84	-3.35	0.001*	0.16
Maddi hasarlı kaza – yaralanmalı kaza	-0.18	-1.52	0.000*	1.26

*0.01 anlam düzeyinde istatistiksel olarak anlamlı

**0.05 anlam düzeyinde istatistiksel olarak anlamlı

Tablo 4.24’te karayolu trafik kazasına karışan araç cinsine göre bağımlı değişken bahis oranlarının ikili karşılaştırılması sunulmuştur. Tablo 4.24’te görüldüğü gibi trafik kazasına karışan aracın kamyon ya da kamyonet olması durumunda ölümlü trafik kazasının meydana gelme olasılığının (OR = 2.01, $p < 0.05$) yaralanmalı ve maddi hasarlı trafik kazasına göre yaklaşık 2 kat daha fazla olduğu tespit edilmiştir. Benzer şekilde, karayolu trafik kazasına karışan aracın otobüs ya da minibüs olması durumunda ölümlü trafik kazasının meydana gelme olasılığının (OR = 3.39, $p < 0.01$) yaralanmalı trafik kazasına göre yaklaşık 3 kat, maddi hasarlı trafik kazasına göre (OR = 4.53, $p < 0.01$) yaklaşık 4.5 kat fazla olduğu görülmektedir. Karayolu trafik kazasına karışan araç cinsinin otobüs ya da minibüs olması durumunda, yaralanmalı trafik kazasının meydana gelme olasılığının (OR = 1.93, $p < 0.01$) maddi hasarlı trafik kazasına göre yaklaşık 2

kat daha fazla olduğu tespit edilmiştir. Öte yandan, karayolu trafik kazasına karışan araç cinsinin arazi taşıtı/traktör/çekici, vb olması durumunda bağımlı değişkene ait referans kategorilerinin değiştirilmesi sonucu elde edilen bahis oranları istatistiksel olarak anlamlı bulunamamıştır.

Tablo 4. 24. Karayolu Trafik Kazasına Karışan Araç Cinsine Göre Bahis Oranlarının İkili Karşılaştırılması

Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = Kamyon/kamyonet			
	B	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	0.70	2.09	0.037**	2.01
Ölümlü kaza – maddi hasarlı kaza	0.66	1.97	0.049**	1.93
Yaralanmalı kaza – ölümlü kaza	-0.70		0.037**	0.50
Yaralanmalı kaza – maddi hasarlı kaza	-0.04	-0.63	0.529	0.96
Maddi hasarlı kaza – ölümlü kaza	-0.66	-1.97	0.049**	0.52
Maddi hasarlı kaza – yaralanmalı kaza	0.04	0.63	0.529	1.04
Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = Otobüs/minibüs			
	B	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	1.22	2.62	0.009*	3.39
Ölümlü kaza – maddi hasarlı kaza	1.51	3.23	0.001*	4.53
Yaralanmalı kaza – ölümlü kaza	-1.22	-2.62	0.009*	0.29
Yaralanmalı kaza – maddi hasarlı kaza	0.29	3.02	0.003*	1.34
Maddi hasarlı kaza – ölümlü kaza	-1.51	-3.23	0.001*	0.22
Maddi hasarlı kaza – yaralanmalı kaza	-0.29	-3.02	0.003*	0.75
Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = Arazi taşıtı/Traktör/Çekici/Diğer			
	B	z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	0.92	1.84	0.066	2.50
Ölümlü kaza – maddi hasarlı kaza	0.93	1.85	0.064	2.53
Yaralanmalı kaza – ölümlü kaza	-0.92	-1.84	0.066	0.40
Yaralanmalı kaza – maddi hasarlı kaza	0.01	0.074	0.941	0.99
Maddi hasarlı kaza – ölümlü kaza	-0.92	-1.85	0.064	0.40
Maddi hasarlı kaza – yaralanmalı kaza	-0.01	-0.074	0.941	1.01

*0.01 anlam düzeyinde istatistiksel olarak anlamlı

**0.05 anlam düzeyinde istatistiksel olarak anlamlı

Tablo 4.25’te karayolu trafik kazasına karışan araç sürücüsünün yaş grubuna göre bahis oranlarının ikili karşılaştırılması sunulmuştur. Buna göre, karayolu trafik kazasına karışan araç sürücüsünün 16 yaşından küçük olması durumunda, ölümlü trafik kazasının meydana gelme olasılığının (OR = 18.70, $p < 0.01$) maddi hasarlı trafik kazasına göre yaklaşık 19 kat, yaralanmalı trafik kazasının meydana gelme olasılığının (OR = 41.50, $p < 0.01$) maddi hasarlı trafik kazasına göre yaklaşık 42 kat daha fazla olduğu görülmektedir. Bağımsız değişken olarak seçilen diğer yaş gruplarına ait bahis oranları karşılaştırmaları istatistiksel olarak anlamlı bulunamamıştır.

Tablo 4. 25. Karayolu Trafik Kazasına Karışan Araç Sürücüsünün Yaş Grubuna Göre Bahis Oranlarının İkili Karşılaştırılması

Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = 16 yaşından küçük			
	b	z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	-0.80	-0.87	0.385	0.45
Ölümlü kaza – maddi hasarlı kaza	2.93	3.09	0.002*	18.70
Yaralanmalı kaza – ölümlü kaza	0.80	0.87	0.385	2.22
Yaralanmalı kaza – maddi hasarlı kaza	3.73	13.732	0.000*	41.50
Maddi hasarlı kaza – ölümlü kaza	-2.93	-3.09	0.002*	0.05
Maddi hasarlı kaza – yaralanmalı kaza	-3.73	13.732	0.000*	0.02
Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = 16-24 yaş arası			
	b	z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	-0.76	-1.17	0.241	0.47
Ölümlü kaza – maddi hasarlı kaza	-0.77	-1.19	0.236	0.46
Yaralanmalı kaza – ölümlü kaza	0.76	1.17	0.241	2.15
Yaralanmalı kaza – maddi hasarlı kaza	-0.01	-0.07	0.945	0.99
Maddi hasarlı kaza – ölümlü kaza	0.01	0.07	0.945	1.01
Maddi hasarlı kaza – yaralanmalı kaza	0.77	1.19	0.236	2.16
Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = 25-34 yaş arası			
	b	Z	P > z	e^b

Ölümlü kaza – yaralanmalı kaza	-0.94	-1.84	0.066	0.39
Ölümlü kaza – maddi hasarlı kaza	-0.92	-1.79	0.074	0.40
Yaralanmalı kaza – ölümlü kaza	0.94	1.84	0.066	2.57
Yaralanmalı kaza – maddi hasarlı kaza	0.03	0.26	0.796	1.03
Maddi hasarlı kaza – ölümlü kaza	0.92	1.79	0.074	2.50
Maddi hasarlı kaza – yaralanmalı kaza	-0.03	-0.26	0.796	0.98
Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = 35-44 yaş arası			
	B	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	-0.04	-0.09	0.929	0.96
Ölümlü kaza – maddi hasarlı kaza	-0.04	-0.08	0.935	0.96
Yaralanmalı kaza – ölümlü kaza	0.04	0.09	0.929	1.04
Yaralanmalı kaza – maddi hasarlı kaza	0.01	0.03	0.975	1.00
Maddi hasarlı kaza – ölümlü kaza	0.04	0.08	0.935	1.04
Maddi hasarlı kaza – yaralanmalı kaza	-0.01	0.03	0.975	1.00
Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = 45-54 yaş arası			
	b	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	0.66	1.40	0.162	1.93
Ölümlü kaza – maddi hasarlı kaza	0.70	1.49	0.137	2.02
Yaralanmalı kaza – ölümlü kaza	-0.66	-1.40	0.162	0.52
Yaralanmalı kaza – maddi hasarlı kaza	0.04	0.37	0.710	1.04
Maddi hasarlı kaza – ölümlü kaza	-0.70	-1.49	0.137	0.50
Maddi hasarlı kaza – yaralanmalı kaza	-0.04	-0.37	0.710	0.96

*0.01 anlam düzeyinde istatistiksel olarak anlamlı

**0.05 anlam düzeyinde istatistiksel olarak anlamlı

Tablo 4.26.'da karayolu trafik kazasına karışan araç sürücüsünün öğrenim durumuna göre bahis oranlarının ikili karşılaştırılması gösterilmiştir. Buna göre, kazaya karışan araç sürücüsünün öğrenim durumunun ilköğretim olması durumunda, yaralanmalı trafik kazalarının meydana gelme olasılığının (OR = 2.09, $p < 0.01$) maddi hasarlı trafik kazalarına göre yaklaşık 2 kat daha fazla olduğu görülmektedir. Benzer şekilde, kazaya

kariřan ara sürücüsünün öğrenim durumunun ortaöğretim olması durumunda, yaralanmalı trafik kazasının meydana gelme olasılığının (OR = 1.54, $p < 0.01$) maddi hasarlı trafik kazasına göre 1.5 kat daha fazla olduğu tespit edilmiştir.

Tablo 4. 26. Karayolu Trafik Kazasına Kariřan Ara Sürücüsünün Öğrenim Durumuna Göre Bahis Oranlarının İkili Karşılaştırılması

Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = İlköğretim			
	B	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	0.05	0.12	0.905	1.05
Ölümlü kaza – maddi hasarlı kaza	0.79	1.92	0.055	2.20
Yaralanmalı kaza – ölümlü kaza	-0.05	0.12	0.905	0.95
Yaralanmalı kaza – maddi hasarlı kaza	0.74	9.93	0.000*	2.09
Maddi hasarlı kaza – ölümlü kaza	-0.79	-1.92	0.055	0.45
Maddi hasarlı kaza – yaralanmalı kaza	-0.74	-9.93	0.000*	0.48
Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = Ortaöğretim			
	B	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	-0.50	-1.09	0.277	0.61
Ölümlü kaza – maddi hasarlı kaza	-0.07	-0.15	0.878	0.93
Yaralanmalı kaza – ölümlü kaza	0.50	1.09	0.277	1.65
Yaralanmalı kaza – maddi hasarlı kaza	0.43	5.98	0.000*	1.54
Maddi hasarlı kaza – ölümlü kaza	0.07	0.15	0.878	1.07
Maddi hasarlı kaza – yaralanmalı kaza	-0.43	-5.98	0.000*	0.65

*0.01 anlam düzeyinde istatistiksel olarak anlamlı

**0.05 anlam düzeyinde istatistiksel olarak anlamlı

Tablo 4.27’de karayolu trafik kazasına kariřan ara sürücüsüne ait sürücü belgesinin sınıfına göre bahis oranlarının ikili karşılaştırılması gösterilmiştir. Buna göre, kazaya kariřan ara sürücüsüne ait sürücü belgesinin E sınıfı olması durumunda, yaralanmalı trafik kazalarının meydana gelme olasılığının (OR = 1.42, $p < 0.01$) maddi hasarlı trafik kazalarına göre yaklaşık 1.5 kat daha fazla olduğu görölmektedir. Benzer şekilde, kazaya kariřan ara sürücüsüne ait sürücü belgesinin B sınıfı olması durumunda, yaralanmalı trafik kazasının meydana gelme olasılığının (OR = 12.05, $p < 0.01$) ölümlü trafik kazasına göre yaklaşık 12 kat daha fazla olduğu tespit edilmiştir. Ayrıca, sürücü

belgesinin B sınıfı olması durumunda, maddi hasarlı trafik kazasının meydana gelme olasılığının (OR = 16.14, $p < 0.01$) ölümlü trafik kazasına göre yaklaşık 16 kat, yaralanmalı trafik kazasına göre (OR = 1.34, $p < 0.01$) yaklaşık 1.5 kat daha fazla olduğu görülmektedir.

