



YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

nitel tercih modl, çoklu logit, probi..

Yüksek Lisans Tezi

özgür sönmez

2/8

**YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**NİTEL TERCİH MODELLERİ, ÇOKLU LOGİT, PROBİT
MODELLER ve BİR UYGULAMA**

1107

İstatistikçi Özgür SÖNMEZ

**FBE İstatistik Anabilim Dalı İstatistik Programında
Hazırlanan**

YÜKSEK LİSANS TEZİ

Tez Danışmanı : Doç. Dr. Ali Hakan BÜYÜKLÜ

İSTANBUL, 2006

YILDIZ TEKNİK ÜNİVERSİTESİ
KÜTÜPHANE VE DOKÜMANTASYON
DAİRE BAŞKANLIĞI

Yer No (DDC): R.360/8

Kayıt No

3080

Geldiği Yer

Fen Bilim Enst.

Tarih

31.05.06

Fiyat

7.50

Fatura No

:

Ayniyat No

1-6

Ek

:

NİTEL TERCİH MODELLERİ, ÇOKLU LOGİT, PROBİT MODELLER ve BİR UYGULAMA

2.1 İki Değerli
2.2 Çoklu Değerli
2.3 Basit Doğrusal
2.4 Çok Değerli

2.5 GÖLGE OLANAK
2.6 Varyans ve Kovaryans
2.7 Gölge Değerleri
2.8 Regresyon

İstatistikçi Özgür SÖNMEZ

FBE İstatistik Anabilim Dalı İstatistik Programında
Hazırlanan

YÜKSEK LİSANS TEZİ

Ali Hakan Büyüklü

Tez Danışmanı : Doç. Dr. Ali Hakan BÜYÜKLÜ

Osman Çetinkaya
Ali Hakan Büyüklü

Ali Hakan Büyüklü
Doç. Dr. Ali Hakan BÜYÜKLÜ

İÇİNDEKİLER

Sayfa

SİMGE LİSTESİ	ii
KISALTMA LİSTESİ	iv
ŞEKİL LİSTESİ	v
ÇİZELGE LİSTESİ	vi
ÖNSÖZ	viii
ÖZET	ix
ABSTRACT	x
1. GİRİŞ	1
2. REGRESYON ANALİZİ	4
2.1 Regresyon Analizinin Tanımı	4
2.2 İki Değişkenli Regresyon Modeli	5
2.3 Çoklu Doğrusal Regresyon	8
2.4 Basit Doğrusal Regresyonun Matris Yardımıyla Çözümü	10
2.5 Çok Değişkenli Doğrusal Regresyonun Matrisle Çözümü	11
3. GÖLGE DEĞİŞKEN	12
3.1 Varyans ve Kovaryans Analizi Modelleri	12
3.2 Gölge Değişken Tuzağı	15
3.3 Regresyon Modellerinin F Testleri Yardımıyla Karşılaştırılması	17
4. NİTEL TERCİH MODELLERİ	21
4.1 İkili Tercih Modelleri	21
4.1.1 Doğrusal Olasılık Modeli	21
4.1.2 Lojistik Regresyon ve Logit Model	27
4.1.2.1 Lojistik Regresyonda Genel İfadeler	27
4.1.2.2 Bahis Oranı (Odds Ratio) Hesaplanması	30
4.1.2.3 Logit Modelinin Tahmin Edilmesi	32
4.1.2.4 Uyum İyiliği Ölçüleri	35
4.1.3 Probit Model	41
4.2 Çoklu Tercih Modelleri	50
4.2.1 Sıralı Olmayan Tercih Modelleri	51
4.2.1.1 Çok Durumlu Doğrusal Olasılık Modeli	51
4.2.1.2 Çok Durumlu Probit Model	52
4.2.1.3 Çok Durumlu Logit Model	53
4.2.1.4 Koşullu Logit Model (Conditional Logit Model)	55
4.2.1.5 Yuvalanmış Logit Model	56
4.2.1.5.1 Genelleştirilmiş Uç Değer Modeli (Generalized Extreme Value Model)	56
4.2.1.5.2 Yuvalanmış Model (Nested Model)	57
4.2.2 Sıralı (Ordered) Tercih Modelleri	60
4.2.3 Ardışık (Sequential) Tercih Modelleri	64
4.2.4 Tercih Modellerine İlişkin Genel İfadeler	65

5.	EKONOMİK TÜKETİCİ TEORİSİYLE MODEL KURMA	68
5.1	Ekonomik Tüketici Teorisi	68
5.2	Rasgele Fayda (Random Utility)	69
5.3	İkili Seçim Modelleri	70
5.3.1	İkili Seçim Model Çeşitleri	70
5.3.1.1	İkili Doğrusal Olasılık Modeli	70
5.3.1.2	İkili Logit	72
5.3.1.3	İkili Probit	73
5.3.1.4	Diğer İkili Seçim Modelleri	74
5.3.1.4.1	Arctan Olasılık Modeli	74
5.3.1.4.2	Sağdan Kesikli Üstel Model	74
5.3.1.4.3	Soldan Kesikli Üstel Model	75
5.3.2	İkili Seçim Modellerinde Sınırlı Durumlar (Limiting Cases)	75
5.3.3	İkili Seçim Modellerinde En Çok Benzerlik Tahmin Yöntemi	76
5.3.3.1	Tahmin Yöntemi İçin Genel Gösterim	76
5.3.3.2	İkili Logit Modelde En Çok Benzerlik Tahmini	79
5.3.3.3	İkili Probit Modelde En Çok Benzerlik Tahmini	80
5.4	Çoklu Seçim Modelleri	80
5.4.1	Çoklu Seçim Teorisi	80
5.4.2	Çoklu Logit Model	81
5.4.2.1	Çoklu Logit Modelin Tanımlanması	81
5.4.2.2	Gumbel Dağılımına İlişkin Genel Özellikler	82
5.4.2.3	Çoklu Logit Modelin Elde Edilmesi	83
5.4.2.4	Çoklu Logit Modelde Sınırlı Durumlar (Limiting Cases)	84
5.4.2.5	Parametrelerinde Doğrusal Çoklu Logit Model	84
5.4.2.6	Logit Modelin Özellikleri	85
5.4.2.6.1	Alternatifler Arası Bağımsızlık (Independent from Irrelevant Alternatives)	85
5.4.2.6.2	Çoklu Logit Modelde Elastikiyetler (Esneklikler)	85
5.4.2.7	Arttırımlı (Incremental) Çoklu Logit Model	87
5.4.2.8	Çoklu Logit Modelin Tahmin Edilmesi	87
5.4.2.8.1	Çoklu Logit Modelde En Çok Benzerlik Yöntemi	87
5.4.2.8.2	Tekrarlı Gözlemler İçin En Çok Benzerlik Yöntemi	88
6.	ÇALIŞMA GRUPLARINI ETKİLEYEN ETMENLERİN ARACILIĞI İLE GRUPLARA AİT OLMA OLASILIKLARININ BELİRLENMESİNDE SIRALI LOGİT ve SIRALI PROBİT MODELLERİNİN KULLANIMI	89
6.1	İkili Model Analizleri	90
6.1.1	İkili Doğrusal Olasılık Modeli Sonuçları	91
6.1.2	İkili Logit Model Sonuçları	94
6.1.3	İkili Probit Model Sonuçları	100
6.2	Çoklu Model Analizleri	104
6.2.1	Çoklu Logit Model Sonuçları	104
6.2.2	Çoklu Probit Model Sonuçları	108
7.	SONUÇLAR	118
	KAYNAKLAR	123