Tablo 4. 27. Karayolu Trafik Kazasına Karışan Araç Sürücüsünün Sürücü Belgesinin Sınıfına Göre Bahis Oranlarının İkili Karşılaştırılması

Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = E sınıfı sürücü belgesi			
	B	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	0.33	0.67	0.505	1.40
Ölümlü kaza – maddi hasarlı kaza	0.69	1.37	0.172	1.98
Yaralanmalı kaza – ölümlü kaza	-0.33	-0.67	0.505	0.72

Tablo 4.27. (Devam)

Yaralanmalı kaza – maddi hasarlı kaza	0.35	3.64	0.000*	1.42
Maddi hasarlı kaza – ölümlü kaza	-0.69	-1.37	0.172	0.50
Maddi hasarlı kaza – yaralanmalı kaza	-0.35	-3.64	0.000*	0.70
Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = B sınıfı sürücü belgesi			
	B	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	-2.49	-2.80	0.005*	0.08
Ölümlü kaza – maddi hasarlı kaza	-2.78	-3.17	0.002*	0.06
Yaralanmalı kaza – ölümlü kaza	2.49	2.80	0.005*	12.05
Yaralanmalı kaza – maddi hasarlı kaza	-0.29	-2.84	0.005*	0.75
Maddi hasarlı kaza – ölümlü kaza	2.78	3.17	0.002*	16.14
Maddi hasarlı kaza – yaralanmalı kaza	0.29	2.84	0.005*	1.34

*0.01 anlam düzeyinde istatistiksel olarak anlamlı

**0.05 anlam düzeyinde istatistiksel olarak anlamlı

Son olarak, Tablo 4.28.'de karayolu trafik kazasına karışan araç sürücüsünün ihlal ettiği trafik kurallarına göre bahis oranlarının ikili karşılaştırılması gösterilmiştir. Buna göre, kazaya karışan araç sürücüsünün ihlal ettiği trafik kuralının aşırı hız durumunda, yaralanmalı trafik kazalarının meydana gelme olasılığının (OR = 3.90, $p < 0.01$) maddi hasarlı trafik kazalarına göre yaklaşık 4 kat daha fazla olduğu görülmektedir. Benzer

şekilde, kazaya karışan araç sürücüsünün ihlal ettiği trafik kuralının “geçiş üstünlüğüne riayet etmemek” olması durumunda, yaralanmalı trafik kazasının meydana gelme olasılığının (OR = 1.69, $p < 0.01$) maddi hasarlı trafik kazasına göre yaklaşık 2 kat daha fazla olduğu tespit edilmiştir. Ayrıca, araç sürücüsüne ait sürücü kusurunun dikkatsizlik ya da diğer ihlaller olması durumunda, yaralanmalı trafik kazasının meydana gelme olasılığının (OR = 2.90, $p < 0.01$) maddi hasarlı trafik kazasına göre yaklaşık 3 kat daha fazla olduğu görülmektedir.

Tablo 4. 28. Karayolu Trafik Kazasına Karışan Araç Sürücüsünün İhlal Ettiği Trafik Kurallarına Göre Bahis Oranlarının İkili Karşılaştırılması

Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = Hız			
	B	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	0.35	0.32	0.748	1.41
Ölümlü kaza – maddi hasarlı kaza	1.71	1.58	0.114	5.52
Yaralanmalı kaza – ölümlü kaza	-0.35	-0.32	0.748	0.71
Yaralanmalı kaza – maddi hasarlı kaza	1.36	12.8	0.000*	3.90
Maddi hasarlı kaza – ölümlü kaza	-1.71	-1.58	0.114	0.18
Maddi hasarlı kaza – yaralanmalı kaza	-1.36	-12.8	0.000*	0.26
Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = Geçiş üstünlüğüne riayet etmemek			
	B	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	0.78	0.73	0.465	2.19
Ölümlü kaza – maddi hasarlı kaza	1.31	1.22	0.222	3.69
Yaralanmalı kaza – ölümlü kaza	-0.78	-0.73	0.465	0.46
Yaralanmalı kaza – maddi hasarlı kaza	0.52	5.16	0.000*	1.69
Maddi hasarlı kaza – ölümlü kaza	-1.31	-1.22	0.222	0.27
Maddi hasarlı kaza – yaralanmalı kaza	-0.52	-5.16	0.000*	0.59
Bağımlı Değişken Bahis Oranları karşılaştırması	Bağımsız Değişken = Dikkatsizlik ve Diğer ihlaller			
	B	Z	P > z	e^b
Ölümlü kaza – yaralanmalı kaza	0.95	0.90	0.379	2.60
Ölümlü kaza – maddi hasarlı kaza	2.02	1.86	0.062	7.55
Yaralanmalı kaza – ölümlü kaza	-0.95	-0.90	0.379	0.30
Yaralanmalı kaza – maddi hasarlı kaza	1.07	8.95	0.000*	2.90
Maddi hasarlı kaza – ölümlü kaza	-2.02	-1.86	0.062	0.13
Maddi hasarlı kaza – yaralanmalı kaza	-1.07	-8.95	0.000*	0.34

*0.01 anlam düzeyinde istatistiksel olarak anlamlı

**0.05 anlam düzeyinde istatistiksel olarak anlamlı

4.5.2.4.9. Kurulan Modele ait Marjinal Etkiler

Kurulan bir logit modeli için x_{ik} üzerindeki marjinal etkiler aşağıdaki eşitlikteki gibi ifade edilmektedir:

$$\frac{\partial F(x'_i\beta)}{\partial x_{ik}} = \Lambda_i(1 - \Lambda_i)\beta_k \quad (4.1)$$

Bu eşitlikte $\Lambda_i = \Lambda(x_i\beta)$ olarak tanımlanmaktadır (Powers ve Xie, 2000: 78). Araştırma kapsamında kurulan modeldeki bağımlı değişkene ait ölümlü kazalar için bağımsız değişkenlerin marjinal etkileri belirlemek için Stata 11.2 paket programında *mfx compute, predict(outcome(1))* komutu çalıştırılmıştır. Tablo 4.29.'da ölümlü kazalar için bağımsız değişkenlere ait marjinal etki değerleri sunulmuştur. Buna göre, karayolu trafik kazasının meydana geldiği yolda yaya kaldırımı ve aydınlatmanın olup olmaması, kazaya karışan araç sürücüsüne ait sürücü belgesinin B sınıfı olması ve diğer kusurlar değişkenleri istatistiksel olarak anlamlı bulunmuştur.

Tablo 4. 29. Ölümlü Kaza Türüne Ait Marjinal Etkiler

$y = P(\text{bağımlı değişken} = \text{ölümlü kaza}) = 0.0009$					
Bağımsız Değişken	dy/dx	Standart Hata	z	P > z	[% 95 G.A.]
Mevsim					
Kış	(-0.00482)	0.00038	-1.27	0.206	(-0.0013) - 0.0020
İlkbahar	0.00052	0.00037	0.14	0.889	(-0.0007) - 0.0008
Yaz	0.00041	0.00043	- 1.49	0.340	(-0.0004) - 0.0012
Saat					
06:00-11:59	(-0.00008)	0.00057	-0.15	0.885	(-0.0012) - 0.0010
12:00-17:59	(-0.00090)	0.00063	- 1.44	0.151	(-0.0021) - 0.0003
18:00-23:59	(-0.00044)	0.00038	-1.17	0.243	(-0.0012) - 0.0003
Haftaici/Haftasonu					
Haftaici	(-0.00025)	0.00032	-0.80	0.423	(-0.0008) - 0.0004
Hava durumu					
Açık	(-0.00219)	0.00136	-1.60	0.109	(-0.0048) - 0.0005
Yağmurlu/Bulutlu	(-0.00077)	0.00042	-1.82	0.068	(-0.0016) - 0.0001
Gün durumu					
Gündüz	(-0.00047)	0.00057	- 0.84	0.403	(-0.0016) - 0.0006
Yaya kaldırımı	(-0.00175)	0.00075	-2.33	0.020*	(-0.0032) - 0.0003
Yol Yüzeyi					
Kuru/Tozlu	0.00095	0.00057	1.66	0.097	(-0.0002) - 0.0021
Islak/Çamurlu/ Su ya da yağ birikintili	0.00090	0.00092	0.99	0.324	(-0.0009) - 0.0027
Aydınlatma	(-0.00164)	0.00068	-2.41	0.016*	(-0.0030) - (-0.0003)
Araç sayısı					
Tek araçlı	0.00090	0.00052	1.72	0.085	(-0.0013) - 0.0019

Zincirleme	0.00401	0.00263	1.53	0.127	(-0.0011) - 0.0092
Araç cinsi					
Kamyonet/Kamyon	0.00071	0.00048	1.48	0.139	(-0.0002) - 0.0017
Otobüs/Minibüs	0.00246	0.00159	1.54	0.122	(-0.0007) - 0.0056
Diğer	0.00132	0.00118	1.12	0.262	(-0.0010) - 0.0036
Yaş grubu					
16 yaşından küçük	0.00046	0.00115	0.41	0.685	(-0.0018) - 0.0027
16-24 yaş arası	(-0.00052)	0.00038	-1.37	0.169	(-0.0013) - 0.0002
25-34 yaş arası	(-0.00078)	0.00049	-1.60	0.110	(-0.0017) - 0.0002
35-44 yaş arası	(-0.00004)	0.00042	-0.09	0.932	(-0.0009) - 0.0008
45-54 yaş arası	0.00081	0.00077	1.05	0.292	(-0.0007) - 0.0023
Öğrenim durumu					
İlköğretim	0.00051	0.00045	1.15	0.250	(-0.0004) - 0.0014
Ortaöğretim	(-0.00018)	0.00039	-0.46	0.643	(-0.0009) - 0.0006
Sürücü belgesi					
E Sınıfı	0.00050	0.00045	1.10	0.271	(-0.0004) - 0.0014
B Sınıfı	(-0.00201)	0.00074	-2.74	0.006*	(-0.0030) - (-0.0003)
Kusur					
Hız	0.00130	0.00140	0.92	0.356	(-0.0015) - 0.0040
Geçiş üstünlüğü ihlali	0.00120	0.00134	0.90	0.371	(-0.0014) - 0.0038
Diğer kusurlar	0.00294	0.00358	0.82	0.020**	(-0.0041) - 0.0100

*0.01 anlam düzeyinde istatistiksel olarak anlamlı

**0.05 anlam düzeyinde istatistiksel olarak anlamlı

Yaralanmalı kazalar için bağımsız değişkenlere ait marjinal etki değerlerini belirlemek üzere Stata 11.2 paket programında *mfx compute, predict(outcome(2))* komutu çalıştırılmıştır. Tablo 4.30.'da söz konusu bağımsız değişkenlere ait marjinal etki değerleri sunulmuştur. Buna göre, mevsimin kış ya da ilkbahar olması, yolda yaya kaldırımının olup olmaması, yol yüzeyinin kuru/tozlu ya da ıslak/çamurlu olması, tek araçlı kazaların meydana gelmesi, araç cinsinin kamyon/kamyonet olması, sürücünün yaşının 16'dan küçük olması, tüm sürücü belgesi ve sürücü kusuru değişkenleri istatistiksel olarak anlamlı bulunmuştur.

Tablo 4. 30. Yaralanmalı Kaza Türüne Ait Marjinal Etkiler

$y = P(\text{bağımlı değişken} = \text{yaralanmalı kaza}) = 0.304$					
Bağımsız Değişken	dy/dx	Standart Hata	z	P > z	[% 95 G.A.]
Mevsim					
Kış	(-0.130)	0.013	-8.59	0.000*	(-0.16) - (-0.10)
İlkbahar	0.055	0.026	- 3.84	0.000*	(-0.08) - (-0.03)
Yaz	(-0.010)	0.015	-0.67	0.505	(-0.04) - 0.02
Saat					
06:00-11:59	0.022	0.031	0.73	0.465	(-0.04) - 0.08
12:00-17:59	0.004	0.028	0.14	0.887	(-0.05) - 0.06
18:00-23:59	0.017	0.025	0.69	0.491	(-0.03) - 0.07

Haftaiçi/Haftasonu					
Haftaiçi	(-0.005)	0.012	-0.40	0.689	(-0.03) - 0.02
Hava durumu					
Açık	(-0.050)	0.027	-1.91	0.057	(-0.10) - 0.001
Yağmurlu/Bulutlu	(-0.031)	0.025	-1.21	0.225	(-0.08) - 0.02
Gün durumu					
Gündüz	(-0.006)	0.019	- 0.33	0.743	(-0.05) - 0.03
Yaya kaldırımı	(-0.097)	0.012	-8.28	0.000*	(-0.12) - (-0.07)
Yol Yüzeyi					
Kuru/Tozlu	0.177	0.020	9.02	0.000*	0.14 - 0.22
Islak/Çamurlu/ Su ya da yağ birikintili	0.160	0.024	6.55	0.000*	(-0.11) - 0.21
Aydınlatma	(-0.013)	0.012	-1.11	0.267	(-0.04) - 0.01
Araç sayısı					
Tek araçlı	0.049	0.013	3.82	0.000*	0.02 - 0.07
Zincirleme	0.038	0.026	1.43	0.152	(-0.01) - 0.09
Araç cinsi					
Kamyonet/Kamyon	(-0.008)	0.013	-0.65	0.516	(-0.03) - 0.02
Otobüs/Minibüs	0.063	0.022	2.87	0.004*	0.02 - 0.11
Diğer	0.002	0.031	0.06	0.952	(-0.06) - 0.06
Yaş grubu					
16 yaşından küçük	0.668	0.017	40.03	0.000*	0.63 - 0.70
16-24 yaş arası	(-0.002)	0.025	-0.06	0.950	(-0.05) - 0.05
25-34 yaş arası	0.006	0.024	0.27	0.787	(-0.04) - 0.05
35-44 yaş arası	0.001	0.021	0.03	0.975	(-0.04) - 0.04
45-54 yaş arası	0.009	0.022	0.36	0.719	(-0.04) - 0.06
Öğrenim durumu					
İlköğretim	0.158	0.016	9.89	0.000*	0.13-0.19
Ortaöğretim	0.093	0.016	5.88	0.000*	0.06-0.12
Sürücü belgesi					
E Sınıfı	0.073	0.020	3.69	0.000*	0.03-0.11
B Sınıfı	(-0.060)	0.021	-2.88	0.004*	(-0.10)- (-0.02)
Kusur					
Hız	0.296	0.023	12.99	0.000*	0.25-0.34
Geçiş üstünlüğü ihlali	0.112	0.022	5.09	0.000*	0.07-0.16
Diğer kusurlar	0.249	0.029	8.60	0.000*	(-0.12) – (-0.07)

*0.01 anlam düzeyinde istatistiksel olarak anlamlı

**0.05 anlam düzeyinde istatistiksel olarak anlamlı

Son olarak maddi hasarlı kazalar için modeldeki bağımsız değişkenlere ait marjinal etkileri belirlemek üzere Stata 11.2 paket programında *mfx compute, predict(outcome(3))* komutu çalıştırılmıştır. Tablo 4.31’de maddi kazalar için bağımsız değişkenlere ait marjinal etki değerleri sunulmuştur. Buna göre, mevsimin kış ya da ilkbahar olması, hava durumunun açık olması, yolda yaya kaldırımının olup olmaması, yol yüzeyinin kuru/tozlu ya da ıslak/çamurlu olması, araç cinsinin otobüs/minibüs olması, sürücünün yaşının 16’dan küçük olması, tüm öğrenim durumu, sürücü belgesi ve sürücü kusuru değişkenleri istatistiksel olarak anlamlı bulunmuştur.

Tablo 4. 31. Maddi Hasarlı Kaza Türüne Ait Marjinal Etkiler

$y = P(\text{bağımlı değişken} = \text{maddi hasarlı kaza}) = 0.703$					
Bağımsız Değişken	dy/dx	Standart Hata	Z	P > z	[% 95 G.A.]
Mevsim					
Kış	0.13	0.015	8.61	0.000*	0.10 - 0.16
İlkbahar	0.05	0.014	3.83	0.000*	0.03 - (-0.08)
Yaz	0.01	0.014	0.64	0.524	(-0.02) - 0.04
Saat					
06:00-11:59	(-0.02)	0.031	0.42	0.467	(-0.08) - 0.04
12:00-17:59	(-0.003)	0.028	-0.73	0.912	(-0.06) - 0.05
18:00-23:59	(-0.02)	0.025	-0.11	0.502	(-0.06) - 0.03
Haftaiçi/Haftasonu					
Haftaiçi	0.066	0.012	0.64	0.524	(-0.02) - 0.03
Hava durumu					
Açık	0.05	0.027	1.98	0.047**	0.001 - 0.11
Yağmurlu/Bulutlu	0.03	0.025	1.24	0.214	(-0.02) - 0.08
Gün durumu					
Gündüz	0.007	0.020	0.35	0.725	(-0.03) - 0.05
Yaya kaldırımı	0.10	0.012	8.43	0.000*	0.08 - 0.12
Yol Yüzeyi					
Kuru/Tozlu	(-0.18)	0.020	-9.07	0.000*	(-0.22) - (-0.14)
Islak/Çamurlu/ Su ya da yağ birikintili	(-0.16)	0.024	-6.59	0.000*	(-0.21) - (-0.11)
Aydınlatma	0.01	0.012	1.25	0.211	(-0.01) - 0.04
Araç sayısı					
Tek araçlı	(-0.05)	0.013	-3.89	0.000*	(-0.08) - (-0.02)
Zincirleme	(-0.04)	0.026	-1.58	0.115	(-0.09) - 0.01
Araç cinsi					
Kamyonet/Kamyon	0.01	0.013	0.59	0.554	(-0.02) - 0.03
Otobüs/Minibüs	(-0.07)	0.022	-2.98	0.003*	(-0.11) - (-0.02)
Diğer	(-0.003)	0.032	-0.10	0.918	(-0.06) - 0.06
Yaş grubu					
16 yaşından küçük	(-0.67)	0.017	-40.16	0.000*	(-0.70) - (-0.64)
16-24 yaş arası	0.002	0.024	0.08	0.933	(-0.05) - 0.05
25-34 yaş arası	(-0.005)	0.021	-0.23	0.817	(-0.05) - 0.04
35-44 yaş arası	(-0.001)	0.022	-0.03	0.976	(-0.04) - 0.04
45-54 yaş arası	(-0.01)	0.024	-0.39	0.695	(-0.06) - 0.04
Öğrenim durumu					
İlköğretim	(-0.16)	0.016	-9.92	0.000*	(-0.19) - (-0.13)
Ortaöğretim	(-0.10)	0.016	-5.87	0.000*	(-0.12) - (-0.06)
Sürücü belgesi					
E Sınıfı	(-0.08)	0.020	-3.71	0.000*	(-0.11) - (-0.03)
B Sınıfı	0.06	0.021	2.97	0.003*	0.02 - 0.10
Kusur					
Hız	(-0.30)	0.023	-13.05	0.000*	(-0.34) - (-0.25)

Geçiş üstünlüğü ihlali	(-0.11)	0.022	-5.14	0.000*	(-0.16) - (-0.07)
Diğer kusurlar	(-0.25)	0.029	-8.73	0.000*	(-0.31) - 0.12

*0.01 anlam düzeyinde istatistiksel olarak anlamlı

**0.05 anlam düzeyinde istatistiksel olarak anlamlı

Çok durumlu modeller doğrusal olmayan modeller olduğundan x_k gibi bir değişkenin P_{ij} üzerindeki etkisi sabit olmamaktadır. Bu sebeple, marjinal etkilerin çok durumlu modeller için çok kullanışlı olduğu söylenememektedir. İki durumlu modeller için x_k değeri arttığında marjinal etkiler katsayılarında pozitif bir artış sağlamakta, ancak çok durumlu modellerde, $P(y_i = j)$ olasılığındaki değişim, β_{jk} parametresinde aynı işaretle bir artışa sebep olamayabilmektedir. Örneğin, $P(y_i = 2)$ olasılığındaki bir değişim β_{jk} parametresi pozitif değer olsa bile, x_k değeri azalabilmektedir. Böylece, bazı ekonometrik çalışmalarda çok durumlu modellerin marjinal etkiler vasıtasıyla yorumlanmasına dikkat çekilse bile, genellikle çok durumlu modellerin göreceli risk ve bahis oranları ile değerlendirilmesi tavsiye edilmektedir (Powers ve Xie, 2000: 231).