EKLER.....	126
Ek 1 Veri Seti	127
Ek 2 İkili Logit Modelde Gözlem, Tahmin ve Kalıntı Değerleri	131
Ek 3 Sıralı Probit Modelde Alternatiflere İlişkin Olasılıklar	135
ÖZGEÇMİŞ.....	141

SİMGE LİSTESİ

AIC	Akaike Bilgi Kriteri
b_j^*	Standartlaştırılmış regresyon katsayısı
β_i	Parametre
$\hat{\beta}_i$	Parametre tahmini
C	Evrensel küme
C_n	Alternatiflere ilişkin seçim kümesi
$\tilde{C}_n$	Seçim kümesinin alt kümesi
Cov	Kovaryans
df	Serbestlik derecesi (degree of freedom)
D_i	Gölge değişken (dummy variable)
D_{in}	n . bireyin i . alternatifi seçme sayısı
e_i	Örneklem hata payı (kalıntı)
$\exp(x)$	e^x
ε_{ij}	Faydanın gözlenemeyen yani rasgele kısmı (disturbans)
E	Beklenen değer operatörü
E_X^P	Esneklik, elastikiyet
$f(x)$	Olasılık yoğunluk fonksiyonu
$F(x)$	Birikimli dağılım fonksiyonu
HQ	Hannan-Quinn kriteri
I_i	Endeks değeri (fayda endeksi)
Inc	Gelir
IV	Dahili değer (inclusive value)
L_i	Logit
LL	Benzerlik fonksiyonu (likelihood)
m_i	İterasyon sayacı
$n.e.s$	Normal eşdeğeri sapma (normit)
$NOWP$	Yanlış tahmin edilenlerin sayısı (number of wrong predictions)
$odds(A)$	A olayına ilişkin bahis değeri
OO	Olabilirlik oranı
OR	Bahis oranı (odds ratio)
P_{ji}	Olasılıkların ortalamalardan yani $\bar{P}_j$ ' den farkları
prc_i	Fiyat
PCP	Doğru tahminlerin oranı (percent of correct predictions)
P_{ji}	i . bireyin j . alternatifi seçme olasılığı
$P_n(j)$	n . bireyin j . alternatifi seçme olasılığı
$\bar{P}_j$	j . alternatifin seçilme olasılıklarının ortalaması
q_i	Mal veya hizmet miktarı
R_{EF}^2	Efron R^2
R_{MF}^2	McFadden R^2
R_{MD}^2	Maddala R^2

R_{AN}^2	Aldrich Nelson R^2
R_{PSD}^2	Pseudo R^2
R_{MZ}^2	McKelvey ve Zovoina R^2
R_{VZ}^2	Veall-Zimmerman R^2
se	Standart hata (Standard error)
S_{qcc}	Kareli korelasyon katsayısı
S_n	Bireyin karakteristik özellikleri
SC	Schwarz kriteri
u_i	Anakütle hata payı
U_{ij}	i . bireyin j . alternatiften elde ettiği fayda
Var	Varyans
V_{ij}	Faydanın sistematik kısmı
w_i	Ağırlıklı en küçük kareler yönteminde kullanılan ağırlıklar
$WSSR$	Ağırlıklı hata kareleri toplamı
X_i	Bağımsız Değişken
Y_i	Bağımlı değişken
z	Standart normal değişken
z_{in}	Tercihin kendisi ile ilgili olan ve n . birey tarafından değerlendirilen özellikleri
Φ	Birikimli normal dağılım fonksiyonu
ϕ	Standartlaştırılmış normal yoğunluk fonksiyonu
Λ	Birikimli lojistik dağılım
η_i	Gumbel dağılımına ilişkin mod değeri
δ_{ij}	Kronecker delta fonksiyonu

KISALTMA LİSTESİ

a.g.e	Adı Geçen Eser	Sayfa
ANOVA	Varyans Analizi (Analysis of Variance)	8
ANCOVA	Kovaryans Analizi (Analysis of Covariance)	23
BDF	Birikimli Dağılım Fonksiyonu	27
DOM	Doğrusal Olasılık Modeli	38
DORM	Doğrusal Olmayan Regresyon Modeli	41
DRM	Doğrusal Regresyon Modeli	44
EYO	En Yüksek Olabilirlik (En Çok Benzerlik)	45
GEV	Genelleştirilmiş Uç Değer (Generalized Extreme Value)	46
IIA	Alternatifler Arası Bağımsızlık (Independence of Irrelevant Alternatives)	46
OYF	Olasılık Yoğunluk Fonksiyonu	62
SSE	Regresyonla Açıklanamayan Değişkenlik(Error Sum of Squares)	70
SSTO	Toplam Değişkenlik(Total Sum of Squares)	71
SSR	Regresyonla Açıklanan Değişkenlik(Regression Sum of Squares)	72
s.s	Sayfa sıralı	73
YTL	Yeni Türk Lirası	74