SONUÇ

Bu çalışmada karayolu trafik kazalarını etkileyen risk faktörlerinin istatistiksel bir yöntem kullanılarak belirlenmesi amaçlanmıştır. Bağımlı değişken olarak seçilen kaza türü ölümlü, yaralanmalı ve maddi hasarlı kazalar olmak üzere üç farklı kaza türünü barındırdığı için sözü edilen risk faktörlerinin tespit edilmesi sürecinde üç ya da daha fazla kategoriye sahip bağımlı değişken ile birçok bağımsız değişken arasındaki ilişkiyi ortaya koyan çok durumlu logit modeli kullanılmıştır. Çalışma kapsamında Erzurum ve Kars illerinde 2008-2012 yılları arasında meydana gelen 9162 karayolu trafik kazası tespit tutanağı incelenmiş ve elde edilen veriler bilgisayar ortamına aktarılmıştır. Çoklu bağlantı sorununa sebep olan ve eksik bilgiye sahip bağımsız değişkenler model dışında tutulduktan sonra kurulacak model için karayolu trafik kazalarını etkileyeceği düşünülen 14 farklı risk faktörüne ait 31 adet gölge değişken tanımlanmıştır. Sözü edilen risk faktörlerine ait kategorilerin belirlenmesinde literatürdeki ilgili çalışmalar dikkate alınmıştır. Bağımlı değişkenler ve bağımsız değişkenlerin belirlenerek kurulan çok durumlu logit modelinde maddi hasarlı kazalar referans kategori olarak kabul edilmiştir.

Erzurum ve Kars illerinde 2008-2012 yılları arasında meydana gelen karayolu trafik kazalarını etkileyen risk faktörlerinin tespit edilmesi adına kurulan çok durumlu logit modelinin öncelikle ilişkisiz alternatiflerin bağımsızlığı varsayımını sağlayıp sağlamadığı test edilmiştir. Bu varsayımın kontrolü için Hausman testi kullanılmış olup sonuç olarak bağımlı değişkene ait üç kaza türü kategorisinin de alternatiflerin bağımsızlığını ve ilişkisizliğini savunan sıfır hipotezini kabul ettiği ve bu sayede ilişkisiz alternatiflerin bağımsızlığı varsayımını sağladığı tespit edilmiştir. Kurulan model çoklu bağlantı sorunu testini de sağlamaktadır. Kurulan model uyum iyiliği testlerine tabi tutulduğunda ise düşük Akaike bilgi kriteri ve çok yüksek negatif Bayes bilgi kriteri değerlerine sahip modelin uyum iyiliği testlerini de karşıladığı görülmüştür. Bir sonraki aşamada ise birçok bağımsız değişkenin istatistiksel olarak anlamlı olması ile olabilirlik oranı ve Wald testlerini de geçen çok durumlu logit modele ait sonuçların göreceli risk oranı değerleri vasıtasıyla değerlendirilmesi yoluna gidilmiştir.

Kurulan modelin log-olabilirlik değeri -5035.374 olup modele ait pseudo- χ^2 değeri istatistiksel olarak anlamlıdır. Çok durumlu logit modelinde elde edilen göreceli

risk oranları iki durumlu logit modelinde sıklıkla kullanılan bahis oranı değerlerinin çok durumlu modellerdeki karşılığı olarak görülmekte ve görel risk oranlarının yorumlanması da bahis oranlarında yapılan yorumlara benzemektedir. Modelde maddi hasarlı kazalar referans kategori olarak değerlendirildiğinde istatistiksel olarak anlamlı bulunan bağımsız değişkenlerin görel risk oranları değerlendirilmesi yapılmıştır. Kurulan modeldeki istatistiksel olarak anlamlı bulunan bağımsız değişkenlere ait görel risk oranları değerleri yorumlandığında, hem ölümlü hem de yaralanmalı kazalar için karayolu trafik kazalarını en çok etkileyen risk faktörünün sürücünün yaşının 16'dan küçük olması olduğu görülmüştür. Bu bağımsız değişken belirlenen yıllar arasında Erzurum ve Kars illerinde ölümlü ve yaralanmalı kazaların meydana gelme olasılığını önemli derecede artırmaktadır. Karayolu trafik kazalarını en çok etkileyen bir diğer faktör karayolu trafik kazasının zincirleme kaza olmasıdır. Zincirleme trafik kazaları ölümlü trafik kazalarının meydana gelme olasılığını yaklaşık 6 kat artırmaktadır. Bunun yanında karayolu trafik kazasına karışan aracın otobüs ya da minibüs olması ölümlü kazalarının meydana gelme olasılığını yaklaşık 5 kat artırmaktadır. Karayolu trafik kazasının meydana geldiği yolun kuru ya da tozlu olması da ölümlü trafik kazası meydana gelme olasılığını yaklaşık 4 kat artırmaktadır. Belirlenen yıllar arasında sürücünün karayolu trafik kazasının meydana geldiği sırada aşırı hızlı olması yaralanma trafik kazalarının meydana gelme olasılığını yaklaşık 4 kat, dikkatsizlik gibi diğer kusurlar ise 3 kat artırmaktadır. Zincirleme kazalar gibi tek araçlı kazalar da ölümlü kazaların meydana gelme olasılığını yaklaşık 3 kat artırmaktadır. Bu sonucun daha önce tartışılan dünya geneli trafik kazaları sonuçlarında olduğu gibi daha çok yaya ölümlerini işaret ettiği düşünülmektedir. Yaralanmalı karayolu trafik kazalarının meydana gelme olasılığını yaklaşık 2 kat artıran yolun kuru/tozlu ya da ıslak/çamurlu/su birikintili olması, sürücünün öğrenim durumunun ilköğretim ya ortaöğretim olması, sürücünün geçiş üstünlüğü kuralını ihlal etmesi gibi değişkenler karayolu trafik kazalarını etkileyen diğer risk faktörleridir.

Kurulan çok durumlu modelin sonuçlarına göre görel risk oranları birbirine yakın olsa da yaralanmalı karayolu trafik kazasının meydana gelme olasılığını en çok azaltan faktör kazanın meydana geldiği günde havanın açık olmasıdır. Bu faktörü kazanın meydana geldiği mevsimin ilkbahar olması ve araç sürücüsünün B sınıfı sürücü belgesine sahip olması gibi değişkenler izlemektedir. Dikkat edilmesi gereken bir diğer

sonuç ise kazanın meydana geldiği yolda yaya kaldırımının olmasının yaralanmalı karayolu trafik kazasını azaltmasıdır. Yine kış mevsiminde yaralanmalı kazaların görülme olasılığı azalmaktadır. Öte yandan, karayolunun meydana geldiği yolda aydınlatmanın ve yaya kaldırımının olması ölümlü trafik kazalarının meydana gelme olasılığını azaltmaktadır. Aynı zamanda havanın açık olması ve sürücü belgesinin B sınıfı olması ölümlü trafik kazalarının meydana gelme olasılığını azaltmaktadır. Genel bir değerlendirme yapıldığında, karayolu trafik kazalarını etkileyen risk faktörlerinin belirli bir grup içerisinde kümelenmediği; araç, yol, hava durumu, sürücü gibi birçok risk faktörünün karayolu trafik kazalarında etkili olduğu söylenebilmektedir.

Kurulan modelde ikili etkilerin görülmesi adına referans kategorilerinin değiştirilmesi yoluna gidilmiş ve elde edilen sonuçlar bahis oranı değerleri vasıtasıyla yorumlanmıştır. Buna göre, bahis oranı değeri en yüksek olan risk faktörü yine sürücünün yaşının 16'dan küçük olmasıdır. Sürücünün yaşının 16'dan küçük olması durumunda yaralanmalı ve ölümlü kazaların meydana gelme olasılığı maddi hasarlı kazalara göre daha fazla olduğu görülmektedir. Referans kategorilerinin değiştirilmesi sonucu karayolu trafik kazalarını en çok etkileyen bir diğer faktör araç sürücüsüne ait sürücü belgesinin B sınıfı olmasıdır. Sürücü belgesinin B sınıfı olması durumunda, maddi hasarlı ve yaralanmalı trafik kazalarının ölümlü kazalara göre meydana gelme olasılığı daha fazla olmaktadır. Zincirleme kaza olması durumunda ölümlü kazaların meydana gelme olasılığı maddi hasarlı ve yaralanmalı kazalara göre daha fazla olmaktadır. Öte yandan, karayolu trafik kazasının meydana geldiği günde havanın açık olması ölümlü trafik kazalarının meydana gelme olasılığı azaltmaktadır. Havanın açık olması durumunda maddi hasarlı kazaların meydana gelme olasılığı daha fazla olmaktadır. Benzer şekilde kazanın meydana geldiği yolda aydınlatmanın olması durumunda, ölümlü trafik kazalarının meydana gelme olasılığının yaralanmalı ve maddi hasarlı kazalara göre azaldığı tespit edilmiştir. Bu durumda yine maddi hasarlı kazaların meydana gelme olasılığı daha fazladır. Kazaya karışan aracın otobüs ya da minibüs olması durumunda, ölümlü kazaların meydana gelme olasılığının maddi hasarlı kazalara göre daha fazla olduğu görülmüştür. Yolcularla birlikte kazaya daha fazla kişinin karıştığı düşünüldüğünde bu sonucun beklenen bir durum olduğu söylenebilmektedir. Kazanın meydana geldiği yolun kuru ya da tozlu olması durumunda ölümlü kazaların meydana gelme olasılığının maddi hasarlı kazalara göre; araç sürücüsünün aşırı hız

yapması durumunda yaralanmalı kazaların meydana gelme olasılığının maddi hasarlı kazalara göre daha fazla olduğu belirlenmiştir. Kazanın meydana geldiği yolda yaya kaldırımının olmasının ölümlü kazaların meydana gelme olasılığını düşürdüğü tespit edilmiştir. Bu sonuç da savunmasız yol kullanıcılarının karayolu trafik kazalarında karşı karşıya kaldığı olumsuz etkilerin önlenmesi adına dikkat çekicidir. Son olarak araç sürücüsünün öğrenim durumunun ilköğretim olması, yaralanmalı trafik kazalarının meydana gelme olasılığını arttırmaktadır.

Son olarak kurulan modelin marjinal etki değerlerine bakılmış ve buna göre mevsimin kış ya da ilkbahar olması, hava durumunun açık olması, yolda yaya kaldırımının olup olmaması, yol yüzeyinin kuru/tozlu ya da ıslak/çamurlu olması, araç cinsinin otobüs/minibüs olması, sürücünün yaşının 16'dan küçük olması, tüm öğrenim durumu, sürücü belgesi ve sürücü kusuru istatistiksel olarak anlamlı bulunmuştur. Ancak çok durumlu logit modeli doğrusal olmayan bir model olduğu için marjinal etkilerin yerine göreceli risk ve bahis oranlarının yorumlanmasının daha çarpıcı olacağı düşünülmüştür.

Araştırma kapsamında yapılan çok durumlu logit modeli analizi sonucu elde edilen bulgular ışığında ilk olarak araç sürücüsünün yaşının küçük olması durumunda ileriki yıllarda atılması gereken adımlar ve izlenmesi gereken politikalar düşünülmelidir. Özellikle ölümlü kazaları önemli derecede artıran bu risk faktörünün etkisinin azaltılabilmesi adına karar vericilere, politika uygulayıcılara ve ebeveynlere bazı görevler düşmektedir. Reşit olmayan gençlerin araç başına geçmelerinin ilk olarak ebeveynlerce önlenmesi karayolu trafik kazalarının olumsuz etkilerin ve özellikle ölümlü kazaların azaltılması hususunda faydalı olacağı düşünülmektedir. Benzer şekilde, caydırıcı kurallar ve yasalarla birlikte gençlerin araç sürücüsü olma şartlarını yerine getirmeden direksiyon başına geçmelerinin önlenmesi bir diğer trafik güvenliği politikası olabilecektir. Bu politikanın sadece genç bireyler için değil de tüm sürücü adaylarına dikkatle uygulanması sürücü belgesi faktörünün karayolu trafik kazalarına etkisinin azaltılması yolunda önemli bir aşama olduğu görülebilecektir. Bu bağlamda, hem sürücü hem de diğer karayolu kullanıcıları olmak üzere her açıdan daha eğitilmiş bireylerin trafiğe çıkmasının sağlanması da karayolu trafik kazalarının olumsuz etkilerini azaltabilecektir. Sürücülerin aşırı hız ve geçiş üstünlüğü gibi sıklıkla görmezden geldiği trafik kurallarının ihlalinin önlenmesi de geliştirilmesi gereken bu

politikalar kapsamında değerdendirilebilecek ve zincirleme kazaların azaltılmasında da faydalı olabilecektir.

Araştırma sonuçlarına göre; yaya kaldırımı, aydınlatma ve diğer yol değışkenlerinin karayolu trafik kazalarının azaltılmasında ve trafik güvenliğinin sağlanmasında olumlu etkilerinin olduğu görüldüğünden karayollarının yoğun trafik akışlarında bile yeterli yol altyapısına ve işaretlerine sahip olmaları karayollarının güvenliğini artıracak ve trafik kazalarının olumsuz sonuçlarını azaltabilecektir. Özellikle karayolu trafiğinin seyretmesinin zorlaştığı bölgelerde yol gereksinimlerinin karşılanması savunmasız yol kullanıcılarının karşılaşacağı zararları azaltma adına yararlı olabilecektir. Olumsuz hava durumlarında dahi karayollarının yapısının daha dayanıklı hale getirilmesi sürücü kusurlarında ifade edildiği zincirleme kazalarının azaltılması hususunda da yararlı olabilecektir.

Dikkat çekilmesi gereken bir diğer önemli nokta ise karayolu trafik kazalarına ait tespitlerin uygun şekilde rapor edilmemesidir. Karayolu trafik kazalarının sebeplerinin belirlenmesi çözümlere ait atılacak adımlar için önemli bir işaret olmaktadır. Ancak, özellikle emniyet kemeri ve alkol durumu gibi karayolu trafik kazaları için önemli risk faktörü olabilecek unsurların kontrolü ve raporlanmasının daha etkin yapılabilmesi bu faktörlerin etkilerinin de tespit edilmesini kolaylaştırılacaktır. Ayrıca, mevcut kurallara göre bir kazanın ölümlü trafik kazası olarak raporlanması kazaya karışan bireylerin kaza anında yaşamını yitirmiş olması ile değerdendirilmektedir. Hastanede geçen süreç ve bireyin sağlık durumu ile ilgili süreç takip edilmediği için ölümlü trafik kaza sayılarına ait doğru bilgiler elde edilememektedir. Dünyada birçok ülkede geçerli olan hasta takibinin 30 gün süreyle yapılması uygulamasının Türkiye’de de geçerlilik kazanması ve bu noktada sağlık kurumları ile sürekli irtibatın sağlanması daha etkin karayolu politikalarının izlenmesi adına faydalı olabilecektir.

KAYNAKLAR

- Abbe, E., Bierlaire, M. ve Toledo, T. (2007). "Normalization and Correlation of Cross-nested Logit Models". *Transportation Research Part B*, 41(7), 795-808.
- Abdel-Aty, M. A., Chen, C. L., & Schott, J. R. (1998). An Assessment of the Effect of Driver Age on Traffic Accident Involvement Using Log-linear Models. *Accident Analysis & Prevention*, 30(6), 851-861.
- Akaike, H. (1973). "Information Theory and an Extension of the Maximum Likelihood Principle". B. N. Petrov ve F. Csaki (Ed.). *Second International Symposium on Information Theory* (ss. 267-281). Budapest: Akademiai Kiado.
- Ashford, J. R. ve Sowden, R. R. (1970). "Multi-Variate Probit Analysis". *Biometrics*, 535-546.
- Agresti, A. (1996). *An Introduction to Categorical Data Analysis*, New York: John Wiley & Sons. Inc.
- , (1999). "Modelling Ordered Categorical Data: Recent Advances and Future Challenges". *Statistics in Medicine*, 18(17-18), 2191-2207.
- , (2002). *Categorical Data Analysis*, New York: John Wiley & Sons, Inc.
- , (2010). *Analysis of Ordinal Categorical Data*, New York: John Wiley & Sons. Inc.
- Agresti, A., Booth J. G., Hobert, J. P., Cappel, B. (2000). "Random-Effects Modeling of Categorical Response Data". *Sociological Methodology*, 30(1), 27-80.
- Agresti, A. ve Finlay, B. (2009). *Statistical Methods for the Social Sciences*, New Jersey: Prentice Hall, Inc.
- Aitchison, J. ve Silvey, S. (1957). "The Generalization of Probit Analysis to the Case of Multiple Responses". *Biometrika*, 44(1), 131-140.

- Aitchison, J. ve Bennet, J. (1970). "Polychotomous Quantal Response by Maximum Indicant". *Biometrika*, 57(2), 253-262.
- Akdeniz, F. (2013). *Olasılık ve İstatistik*, Ankara: Nobel Kitabevi.
- Akgül, A. ve Çevik, O. (2003). *İstatistiksel Analiz Teknikleri: SPSS'te İşletme Yönetimi Uygulamaları*, Ankara: Emek Ofset Ltd. Şti.
- Akın, F. (2002). *Kalitatif Tercih Modelleri Analizi*, Bursa: Ekin Kitabevi.
- Akkaya, Ş. ve Pazarlıoğlu, M. V. (1998). *Ekonometri II*, İstanbul: Erkam Matbaacılık.
- Albayrak, A. S. (2008). "Lojistik Regresyon Analizi". Şeref Kalaycı (Ed.). *SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri* (ss. 273-298). Ankara: Asil Yayın Dağıtım Ltd. Şti.
- Aldrich, J. H. ve Nelson, F. D. (1984). *Linear Probability, Logit, and Probit Models*, California: SAGE Publications, Inc.
- Al-Ghamdi, A. S. (2002). "Using Logistic Regression to Estimate the Influence of Accident Factors on Accident Severity". *Accident Analysis and Prevention*, 34(6), 729-741.
- Amemiya, T. (1974). "Bivariate Probit Analysis: Minimum Chi-Square Methods". *Journal of the American Statistical Association*, 69(348), 940-944.
- , (1975). "Qualitative Response Models". *Annals of Economic and Social Measurement*, 4(3), 363-372.
- , (1977). "Some Theorems in the Linear Probability Model". *International Economic Review*, 18(3), 645-650.
- , (1981). "Qualitative Response Models: A Survey". *Journal of Economic Literature*, 19(4), 1483-1536.
- , (1985). *Advanced Econometrics*, Massachusetts: Harvard University Press.
- , (1994). *Introduction to Statistics and Econometrics*, Massachusetts: Harvard University Press.