ŞEKİL LİSTESİ

	Sayfa
Şekil 2.1	Anakütle regresyon doğrusu 7
Şekil 2.2	Anakütle ve örnek regresyon doğruları 8
Şekil 4.1	Doğrusal olasılık modeli..... 23
Şekil 4.2	Birikimli dağılım fonksiyonu 27
Şekil 4.3	Logit ve probit model birikimli dağılım eğrileri..... 28
Şekil 4.4	Probit model..... 41
Şekil 4.5	Endeks değerine bağlı olarak olasılık bulunması..... 44
Şekil 4.6	Olasılık değerine bağlı olarak endeks değerinin bulunması 45
Şekil 4.7	Standart normal olasılık yoğunluk fonksiyonu 45
Şekil 4.8	İki seviyeli yuvalanmış logit model 58
Şekil 4.9	Sıralı probit modelde olasılıklar..... 62
Şekil 5.1	Düzgün dağılım..... 70
Şekil 5.2	Doğrusal olasılık modeli..... 71
Şekil 5.3	Düzgün dağılım olasılık yoğunluk fonksiyonu..... 72
Şekil 5.4	İkili probit model 73
Şekil 5.5	Sağdan kesikli üstel model 74
Şekil 5.6	Soldan kesikli üstel model 75
Şekil 5.7	İkili seçim modellerinde sınırlı durumlar 76
Şekil 6.1	İkili DOM' da kalıntı, gözlenen ve tahmin edilen değerler..... 92
Şekil 6.2	Jarque – Bera testi için frekans tablosu 93
Şekil 6.3	İkili logit modelde kalıntı, gözlenen ve tahmin edilen değerler 96
Şekil 6.4	İkili probit modelde kalıntı, gözlenen ve tahmin edilen değerler 101

ÇİZELGE LİSTESİ

Sayfa

Çizelge 2.1	Doğrusal ve doğrusal olmayan regresyon modelinin belirlenmesi.....	5
Çizelge 2.2	Bağımlı ve bağımsız değişkenlerin dağılımı	6
Çizelge 2.3	Çoklu regresyon modelinde verilerin gösterimi	9
Çizelge 4.1	Endeks değerine bağlı olarak normal ve lojistik bdf değerleri	47
Çizelge 4.2	İkili tercih modellerinde tahminciler arasındaki ilişkiler	47
Çizelge 6.1	Modeldeki değişkenlerin tanımlayıcı istatistik değerleri	90
Çizelge 6.2	İkili doğrusal olasılık modelinde parametre tahminleri	91
Çizelge 6.3	İkili DOM' da sorunlu gözlemler ve tahmin değerleri.....	93
Çizelge 6.4	İkili logit modelde parametre tahminleri.....	94
Çizelge 6.5	İkili logit modelde değişkenlere ilişkin bahis oranları.....	95
Çizelge 6.6	İkili logit modelde açıklanan değişkene göre tanımlayıcı istatistikler.....	97
Çizelge 6.7	İkili logit modelde eşik değere göre gruplandırma ve modelin doğruluğu	98
Çizelge 6.8	İkili logit modelde Andrews ve Hosmer–Lemeshow uygunluk testi sonuçları.....	99
Çizelge 6.9	İkili probit modelde parametre tahminleri.....	100
Çizelge 6.10	İkili probit modelde eşik değere göre gruplandırma ve modelin uygunluğu ..	102
Çizelge 6.11	İkili probit modelde Andrews ve Hosmer – Lemeshow uygunluk testi sonuçları	102
Çizelge 6.12	Çoklu modelde bağımlı değişkenin tanımlayıcı istatistik değerleri.....	104
Çizelge 6.13	Sıralı modelde bağımlı değişkenin gruplara göre dağılımı	105
Çizelge 6.14	Sıralı logit modelde parametre ve limit değerleri tahminleri	105
Çizelge 6.15	Sıralı logit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları	107
Çizelge 6.16	Sorunlu değişkenin çıkartılmasıyla elde edilen sıralı logit modelde parametre ve limit değerleri	107
Çizelge 6.17	Sorunlu değişkenin çıkartılmasıyla elde edilen sıralı logit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları	108
Çizelge 6.18	Sıralı probit modelde parametre ve limit değerleri tahminleri	109
Çizelge 6.19	Sıralı probit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları	110
Çizelge 6.20	Sorunlu değişkenin çıkartılmasıyla elde edilen sıralı probit modelde parametre ve limit değerleri	111
Çizelge 6.21	Sorunlu değişkenin çıkartılmasıyla elde edilen sıralı probit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları	111
Çizelge 6.22	Sıralı logit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları	112
Çizelge 6.23	Sıralı probit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları	112
Çizelge 6.24	Sıralı logit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları	113
Çizelge 6.25	Sıralı probit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları	113
Çizelge 6.26	Sıralı logit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları	114
Çizelge 6.27	Sıralı probit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları	114
Çizelge 6.28	Sıralı logit modelde parametre ve limit değerleri tahminleri (örneklem 1)....	115
Çizelge 6.29	Sıralı probit modelde parametre ve limit değerleri tahminleri (örneklem 1) ..	115
Çizelge 6.30	Sıralı logit modelde parametre ve limit değerleri tahminleri (örneklem 2)....	116

Çizelge 6.31	Sıralı probit modelde parametre ve limit değerleri tahminleri (örneklem 2) .	116
Çizelge 6.32	Sıralı logit modelde parametre ve limit değerleri tahminleri (örneklem 3)....	117
Çizelge 6.33	Sıralı probit modelde parametre ve limit değerleri tahminleri (örneklem 3) .	117

ÖNSÖZ

Çalışma konusunun belirlenmesinde, anlaşılmasında, ilerlemesinde ve tüm aksiliklere rağmen bu seviyeye getirilmesinde katkısı olan, yardımlarını ve emeklerini hiçbir zaman unutamayacağım sayın danışmanım Doç. Dr. Ali Hakan BÜYÜKLÜ' ye;

Hayatımın her döneminde en büyük yardımcım olan ve aksiliklerime rağmen maddi, manevi desteklerini her zaman gösteren annem Fatma SÖNMEZ, babam Şahin SÖNMEZ ve kardeşim Özlem SÖNMEZ' e;

Varlığı, güleryüzü ve tüm sevecenliğiyle her anında yanımda olan nişanlım Derya ALTIN' a;

İstatistik bölümünde görev alan tüm öğretim elemanlarına ve çalışanlarına;

Gerek iş yaşamımda gerekse de okul yaşamımda bilgi, beceri ve destekleriyle bana yardımcı olan arkadaşlarıma;

teşekkür ederim.

Ocak, 2006

Özgür SÖNMEZ

ÖZET

İki veya daha fazla değer alan gölge değişkenler bağımlı değişken olarak regresyon modellerinde yer alabilmektedir. Bağımlı değişkenin nitel özellikte olduğu nitel tercih modelleri ikili ve çoklu olmak üzere iki gruba ayrılmaktadır. Her iki tip modelde de temel amaç tercih olasılığının belirlenmesidir.

İkili tercih modelleri karar birimlerinin iki alternatif arasından bir seçim yapmakla karşı karşıya olduğunu ve seçimin karar birimlerinin karakteristik özelliklerine bağlı olduğunu varsaymaktadır. Probit ve logit model arasındaki fark, probit model normal dağılım ile bağlantılı iken, logit model lojistik dağılım ile bağlantılıdır.

Çoklu tercih modellerinde karar birimleri ikiden fazla tercih ile karşı karşıyadır. Kategorik değişkenler sıralı, sıralı olmayan ve ardışık olmak üzere üç grup halinde sınıflandırılmaktadır. Bu kategorik değişkenler kullanılarak çeşitli çoklu tercih modelleri oluşturulmaktadır. Karar birimlerinin tercihlerinin açıklanmasında rasgele fayda modelinden yararlanılmaktadır. Bir alternatif ile ilişkin rasgele fayda, gözlenen veya sistematik bileşen ile gözlenemeyen bileşenlerin toplamı olarak ifade edilmektedir. i . alternatifin seçilme olasılığı, bu alternatifin seçilmesiyle elde edilecek fayda olan U_{in} ' in seçim sepetindeki diğer tüm alternatiflerin seçilmeleriyle elde edilecek olan faydaya eşit veya bu faydadan büyük olma olasılığına eşittir. Çoklu logit modelde hata terimlerinin birbirinden bağımsız ve özdeş olarak Weibull dağılımlı oldukları varsayılmakta iken çoklu probit modelde hata terimlerinin çok değişkenli normal dağılıma uygun oldukları varsayılmaktadır.

Bu uygulamada bireyler çalışma statülerine göre beş gruba ayrıldı ve bu statü değerleri bağımlı değişken olarak kullanıldı. Sıralı probit ile sıralı logit model uygulamaları yapıldı ve her birey için gruplara ilişkin olasılıklar hesaplandı.

Anahtar kelimeler: Gölge değişken, nitel tercih modelleri, sıralı probit, sıralı logit, rasgele fayda.

ABSTRACT

Dummy variables that take two or more values may be dependent variable in the regression models. Qualitative choice models, where the dependent variable is qualitative, are separated two groups as binary and multinomial. The basis aim in two groups of models is to determine the choice probabilities.

Binary choice models assume that decision makers are faced with a choice between two alternatives and that their choice depends on their characteristics. The difference between probit and logit models is that while probit is related to the normal distribution, logit is related to the logistic distribution.

Decision makers are faced with more than two alternatives in multinomial choice models. Categorical variables are classified into three groups as ordered, unordered and sequential. The various multinomial choice models are formed by using these categorical variables. The random utility model is utilized to explain the choices of the decision makers. The random utility of an alternative is expressed as a sum of observable or systematic and unobservable components. The choice probability of alternative i is equal to the probability that the utility of alternative i , U_{in} , is greater than or equal to the utilities of all other alternatives in the choice set. The error terms are assumed to be distributed multivariate-normally in the multinomial probit model while the error terms are assumed to be independently and identically distributed as a Weibull distribution in the multinomial logit model.

In this application, individuals were separated into five groups according to the working status and working status were used as dependent variables. The ordered probit and ordered logit models were applied and for each group, the probabilities that were associated with the groups were calculated.

Keywords: Dummy variable, qualitative choice models, ordered probit, ordered logit, random utility

1. GİRİŞ

Değişkenler arasındaki ilişkilerin belirlenmesinde kullanılan yöntemlerden biri olan regresyon analizi, bağımlı bir değişken ile bir veya daha fazla bağımsız değişken arasındaki ilişkiyi açıklama ve değerlendirme ile ilgilidir. Regresyon analizinde değişkenler arasındaki bağımlılık fonksiyonel veya deterministik değil, istatistikseldir.

Çalışmanın regresyon analizi başlıklı ilk bölümünde, konuya bir temel olması açısından regresyon analizi, iki değişkenli regresyon modeli, doğrusal ve doğrusal olmayan modeller, anakütle regresyon doğrusu, örnek regresyon denklemi, çoklu doğrusal regresyon ve doğrusal regresyon çözümünde matris yaklaşımı konuları incelenmektedir.

Regresyon analizinde bağımlı ve bağımsız değişkenlerin gelir, fiyat gibi nicel (kantitatif) özellikte olmasının yanı sıra bu değişkenler cinsiyet, eğitim durumu, din gibi nitel (kalitatif) özellikte de olabilmektedir. Nitel değişkenlerin regresyon analizinde kullanılması sırasında bu değişkenler 0, 1 ya da 2 olarak nicel hale getirilmekte ve böylece yapay değişkenler oluşturulmaktadır. Bu yapay değişkenler gölge değişken olarak adlandırılmaktadır.

Gölge değişken başlığı altında ele alınan bölümde gölge değişkene ilişkin tanımlamalar, varyans ve kovaryans analizi modelleri, gölge değişken tuzağı ve bu tuzağa düşmemek için uyulması gereken kurallara ilişkin bir örnek, gölge değişken içeren regresyon modellerinin F testleri yardımıyla karşılaştırılması konularına değinilmektedir.

Gölge değişken başlıklı bölümde yer alan modellerde bağımlı değişken nicel özellikte iken bağımsız değişken ya nicel, ya da nitel özellikte olmaktadır. Bağımlı değişkenin nitel özellikte olduğu nitel tercih modelleri (qualitative response models) olarak bilinen gölge bağımlı değişkenli modeller ikili ve çoklu olmak üzere 2 ana grup altında toplanmaktadır. Nitel tercih modellerinde temel amaç bireyin alternatifi seçme ya da tercih etme olasılığının belirlenmesidir. Karar sürecinin olasılıklı yapısı hakkında geliştirilen varsayımlara dayalı olarak çeşitli tercih modelleri oluşturulmaktadır.

İkili tercih modelleri başlıklı bölümde sırasıyla doğrusal olasılık modeli, logit ve probit modeller incelenmektedir. Bu bölümde doğrusal olasılık modeline ilişkin tanımlamalar, modelin genel özellikleri ve bu modelin kullanılmasıyla ortaya çıkan birtakım sorunlar ele alınmaktadır. Doğrusal olasılık modelinde değinilen sorunlardan kurtulmak için birikimli dağılım fonksiyonları yardımıyla kurulan modellerden en çok kullanılanlar logit ve probit modellerdir. Lojistik ve normal birikimli dağılım fonksiyonlarının kullanıldığı modeller sırasıyla logit ve probit model olarak bilinmektedir. Ayrıca bu bölümde lojistik regresyonla ilgili genel ifadeler, logit ve bahis oranı kavramları, modelin parametrelerinin tahmin

edilmesi, uyum iyiliği ölçüleri, fayda endeksi, marjinal etkilerin hesaplanması, katsayıların karşılaştırılması konularına yer verilmektedir.