- Anderson, J. A. ve Philips, P. R. (1981). "Regression, Discrimination and Measurement Models for Ordered Categorical Variables". *Journal of the Royal Statistical Society Series C (Applied Statistics)*, 30(1), 22-31.
- Ananth, C. V. ve Kleinbaum, D. G. (1997). "Regression Models for Ordinal Responses: A Review of Methods and Applications". *International Journal of Epidemiology*, 26(6), 1323-1333.
- Arıcıgil Çilan, Ç. (2009). *Sosyal Bilimlerde Kategorik Verilerle İlişki Analizi*, Ankara: Pegem Akademi.
- Arrow, K. J. (1951). *Social Choice and Individual Values*, New York: John & Wiley Sons, Inc.
- Awadzi, K. D., Classen, S., Hall, A., Duncan, R. P. ve Garvan, C. W. (2008). "Predictors of Injury Among Younger and Older Adults in Fatal Motor Vehicle Crashes". *Accident Analysis and Prevention*, 40(6), 1804-1810.
- Baltagi, B. H. (1997). *Econometrics*, New York: Springer.
- Baltas, G. ve Doyle, P. (2001). "Random Utility Models in Marketing Research: A Survey". *Journal of Business Research*, 51(2), 115-125.
- Başar, A. ve Oktay, E. (2010). *Uygulamalı İstatistik I*, Erzurum: Erzurum Kültür ve Eğitim Vakfı Yayınevi.
- Ben-Akiva, M. (1973). *Structure of Passenger Travel Demand Models*. (Unpublished Ph.D. Dissertation). Massachusetts: Massachusetts Institute of Technology Department of Civil Engineering.
- Ben-Akiva, M. ve Bierlaire, M. (1999). "Discrete Choice Methods and Their Applications to Short Term Travel Decisions". Randolph W. Hall (Ed.). *Handbook of Transportation Science* (ss. 5-33). New York: Springer Science + Business Media.
- Ben-Akiva, M., McFadden, D., Abe, M., Böckenholt, U., Bolduc, D., Gopinath, D., Morikawa, T., Ramaswamy, V., Rao, V., Revelt, D. ve Steinberg,

- D. (1997). "Modeling Methods for Discrete Choice Analysis". *Marketing Letters*, 8(3), 273-286.
- Ben-Akiva, M., McFadden, D., Gärling, T., Gopinath, D., Walker, J., Bolduc, D., Börsch-Supan, A., Delquié, P., Larichev, O., Morikawa, T., Polydoropoulou, A. ve Rao, V. (1999). "Extended Framework for Modeling Choice Behavior". *Marketing Letters*, 10(3), 187-203.
- Berkson, J. (1955). "Maximum Likelihood and Minimum Chi-Square Estimates of the Logistic Function". *Journal of the American Statistical Association*, 50(269), 130-162.
- Bishop, Y. M., Fienberg, S. E. ve Holland, P. W. (1975). *Discrete Multivariate Analysis: Theory and Practice*, Massachusetts: The MIT Press.
- Bock, D. (1975). *Multivariate Methods in Behavioral Research*, New York: McGraw-Hill, Inc.
- Boes, S. ve Winkelmann, R. (2006). "Ordered Response Models". *Allgemeines Statistisches Archiv*, 90(1), 167-181.
- Borooah, V. K. (2002). *Logit and Probit: Ordered and Multinomial Models*, Thousand Oaks: SAGE Publications, Inc.
- Brown, B. W. ve Walker, M. B. (1989). "The Random Utility Hypothesis and Inference in Demand Systems". *Econometrica*, 57(4), 815-829.
- Brownstone, D. ve Small, K. A. (1989). "Efficient Estimation of Nested Logit Models". *Journal of Business & Economic Statistics*, 7(1), 67-74.
- Cheng, S. ve Long, J. S. (2007). Testing for IIA in the Multinomial Logit Model". *Sociological Methods & Research*, 35(4), 583-600.
- Chipman, M. L. (1995). "Risk Factors for Injury: Similarities and Differences for Traffic Crashes and Other Causes". *Accident Analysis and Prevention*, 27(5), 699-706.
- Chow, G. C. (1983). *Econometrics*, New York: McGraw-Hill Inc.

- Cox, D. R. (1958). "The Regression Analysis of Binary Sequences". *Journal of the Royal Statistical Society Series B (Methodological)*, 20(2), 215-242.
- .. (1966). "Some Procedures Connected with the Logistic Qualitative Response Curve". F. N. David (Ed.). *Research Papers in Statistics* (ss. 55-71). London: John Wiley & Sons, Inc.
- Cramer, J. S. (1986). *Econometric Applications of Maximum Likelihood Methods*, Cambridge: Cambridge University Press.
- Cressie, N. ve Read, T. R. C. (1984). "Multinomial Goodness-of-Fit Tests". *Journal of the Royal Statistical Society Series B (Methodological)*, 46(3), 440-464.
- Çokluk, Ö., Şekercioğlu, G. ve Büyüköztürk, Ş. (2010). *Sosyal Bilimler için Çok Değişkenli İstatistik: SPSS ve LISREL Uygulamaları*, Ankara: Pegem Akademi.
- Daganzo, C. (1979). *Multinomial Probit: The Theory and Its Application to Demand Forecasting*, London: Academic Press Ltd.
- Davidson, R. ve MacKinnon, J. G. (1999). *Econometric Theory and Methods*, Oxford: Oxford University Press Inc.
- De Boni, R., Cruz, O. G., Weber, E., Hasenack, H., Lucatelli, L., Duarte, P., Gracie, R., Pechansky, F. ve Bastos, F. (2013). "Traffic Crashes and Alcohol Outlets in a Brazilian State Capital". *Traffic Injury Prevention*, 14(1), 86-91.
- DeMaris, A. (2004). *Regression with Social Data: Modeling Continuous and Limited Response Variables*, New Jersey: John Wiley & Sons, Inc.
- Depaire, B., Wets, G. ve Vanhoof, K. (2008). "Traffic Accident Segmentation by Means of Latent Class Clustering". *Accident Analysis and Prevention*, 40(4), 1257-1266.
- Domencich, T. A. ve McFadden, D. (1975). *Urban Travel Demand: A Behavioral Analysis*, Amsterdam: North-Holland Publishing Company.

- Dougherty, C. (2007). *Introduction to Econometrics*, New York: Oxford University Press Inc.
- Drucker, C., Gerberich, S. G., Manser, M. P., Alexander, B. H., Church, T. R., Ryan, A. D. ve Becic, E. (2013). "Factors Associated with Civilian Drivers Involved in Crashes with Emergency Vehicles". *Accident Analysis and Prevention*, 55(1), 116-123.
- Dünya Sağlık Örgütü (2013a). *Global Status Report on Road Safety 2013: Supporting a Decade of Action*. Erişim tarihi: 19 Haziran 2013, http://www.who.int/violence_injury_prevention/road_safety_status/2013/en/index.html
- Eluru, N., Bhat, C. R. ve Hensher, D. A. (2008). "A Mixed Generalized Ordered Response Model for Examining Pedestrian and Bicyclist Injury Severity Level in Traffic Crashes". *Accident Analysis and Prevention*, 40(3), 1033-1054.
- Elvik, R., Fridstrøm, L., Kaminska, J. ve Myer, S. F. (2013). "Effects on Accidents of Changes in the Use of Studded Tyres in Major Cities in Norway: A Long-term Investigation". *Accident Analysis and Prevention*, 54, 15-25.
- Emniyet Genel Müdürlüğü Trafik Hizmetleri Başkanlığı Genel Kaza İstatistikleri Ankara: Emniyet Genel Müdürlüğü (Erişim tarihi: 20 Haziran 2013). <http://www.trafik.gov.tr/Sayfalar/Istatistikler/Genel-Kaza.aspx>
- Fagerland, M. W., Hosmer, D. W. ve Bofin, A. M. (2008). "Multinomial Goodness-of-fit Tests for Logistic Regression Models". *Statistics in Medicine*, 27(21), 4238-4253.
- Finney, D. J. (1971). *Probit Analysis*, Cambridge: Cambridge University Press.
- Fisher, R. A. (1969). *Statistical Methods for Research Workers*, Edinburgh: Oliver and Boyd Publishing.
- Fox, J. (1997). *Applied Regression Analysis, Linear Models, and Related Methods*, California: SAGE Publications, Inc.

- Garrow, L. A. (2010). *Discrete Choice Modelling and Air Travel Demand*, Burlington: Ashgate Publishing Company.
- Genceli, M. (1987). “Nitel Bağımlı Değişken için Doğrusal Olasılık Modeli (IVb)”. *İstanbul Üniversitesi İktisat Fakültesi Mecmuası*, 45(1-4), 213-244.
- Gensch, D. H. ve Ghose, S. (1997). “Differences in Independence of Irrelevant Alternatives at Individual vs Aggregate Levels, and at Single Pair vs Full Choice Set”. *Omega, International Journal of Management Science*, 25(2), 201-214.
- Geweke, J., Keane, M. ve Runkle, D. (1994). “Alternative Computational Approaches to Inference in the Multinomial Probit Model”. *The Review of Economics and Statistics*, 76(4), 609-632.
- Greene, W. H. (2008). *Econometric Analysis*, New Jersey: Pearson Education Inc.
- Greene, W. H. ve Seaks, T. G. (1991). “The Restricted Least Squares Estimator: A Pedagogical Note”. *The Review of Economics and Statistics*, 73(3), 563-567.
- Goldberger, A. S. (1991). *A Course in Econometrics*, Massachusetts: Harvard University Press.
- Gujarati, D. N. (2011). *Temel Ekonometri*, (Çev.: Ümit Şenesen ve Gülay Günlük Şenesen), İstanbul: Literatür Yayıncılık.
- Güriş, S. ve Çağlayan, E. (2010). *Ekonometri Temel Kavramlar*, İstanbul: Der Yayınları.
- Güriş, S., Çağlayan, E. ve Güriş, B. (2011). *E-Views ile Temel Ekonometri*, İstanbul: Der Yayınları.
- Hanushek, E. A. ve Jackson, J. E. (1977). *Statistical Methods for Social Scientists*, New York: Academic Press.
- Hardy, M. A. (1993). *Regression with Dummy Variables*, London: SAGE Publications, Inc.

- Hausman, J. A. (1978). "Specification Tests in Econometrics". *Econometrica*, 46(6), 1251-1271.
- Hausman, J. A. ve McFadden, D. (1984). "Specification Tests for the Multinomial Logit Model". *Econometrica*, 1219-1240.
- Hausman, J. A. ve Wise, D. A. (1978). "A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences". *Econometrica*, 46(2), 403-426.
- Hays, W. L. (1994). *Statistics*, Orlando: Harcourt Brace & Company Inc.
- Hoetker, G. (2007). "The Use of Logit and Probit Models in Strategic Management Research: Critical Issues". *Strategic Management Journal*, 28(4), 331-343.
- Horowitz, J. L. ve Savin, N. E. (2001). "Binary Response Models: Logits, Probits and Semiparametrics". *American Economic Association*, 15(4), 43-56.
- Horrace, W. C. ve Oaxaca, R. L. (2006). "Results on the Bias and Inconsistency of Ordinary Least Squares for the Linear Probability Model". *Economic Letters*, 90(2006), 321-327.
- Hosmer, D. W. ve Lemeshow, S. (1980). "Goodness of Fit Tests for the Multiple Logistic Regression Model". *Communications in Statistics-Theory and Methods*, 9(10), 1043-1069.
- , (2000). *Applied Logistic Regression*, New York: John & Wiley Sons, Inc.
- Islam, S. ve Mannering, F. (2006). "Driver Aging and Its Effect on Male and Female Single-Vehicle Accident Injuries: Some Additional Evidence". *Journal of Safety Research*, 37(3), 267-276.
- İşyar, Y. (1999). *Ekonometrik Modeller*, Bursa: Vipaş A.Ş.
- Johnston, J. ve DiNardo, J. (1997). *Econometric Models*, New York: McGraw Hill.

- Jones, A. P. ve Jørgensen, S. H. (2003). "The Use of Multilevel Models for the Prediction of Road Traffic Accident Outcomes". *Accident Analysis and Prevention*, 35(1), 59-69.
- Karagöz, Y. (2010). *SPSS 18'deki Cross Tabulation Menüsünde Geçen İlişki Katsayıları*, Ankara: Detay Yayıncılık.
- Karayolları Genel Müdürlüğü 12. Bölge Müdürlüğü Erzurum Bölgedeki İller: Erzurum. Ankara: Karayolları Genel Müdürlüğü. Erişim tarihi: 29.06.2013.
<http://www.kgm.gov.tr/Sayfalar/KGM/SiteTr/Bolgeler/12Bolge/Iller/1Erzurum.aspx>
- Karayolları Genel Müdürlüğü 12. Bölge Müdürlüğü Erzurum Genel Bilgi. Ankara: Karayolları Genel Müdürlüğü. Erişim tarihi: 29.06.2013.
<http://www.kgm.gov.tr/Sayfalar/KGM/SiteTr/Bolgeler/12Bolge/Bolge12.aspx>
- Karayolları Genel Müdürlüğü 18. Bölge Müdürlüğü Kars Bölgedeki İller: Kars. Ankara: Karayolları Genel Müdürlüğü. Erişim tarihi: 29.06.2013.
<http://www.kgm.gov.tr/Sayfalar/KGM/SiteTr/Bolgeler/18Bolge/BolgedekiIller.aspx>
- Karayolları Genel Müdürlüğü 18. Bölge Müdürlüğü Kars Genel Bilgi. Ankara: Karayolları Genel Müdürlüğü. Erişim tarihi: 29.06.2013.
<http://www.kgm.gov.tr/Sayfalar/KGM/SiteTr/Bolgeler/18Bolge/Bolge18.aspx>
- Karayolları Trafik Kanunu, (1983), *T. C. Resmi Gazete*, 18195, 18 Ekim 1983.
- Kartal, M., Kutlar, A. ve Beğen, A. (2011). "Logistik Regresyon Tekniği ile Trafik Kazalarını Etkileyen Risk Faktörlerinin İncelenmesi: Sivas, Kayseri, Yozgat Örneği". *Ekonomik ve Sosyal Araştırmalar Dergisi*, 7(2), 45-68.
- Kayış, A. (2008). "Probit Regresyon Modeli". Şeref Kalaycı (Ed.). *SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri* (ss. 301-317). Ankara: Asil Yayın Dağıtım Ltd. Şti.

- Kiefer, N. M. (1982). "Testing for Dependence in Multivariate Probit Models". *Biometrika*, 69(1), 161-166.
- Kleinbaum, D. G., Kupper, L. L., Nizam, A. ve Muller, K. E. (2009). *Applied Regression Analysis*, New Delhi: Cengage Learning India Private Limited.
- Kmenta, J. (1997). *Elements of Econometrics*, Michigan: The University of Michigan Press.
- Knapp, L. G. ve Seaks, T. G. (1992). "An Analysis of the Probability of Default on Federally Guaranteed Student Loans". *The Review of Economics and Statistics*, 74(3), 404-411.
- Koppelman, F. S. ve Sethi, V. (2000). "Closed-Form Discrete-choice Models". David Alan Hensher, Kenneth John Button (Ed.). *Handbook of Transport Modelling* (ss. 211-228). Oxford: Pergamon Press.
- Koutsoyiannis, A. (1992). *Ekonometri Kuramı: Ekonometri Yöntemlerinin Tanıtımına Giriş*, (Çev.: Ümit Şenesen ve Gülay Günlük-Şenesen), İstanbul: İstanbul Teknik Üniversite Matbaası.
- Kutlar, A. (2009). *Uygulamalı Ekonometri*, Ankara: Nobel Yayın Dağıtım.
- Leonard, T. (2000). *A Course in Categorical Data Analysis*, New York: Chatman & Hall/CRC.
- Lesaffre, E. ve Kaufmann, H. (1992). "Existence and Uniqueness of the Maximum Likelihood Estimator for a Multivariate Probit Model". *Journal of the American Statistical Association*, 87(419), 805-811.
- Liu, I ve Agresti, A. (2005). "The Analysis of Ordered Categorical Data: An Overview and a Survey of Recent Developments". *Test*, 14(1), 1-73.
- Long, J. S. (1997). *Regression Models for Categorical and Limited Dependent Variables*, California: SAGE Publications, Inc.
- Long, J. S. ve Freese, J. (2001). *Regression Models for Categorical Dependent Variables Using Stata*, Texas: Stata Press.
- Luce, R. D. (1959). *Individual Choice Behavior: A Theoretical Analysis*, John Wiley & Sons Inc., New York.

- Luce, R. D. ve Suppes, R. (1965). "Preference, Utility, and Subjective Probability". Robert Duncan Luce, Robert R. Bush, E. H. Galanter (Ed.). *Handbook of Mathematical Psychology*, Vol. 3.(ss. 249-410). New York: Wiley & Sons Inc..
- Maddala, G. S. (1983). *Limited-dependent and Qualitative Variables in Econometrics*, Cambridge: Cambridge University Press.
- , (1992). *Introduction to Econometrics*, New York: Macmillan Publishing Company.
- Maddala, G. S. ve Flores-Lagunes, A. (2001). "Qualitative Response Models". Badi H. Baltagi (Ed.). *A Companion to Theoretical Econometrics* (ss. 366-382). Oxford: Blackwell Publishing Ltd.
- Mamčič, S. ve Sivilevičius, H. (2013). "The Analysis of Traffic Accidents on Lithuanian Regional Gravel Roads". *Transport*, 28(1), 108-115.
- Manski, C. F. (1977). "The Structure of Random Utility Models". *Theory and Decision*, 8(3), 229-254.
- Manski, C. F. ve Lerman, S. R. (1977). "The Estimation of Choice Probabilities From Choice Based Samples". *Econometrica*, 45(8), 1977-1988.
- Marschak, J. (1959). "Binary-Choice Constraints and Random Utility Indicators". *Cowles Foundation Paper No 155*, 312-329.
- Mazzanti, M. (2003). "Discrete Choice Models and Valuation Experiments". *Journal of Economic Studies*, 30(6), 584-604.
- McCullagh, P. (1980). "Regression Models for Ordinal Data". *Journal of the Royal Statistical Society Series B (Methodological)*, 42(2), 109-142.
- McCulloch, R. ve Rossi, P. E. (1994). "An Exact Likelihood Analysis of the Multinomial Probit Model". *Journal of Econometrics*, 64(1-2), 207-240.
- McFadden, D. (1974). "Conditional Logit Analysis of Qualitative Choice Behavior". Paul Zarembka (Ed.). *Frontiers in Econometrics* (ss. 105-142). New York: Academic Press.
- , (1976). "Quantal Choice Analysis: A Survey". *Annals of Social Measurement*, 5(4), 363-390.

- , (1978). "Modelling the Choice of Residential Location". Anders Karlqvist, Lars Lundqvist, Folke Snickers, Jorgen H. Weibull (Ed.). *Spatial Interaction and Planning Models* (ss. 75-96). Amsterdam: North-Holland Publishing Company.
- , (1980). "Econometric Models for Probabilistic Choice Among Products". *The Journal of Business*, 53(3), 13-29.
- , (1981). "Econometric Models of Probabilistic Choice". Charles F. Manski, Daniel McFadden (Ed.). *Structural Analysis of Discrete Data with Econometric Applications* (ss. 198-272). Cambridge: MIT Press.
- McKelvey, R. D. ve Zavoina, W. (1975). "A Statistical Model for the Analysis of Ordinal Level Dependent Variables". *Journal of Mathematical Sociology*, 4(1), 103-120.
- Menard, S. (2002). *Applied Logistic Regression Analysis*, London: SAGE Publications Inc.
- Milton, J. C., Shankar, V. N. ve Mannering, F. (2008). "Highway Accident Severities and the Mixed Logit Model: An Exploratory Empirical Analysis". *Accident Analysis and Prevention*, 40(1), 260-266.
- Montgomery, D. C., Peck, E. A. ve Vining, G. G. (2012). *Introduction to Linear Regression Analysis*, New Jersey: John Wiley & Sons, Inc.
- Nerlove, M. ve Press, S. J. (1973). *Univariate and Multivariate Log-linear and Logistic Models*, California: Rand Corporation.
- Oktay, E. (2005). *Kontenjans Tablolarından Elde Edilen İlişki Ölçüleri Öğretim Üyesi Değerleme Çalışması*, Erzurum: Aktif Yayınevi.
- Özer, H. (2004). *Nitel Değişkenli Ekonometrik Modeller*, Ankara: Nobel Yayın Dağıtım.
- Öztürkcan, M. (2009). *Regresyon Analizi*, İstanbul: T.C. Maltepe Üniversitesi Yayınları No: 40 (2009/3).
- Pampel, F. C. (2000). *Logistic Regression: A Primer*, California: SAGE Publications, Inc.
- Petersen, T. (1985). "A Comment on Presenting Results from Logit and Probit Models". *American Sociological Review*, 50(1), 130-131.