Çoklu tercih modellerinde karar birimleri ikiden fazla tercihle karşı karşıyadır ve dolayısıyla bağımlı değişken ikiden fazla değer almaktadır. Çoklu tercih modelleri bölümünde kategorik ve kategorik olmayan değişken ayırımı yapılmaktadır. Kategorik değişkenler sıralı, sıralı olmayan ve ardışık olarak sınıflandırılmakta ve bu değişkenler kullanılarak çoklu tercih modelleri oluşturulmaktadır. 3 alt grupta toplanan çoklu tercih modellerinden ilki olan sıralı olmayan tercih modelleri çok durumlu doğrusal olasılık, çok durumlu probit, çok durumlu logit, koşullu logit ve yuvalanmış logit olmak üzere 5 başlık altında incelenmektedir. Bu modellerde göze çarpan temel nokta, sistematik ve gözlenemeyen olmak üzere 2 bileşenden oluşan-rasgele fayda modelinden yararlanılmasıdır. Sıralı ve ardışık tercih modelleri de hem probit hem de logit model göz önüne alınarak incelenmektedir. Bu bölümde ayrıca ikili ve çoklu tercih modellerinde kullanılan matris notasyonuna değinilmektedir.

Ekonomik tüketici teorisi bireylerin tercihlerinin açıklanmasında yararlanılan yaklaşımlardan biridir. Bireylerin mümkün olan tüm alternatifleri karşılaştırabildikleri farzedilmektedir. Bu karşılaştırma esnasında bireyler bütçe kısıtlamasına uymakla beraber kendileri için en çok fayda sağlayacak olan alternatifi seçmektedirler.

Ekonomik tüketici teorisiyle model kurma başlıklı bölümde fayda fonksiyonu ve bileşenleri, rasgele faydanın tercih olasılıklarının belirlenmesinde kullanımı, ikili seçim modellerinin fayda fonksiyonuyla irdelenmesi, çoklu seçim teorisi, fayda fonksiyonu bileşenleriyle çoklu logit modelin tanımlanması, gumbel dağılım özellikleri ve bu dağılım yardımıyla çoklu logit modelin elde edilmesi, alternatifler arası bağımsızlık (IIA) varsayımı ve elastikiyet özelliklerinin incelenmesi konuları açıklanmaktadır.

Uygulama aşamasında, çalışan kişiler 5 gruba ayrılmakta ve grup değerleri nitel bağımlı değişken olarak tanımlanmaktadır. Ele alınan 5 grup kendi arasında 2 gruba ayrılarak ikili seçim modelleri oluşturulmaktadır. Uygulamada kullanılan bağımsız değişkenler 2' si nicel (piyasa deneyim yılı, resmi kurumlarca onaylı eğitim yılı), 1' i nitel (ırk) olmak üzere 3 adettir. Hem ikili hem de çoklu modellerin kurulmasında E-views 3.0 paket programından yararlanılmaktadır. Uygulamanın ikinci aşamasında çoklu logit ve probit model kurulmaktadır. Bağımlı değişkenin özelliğine, değerlerine bakıldığında sıralı ölçekli olduğu görülmekte ve bu yüzden dolayı sıralı logit ve sıralı probit modellere ilişkin tahmin yöntemleri uygulanarak sonuçlar elde edilmektedir.

Sıralı modelde bağımlı değişkenin gruplara göre dağılımı, parametre ve limit değerleri tahminleri, her gözlem için gruplara ait olma olasılıkları, tahmin edilen grup frekansları hesaplanmaktadır.

Ayrıca, gruplardaki gözlem sayılarının eşit olduğu 3 örneklem ele alınarak sıralı logit ve sıralı probit model analizleri yapılmaktadır.

2. REGRESYON ANALİZİ

2.1 Regresyon Analizinin Tanımı

İstatistiksel araştırmalarda iki veya daha çok değişken arasındaki ilişkinin incelenmesi gerekebilir. Değişkenler arasındaki ilişkinin belirlenmesi için en çok kullanılan yöntemlerden birisi de regresyon analizidir. Regresyon analizi bağımlı bir değişken (Y) ile bir veya daha fazla bağımsız ($X_1, X_2, X_3, \dots, X_k$) değişken arasındaki ilişkiyi açıklama ve değerlendirme ile ilgilidir. Örneğin sporcularda milli olma sayısı ile genel spor kültür puanı arasında bir ilişki olup olmadığı öğrenilmek istenebilir. Bu çalışmada, sporcuların genel spor kültür puanları milli olma sayısına göre değişebileceği, yani milli olma sayısı spor kültür puanlarını etkileyebileceği için, milli olma sayısı bağımsız değişken, genel spor kültür puanı da bağımlı değişken olarak adlandırılır.¹ Bağımlı değişken açıklanan, etkilenen veya sonuç değişken olarak da tanımlanmaktadır. Ayrıca bağımsız değişken ise açıklayan, etkileyen veya neden olan değişken olarak tanımlanmaktadır.

Modelde bağımlı değişkeni etkileyen bir veya daha çok bağımsız değişken yer alabilir. Birinci durumda 2 değişkenli, ikinci durumda ise çok değişkenli regresyon modeli söz konusu olur.

Bağımlı ve bağımsız değişkenler arasındaki ilişki doğrusal ve doğrusal olmayan regresyon modelleriyle açıklanmakta ve doğrusal olmayan regresyon modelleri de dönüşüm ile doğrusallaştırılabilen ve doğrusallaştırılamayan modeller olarak ikiye ayrılmaktadır. Doğrusallığa parametreler ve değişkenler açısından yaklaşılmaktadır. Bir modelin parametreler veya değişkenler açısından doğrusallık koşulu, modeldeki her parametrenin (değişkenin) üssünün 1 olması, parametrelerin (değişkenlerin) üs olarak yer almaması, diğer parametreler (değişkenler) ile çarpılıp bölünmemesidir.²

Buna göre;

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u \quad (2.1)$$

modeli hem parametreler hem de değişkenler açısından doğrusaldır. Buna karşılık aşağıdaki model parametreler açısından doğrusal, değişkenler açısından doğrusal değildir.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_1^2 + u \quad (2.2)$$

¹ Alpar, R., Uygulamalı Çok Değişkenli İstatistiksel Yöntemlere Giriş 1, s.189.

² Genceli, M., Ekonometri ve İstatistik İlkeleri, s.20

Ekonometride doğrusal model denilince parametreler açısından doğrusallık anlaşılmalıdır.³ Doğrusal ve doğrusal olmayan regresyon modellerinin kullanımına ilişkin aşağıdaki çizelge incelenebilir.

Çizelge 2.1 Doğrusal ve doğrusal olmayan regresyon modelinin belirlenmesi

Model parametrelerinde doğrusal mıdır?	Model değişkenlerinde doğrusal mıdır?	
	<i>Evet</i>	<i>Hayır</i>
<i>Evet</i>	DRM	DRM
<i>Hayır</i>	DORM	DORM

Doğrusal modeldeki denklemlerin çözümünün kolaylığı, parametrelerin değerlerinin bulunmasındaki basitlik bu çeşit modellerin geniş ölçüde kullanılmasına sebep olmuştur. Fakat gerçek iktisadi hayatta değişkenler arasındaki bağlantıları her zaman doğrusal denklemlerle ifade etmek mümkün olmaz.⁴

2.2 İki Değişkenli Regresyon Modeli

Regresyon analizinde değişkenler arasındaki bağımlılık fonksiyonel veya deterministik değil, istatistikseldir. Değişkenler arasındaki istatistiksel ilişkilerde olasılık dağılımlarına sahip tesadüfi veya stokastik değişkenler söz konusu olur. İki değişkenli regresyon modelinde bağımsız X değişkeni ile bağımlı Y değişkeni arasındaki ilişki açıklanmaktadır. Bağımsız değişken X_i nin değerlerinin sabit olması, aynı X_i için, ana kütlede, rastlantısal ilişki gereği en azından iki bağımlı değişkenin bulunmasını da zorunlu kılmaktadır. Böylece herhangi bir X_i değişkeni için Y_{ij} değerleri elde edilecek, bu da bir dağılım oluşturacaktır. X_i değişkenine karşılık gelen Y_{ij} değişkenleri alt ana kütle adı verilen gruplarda toplandığında çizelge(2.2) elde edilebilir.

³ Genceli, M., a.g.e , s.21

⁴ Kılıçbay, A., Ekonometri, s.146

Çizelge 2.2 Bağımlı ve bağımsız değişkenlerin dağılımı

$X \longrightarrow$	X_1	X_2	X_3	X_k	
Y	Y_{11}	Y_{21}	Y_{31}	Y_{k1}	
	Y_{12}	Y_{22}	Y_{32}	Y_{k2}	
	.	.	.	.	
$\downarrow$	.	.	.	.	
	Y_{1N_1}	Y_{2N_2}	Y_{3N_3}	Y_{kN_k}	
Toplam	$\sum_j Y_{1j}$	$\sum_j Y_{2j}$	$\sum_j Y_{3j}$	$\sum_j Y_{kj}$	$\sum_j \sum_i Y_{ij}$
Birim Mevcudu	N_1	N_2	N_3	N_k	$\sum N_i = N$
μ_{YX_k}	μ_{YX_1}	μ_{YX_2}	μ_{YX_3}	μ_{YX_k}	

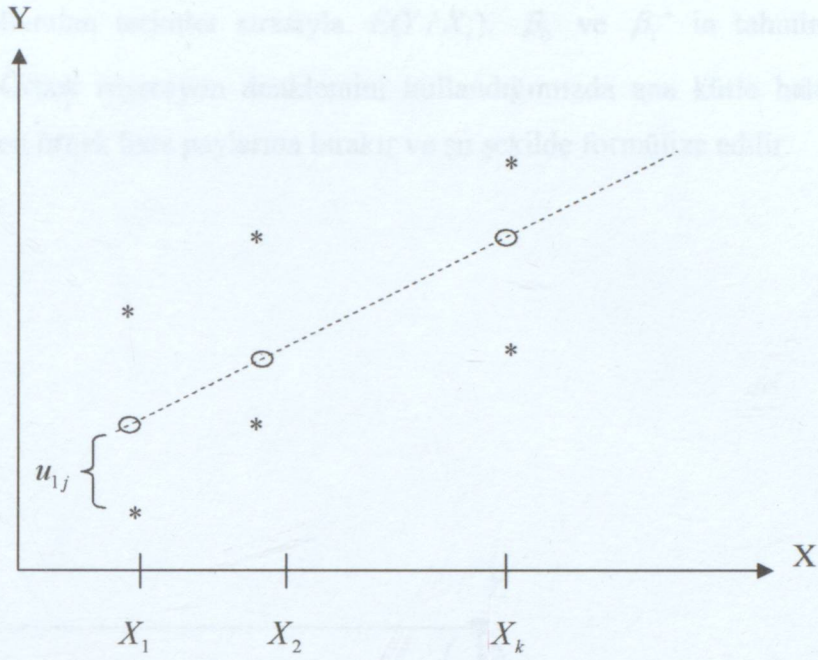
Her alt ana kütleyle ilişkin ortalama değerlerin hesaplanması aşağıda gösterilmektedir.

$$E(Y_k / X_k) = \mu_{YX_k} = \frac{1}{N_k} [(Y_{k1} / X_k) + (Y_{k2} / X_k) + \dots + (Y_{kj} / X_k)] \quad (2.3)$$

Birinci alt ana kütle için j . biriminin kendi alt ana kütlelerinin aritmetik ortalaması ile arasındaki bağıntı şöyle gösterilir.

$$Y_{1j} = \mu_{YX_1} + u_{1j}, j = 1, 2, \dots, N_1 \quad (2.4)$$

Birimler kendi alt ana kütle ortalamaları etrafında bir dağılım göstermektedirler. Daha kolay anlaşılması açısından bu terimler Şekil 2.1 de gösterilmektedir. Şekilde yer alan * simgesi Y_{ij} ve o simgesi ise μ_{YX_i} terimlerini ifade etmektedir.