- Powers, D. A. ve Xie, Y. (1999). *Statistical Methods for Categorical Data Analysis*, New York: Academic Press Inc.
- Quandt, R. E. (1956). "A Probabilistic Theory of Consumer Behavior". *The Quarterly Journal of Consumer Behavior*, 70(4), 507-536.
- Ramanathan, R. (1989). *Introductory Econometrics with Applications*, Orlando: Harcourt Inc.
- Rao, V. R. ve Winter, F. W. (1978). "An Application of the Multivariate Probit Model to Market Segmentation and Product Design". *Journal of Marketing Research*, 15(3), 361-368.
- Reynolds, H. T. (1984). *Analysis of Nominal Data*, Beverly Hills: SAGE Publications, Inc.
- Shankar, V. N., Mannering, F. ve Barfield, W. (1996). "Statistical Analysis of Accident Severity on Rural Freeways". *Accident Analysis and Prevention*, 28(3), 391-401.
- Siegel, A. F. (2012). *Practical Business Statistics*, Burlington: Academic Press.
- Small, K. A. (1987). "A Discrete Choice Model for Ordered Alternatives". *Econometrica*, 55(2), 409-424.
- Small, K. A. ve Hsao, C. (1985). "Multinomial Logit Specification Tests". *International Economic Review*, 26(3), 619-627.
- Snell, E. J. (1964). "A Scaling Procedure for Ordered Categorical Data". *Biometrics*, 20(3), 592-607.
- Stevens, S. S. (1946). "On the Theory of Scales of Measurement", *Science*, 103, 677-680.
- Sze, N. N. ve Wong, S. C. (2007). "Diagnostic Analysis of the Logistic Model for Pedestrian Injury Severity in Traffic Crashes". *Accident Analysis and Prevention*, 39(6), 1267-1278.
- Tarı, R. (2012). *Ekonometri*, Kocaeli: Umuttepe Yayınları.
- Tatlıdil, H. (2002). *Uygulamalı Çok Değişkenli İstatistiksel Analiz*, Ankara: Ziraat Matbaacılık.
- Tay, R., Rifaat, S. M. ve Chin, H. C. (2008). "A Logistic Model of the Effects of Roadway, Environmental, Vehicle, Crash and Driver Characteristics

- on Hit-and-run Crashes”. *Accident Analysis and Prevention*, 40(4), 1330-1336.
- Theil, H. (1971). *Principles of Econometrics*, New York: John Wiley & Sons, Inc.
- Tobin, J. (1955). “The Application of Multivariate Probit Analysis to Economic Survey Data”. *Cowles Foundation Discussion Paper No.1*, 1-21.
- , (1958). “Estimation of Relationships for Limited Dependent Variables”. *Econometrica*, 26(1), 24-36.
- Train, K. E. (2003). *Discrete Choice Methods with Simulation*, Cambridge: Cambridge University Press.
- Tutz, G. (1991). “Sequential Models in Categorical Regression”. *Computational Statistics & Data Analysis*, 11(3), 275-295.
- Türkiye İstatistik Kurumu Adrese Dayalı Nüfus Kayıt Sistemi Sonuçları, 2012 Haber Bülteni (28 Ocak 2013). Ankara: Türkiye İstatistik Kurumu. Erişim tarihi: 29.06.2013. <http://www.tuik.gov.tr/PreHaberBultenleri.do?id=13425>
- Türkiye İstatistik Kurumu Motorlu Kara Taşıtları Nisan 2013 Haber Bülteni (18 Haziran 2013). Ankara: Türkiye İstatistik Kurumu. Erişim tarihi: 20.06.2013. <http://www.tuik.gov.tr/PreHaberBultenleri.do?id=13533>
- Türkiye İstatistik Kurumu (2012). *Karayolu Trafik Kaza İstatistikleri 2011*. Ankara: Türkiye İstatistik Kurumu.
- Üçdoğruk, Ş. (2002). *Multinomial Logit Model: Bölgelerarası İşgücü Tercihleri Analizi*, İzmir: Dokuz Eylül Üniversitesi Yayınları.
- Vovsha, P. (1997). “Application of Cross-Nested Logit Model to Mode Choice in Tel Aviv, Israel, Metropolitan Area”. *Transportation Research Record*, 1607, 6-15.
- Walker, J. ve Ben-Akiva, M. (2002). “Generalized Random Utility Models”. *Mathematical Social Sciences*, 43(3), 303-343.
- Wen, C. ve Koppelman, F. S. (2001). “The Generalized Nested Logit Model”. *Transportation Research Part B*, 35(7), 627-641.
- Yau, K. K. W. (2004). “Risk Factors Affecting the Severity of Single Vehicel Traffic Accidents in Hong Kong”. *Accident Analysis and Prevention*, 36(3), 333-340.

- Zellner, A. ve Lee, T. H. (1965). "Joint Estimation of Relationships Involving Discrete Random Variables". *Econometrica*, 33(2), 382-394.
- Zelterman, D. (2006). *Models for Discrete Data*, Oxford: Oxford University Press.
- Zhang, J., Fraser, S., Lindsay, J., Clarke, K. ve Mao, Y. (1998). "Age-specific Patterns of Factors Related to Fatal Motor Vehicle Traffic Crashes: Focus on Young and Elderly Drivers". *Public Health*, 112(5), 289-295.

EKLER

Ek1. 2918 Sayılı Karayolları Trafik Kanunu’nda Yer Alan ve Sıkça İhlal Edilen Trafik Kuralları (Asli Kusurlar ve Diğer Sürücü Kusurları)

	1- ASLİ KUSURLAR
1	Kırmızı ışıklı trafik işaretinde veya yetkili memurun dur işaretinde geçme
2	“Taşıt giremez” trafik işaretinin bulunduğu karayoluna veya bölünmüş karayolunda karşı yönden gelen trafiğin kullandığı şerit, rampa ve bağlantı yollarına girme
3	İkiden fazla şeritli taşıt yollarında, karşı yoldan gelen trafiğin kullandığı şerit veya yol bölümüne girme
4	Arkadan çarpma
5	Geçme yasağı olan yerlerde geçme
6	Doğrultu değiştirme manevralarını yanlış yapma
7	Şeride tecavüz etme
8	Kavşaklarda geçiş önceliğine uymama
9	Kaplamanın dar olduğu yerlerde geçiş önceliğine uymama
10	Manevraları düzenleyen genel şartlara uymama
11	Yerleşim birimleri dışındaki karayolunun taşıt yolu üzerinde zorunlu haller dışında park etme veya duraklama ve her durumda gerekli tedbirleri almama
12	Park için ayrılmış yerlerde veya taşıt yolu dışında kurallara uygun olarak park etmiş araçlara çarpma
	2- DİĞER KUSURLAR
Madde 47/1-b	İşıklı trafik işaretlerinden kırmızı renkli olanına ve sesli işaretlere uymamak
Madde 47/1-c	Trafik işaret levhaları, cihazları ve yer işaretlemeleri ile belirtilen veya gösterilen hususlara uymamak
Madde 47/1-d	Trafik güvenliği ve düzeniyle ilgili olan diğer kural, yasak, zorunluluk veya yükümlülüklere uymamak
Madde 48/5	Taksi, dolmuş, minibüs, otobüs, kamyon ve çekici gibi ticari amaçla yük

	ve yolcu taşımacılığı yapılan araçlar ile resmi araçları alkollü olarak kullanmak, kanlarındaki alkol miktarının 0.50 promilin üstündeyken diğer araçları kullanmak
Madde 51/2-a	Hız sınırlarını % 10'dan % 50'ye kadar aşmak (% 50 dâhil)
Madde 51/2-b	Hız sınırlarını % 50'den fazla aşmak
Madde 52/1-a	Araçların hızını, kavşaklara yaklaşırken, dönemeçlere girilirken, tepe üstlerine, yaya geçitlerine, hemzemin geçitlerine, tünellere, dar köprü ve menfezlere yaklaşırken, yapım ve onarım alanlarına girerken azaltmamak
Madde 52/1-b	Araçların hızını, aracın yük ve teknik özellikleriyle görüş, yol, hava ve trafik durumunun gerektirdiği şartlara uymamak
Madde 52/1-c	Yol, hava ve trafik durumlarını göz önüne almadan, öndeki aracı güvenli bir mesafe bırakmadan takip etmek
Madde 53/1-a	Sağa dönüş kurallarına riayet etmemek
Madde 53/1-b	Sola dönüş kurallarına riayet etmemek
Madde 54/1-a	Geçme kurallarına riayet etmemek
Madde 54/1-b	Geçmenin yasak olduğu yerlerde geçmek
Madde 56/1-a	Şerit izleme ve değiştirme kurallarına uymamak
Madde 56/1-c	Önde giden bir aracı güvenli ve yeterli mesafeden izlememek
Madde 57/1-a	Kavşaklara yaklaşırken yavaşlamamak ve geçiş hakkı olan araçların önce geçmesine imkân vermemek
Madde 57/1-b	Trafik zabıtası, trafik işaret levhası veya ışıklı trafik cihazları bulunmayan kavşaklarda geçiş hakkına riayet etmemek
Madde 57/1-c	Kontrolsüz kavşaklarda motorsuz araç kullananların motorlu araçlara, motorlu araç kullananlardan soldakinin sağdan gelene geçiş hakkını vermemesi
Madde 58	İndirme ve bindirme veya inme-binme kurallarına riayet etmemek

Madde 61	Taşıt yolu üzerinde ve duraklamanın yasak olduğu yerlerde ve hallerde park etmek
Madde 67	Park yapmış taşıtlar arasından çıkarken, duraklarken taşıt yolunun sağına ve soluna yanaşırken, trafik kural ve yasaklarına uymamak

Kaynak: Karayolları Genel Müdürlüğü Trafik Kazası Tespit Tutanağı, Türkiye İstatistik Kurumu
Karayolu Trafik Kazaları İstatistikleri 2011

ÖZGEÇMİŞ

Kişisel Bilgiler	
Adı Soyadı	Ali Kemal ÇELİK
Doğum Yeri ve Tarihi	Kars, 27.09.1983
Eğitim Durumu	
Lise Öğrenimi	Kars Anadolu Lisesi
Lisans Öğrenimi	Hacettepe Üniversitesi İktisadi ve İdari Bilimler Fakültesi İşletme (İngilizce) Bölümü
Yüksek Lisans Öğrenimi	Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Sayısal Yöntemler Anabilim Dalı
Bildiği Yabancı Diller	İngilizce
İş Deneyimi	
Araştırma Görevlisi	Ardahan Üniversitesi (2010 - 2011)
Araştırma Görevlisi	Atatürk Üniversitesi (2011 - Devam etmekte)
İletişim	
E-Posta Adresi	akemal.celik@atauni.edu.tr
Tarih	05/07/2013