Şekil 2.1 Anakütle regresyon doğrusu

Bağımlı değişken Y ile bağımsız değişken arasındaki doğrusal, rastlantısal bağıntıyı şu şekilde gösterelim:

$$Y_{ij} = \beta_0 + \beta_1 X_i + u_{ij} \quad (2.5)$$

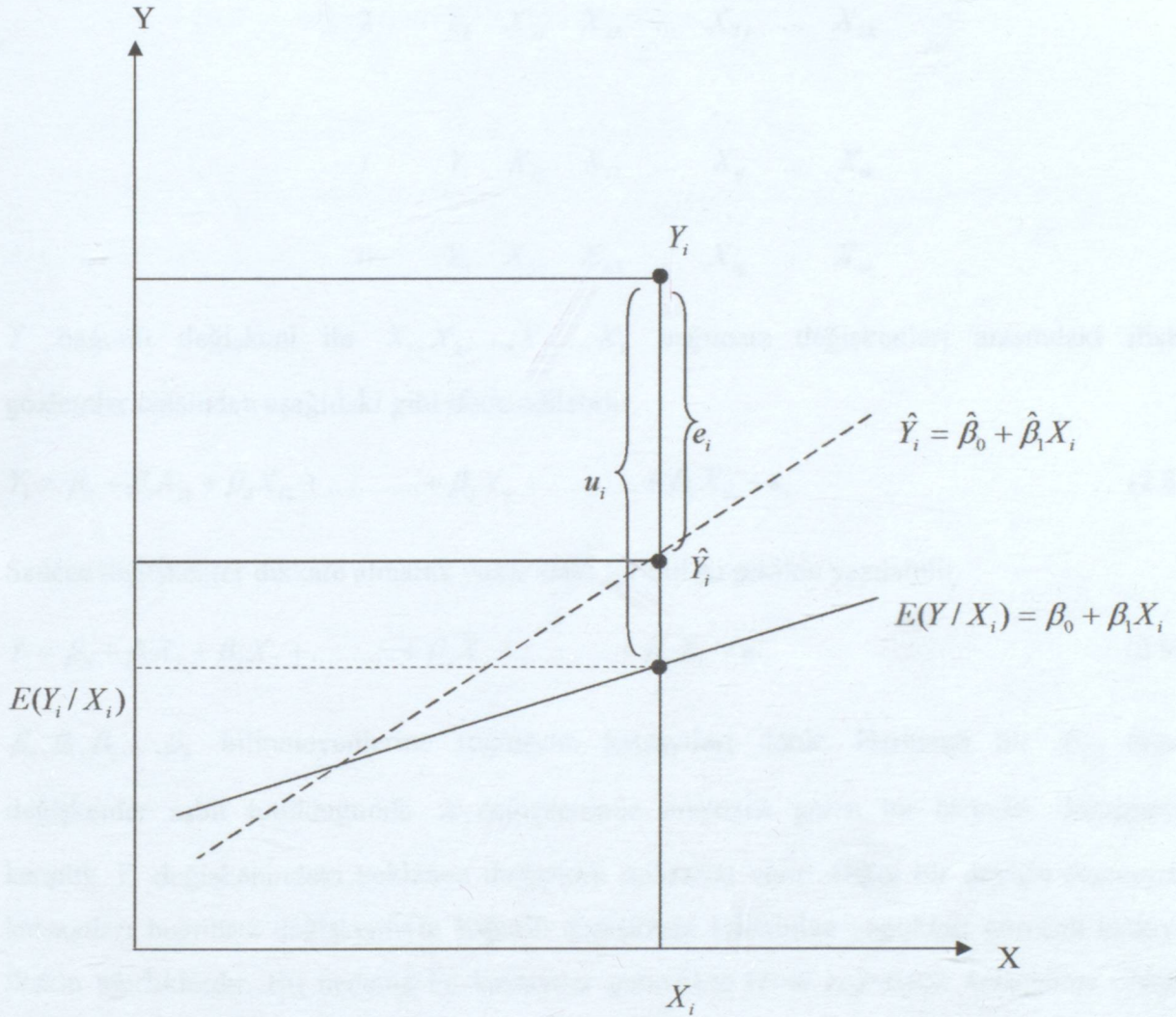
Modeli $(\beta_0 + \beta_1 X_i)$ ile tanımlanan sistematik kısım ve u_{ij} ile gösterilen ana kütle hata payları olarak iki bölümde inceleyelim. Sistematik kısım alt ana kütle ortalamalarından geçen bir doğru olup ana kütle regresyon doğrusu olarak bilinmektedir. β_0 , doğrunun ordinat eksenini (Y eksenini) ile kesiştiği noktayı, β_1 ise doğrunun eğimini vermektedir. β_1 aynı zamanda bağımsız değişken olan X 'teki bir birim değişme karşısında $E(Y/X)$ 'teki değişme miktarını, yani marjinal değişmeyi göstermektedir. u_{ij} 'de Y_{ij} birimlerinin kendi alt ana kütle ortalamalarından artı veya eksi sapmalarını göstermektedir.

Uygulamalarda ana kütlelerin gözlemlenmesi yerine çoğu kez bir örnekten hareket edilmektedir. Bu yüzden, örneklem bilgisine dayanarak ana kütle regresyon fonksiyonunu tahmin etmek durumundayız. Ana kütle regresyon denklemi yerine aşağıda gösterilen örnek regresyon denklemini kullanacağız.

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i \quad (2.6)$$

Denkleimde kullanılan terimler sırasıyla $E(Y/X_i)$, β_0 ve β_1 ' in tahmin edicisi olarak bilinmektedir. Örnek regresyon denklemini kullandığımızda ana kütle hata payları yerini kalıntı adı verilen örnek hata paylarına bırakır ve şu şekilde formülize edilir.

$$e_i = Y_i - \hat{Y}_i \quad (2.7)$$



Şekil 2.2 Anakütle ve örnek regresyon doğruları

2.3 Çoklu Doğrusal Regresyon

Çoklu regresyon modelleri, bir bağımlı değişken ile birden çok bağımsız değişken arasındaki ilişkilerin incelendiği ve aşağıda verilen 2 önemli amacı taşıyan modellerdir.

- 1- Bağımsız değişkenler yardımıyla bağımlı değişkeni tahmin etmek;
- 2- Bağımsız değişkenlerden hangisi ya da hangilerinin bağımlı değişkeni daha çok etkilediğini bulmak ve aralarındaki karmaşık yapıyı tanımlamak.

Çoklu regresyon modellerinde, bağımlı değişken Y , bağımsız değişkenler $X_1, X_2, X_3, \dots, X_k$ ile gösterilir. Verilere ilişkin genel gösterim aşağıdaki çizelgede yer almaktadır.

Çizelge 2.3 Çoklu regresyon modelinde verilerin gösterimi

<u>Gözlem</u>	<u>Y</u>	<u>X_1</u>	<u>X_2</u>	...	<u>X_j</u>	...	<u>X_k</u>
1	Y_1	X_{11}	X_{12}	...	X_{1j}	...	X_{1k}
2	Y_2	X_{21}	X_{22}	...	X_{2j}	...	X_{2k}
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
i	Y_i	X_{i1}	X_{i2}	...	X_{ij}	...	X_{ik}
.	.	.	.	.	.	.	.
n	Y_n	X_{n1}	X_{n2}	.	X_{nj}	.	X_{nk}

Y bağımlı değişkeni ile $X_1, X_2, \dots, X_j, \dots, X_k$ bağımsız değişkenleri arasındaki ilişki gözlemler cinsinden aşağıdaki gibi ifade edilebilir.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_j X_{ij} + \dots + \beta_k X_{ik} + u_i \quad (2.8)$$

Sadece değişkenler dikkate alınarak yukarıdaki formül şu şekilde yazılabilir.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_j X_j + \dots + \beta_k X_k + u \quad (2.9)$$

$\beta_0, \beta_1, \beta_2, \dots, \beta_k$ bilinmeyenlerine regresyon katsayıları denir. Herhangi bir β_j , diğer değişkenler sabit tutulduğunda X_j değişkeninde meydana gelen bir birimlik değişmeye karşılık Y değişkenindeki beklenen değişiklik miktarını verir. Diğer bir deyişle regresyon katsayıları bağımsız değişkenlerin bağımlı değişkenin tahminine yaptıkları göreceli katkıya ilişkin ağırlıklardır. Bu nedenle bu katsayılar genellikle *kısmî regresyon katsayıları* olarak adlandırılır.⁵ Bağımlı ve bağımsız tüm değişkenleri standartlaştırıp elde edilen standartlaştırılmış regresyon katsayılarının büyüklüğünü incelemek değişkenlerin modele yaptıkları katkı konusunda daha kolay yorum yapmamızı sağlar. Bu durumda standartlaştırılmış regresyon katsayılarının en büyüğüne ilişkin değişkenin, modele en fazla katkıyı yapan değişken olduğuna karar verilir. Standartlaştırılmış regresyon katsayılarına BETA katsayıları da denir. Hesaplanması (2.10)' da gösterilmektedir.⁶

⁵ Alpar, R., a.g.e, s.231

⁶ Alpar, R., a.g.e, s.239

$$b_j^* = b_j \frac{s_j}{s_y} \quad (2.10)$$

b_j^* = standartlaştırılmış regresyon katsayısı

s_j = j.değişkene ilişkin standart sapma

s_y = bağımlı değişkene ilişkin standart sapma

2.4 Basit Doğrusal Regresyonun Matris Yardımıyla Çözümü

$Y_i = \beta_0 + \beta_1 X_{i1} + u_i$ şeklinde verilen doğrusal regresyon modeli, matris yaklaşımı ile, n gözlem alınarak aşağıdaki gibi çözümlenir.

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} \\ 1 & X_{21} \\ \vdots & \vdots \\ 1 & X_{n1} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

$$Y = X\beta + u \quad (2.11)$$

Y = (n*1) boyutlu bağımlı değişken vektörü

X = (n*2) boyutlu ilk sütunu 1'lerden, 2.sütunu X 'lerden oluşan matris(veri matrisi)

β = (2*1) boyutlu bilinmeyen katsayılar vektörü

u = (n*1) boyutlu hata vektörü

β katsayılarının tahmininde kullanılan matris gösterimi şu biçimde olmaktadır.

$$\hat{\beta} = (X'X)^{-1} X'Y \quad (2.12)$$

Bu değerlerden hareketle bağımlı değişkene ilişkin tahmin değerleri aşağıdaki gibi elde edilmektedir.

$$\hat{Y} = X\hat{\beta} \quad (2.13)$$

2.5 Çok Değişkenli Doğrusal Regresyonun Matrisle Çözümü

Y bağımlı değişkeni ile k tane açıklayıcı değişken içeren çok değişkenli anakütle regresyon doğrusu denklemini yazalım.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_j X_{ij} + \dots + \beta_k X_{ik} + u_i \quad i = 1, 2, 3, \dots, n \quad (2.14)$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1k} \\ 1 & X_{21} & X_{22} & \dots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \dots & X_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \longrightarrow Y = X\beta + u$$

Hem basit, hem de çoklu doğrusal regresyon çözümlemeleri aynı yolla yapılmaktadır, tek farklılık veri matrisi ve katsayılar matrisinin boyutunda olmaktadır.

3. GÖLGE DEĞİŞKEN

Regresyon analizinde bağımlı değişken, sadece rakamla ifade edilebilen nicel, kantitatif değişkenler (gelir, fiyat, sıcaklık) tarafından etkilenmez, aynı zamanda ırk, cinsiyet, renk, din, uyruk, savaşlar, grevler, siyasi çalkantılar, hükümetin iktisat politikasındaki değişmeler gibi nitel, kalitatif değişkenlerden de etkilenir. Örneğin tüm diğer faktörler sabit iken, erkekler kadınlara oranla daha yüksek gelirlidir, bu cinsiyet değişkeni olarak ortaya çıkar. Ya da savaş yıllarında barış yıllarına nazaran tüketim azalır, bu da savaş değişkeni olarak ortaya çıkar. Böylece cinsiyet, savaş gibi kalitatif değişkenler bağımlı değişkeni etkilerler ve açıklayıcı, bağımsız değişken olarak modele alınırlar.⁷ Bu tür nitel değişkenler genellikle istenilen özellik veya kalitenin varlık ya da yokluğunu ifade ederler. Bu özellikleri nicel hale getirmek için 0 ya da 1 değerlerini alan yapay değişkenler oluşturmak gerekmektedir. Örneğin 1 değeri kişinin erkek olduğunu, 0 değeri kişinin kadın olduğunu ya da 1 savaş yılını 0 ise barış yılını gösterir. Böylece 1 ve 0 değerlerini alan değişkenlere **gölge değişken** denir. Gölge değişkene alternatif olarak kukla, gösterge, iki değerli, kategorik ya da dikotomi değişken adları da verilmektedir.⁸ Bazı durumlarda nicel değişkenler de gölge değişken olarak modele alınabilmektedir. Örneğin, geliri 500 YTL ve altında olanlarla, bu değer üzerinde olanlar diye tüketiciler 2 gruba ayrılarak gölge değişken olarak modelde yer alabilirler. Ya da bireyler 20-35 yaş arasında olanlar ve 35 yaştan büyük olanlar şeklinde iki gruba ayrılarak, tüketim ve tasarruf kalıplarının belirlenmesinde yaş değişkeni gölge değişken olarak kullanılabilir.⁹

3.1 Varyans ve Kovaryans Analizi Modelleri

Bir modelde bağımsız değişken(ler) sadece gölge değişken(ler) den oluşuyorsa bu modele **varyans analizi modeli** (ANOVA-Analysis of Variance) denir.¹⁰ Erkek ve kadınlar için sigara tüketim harcaması modelini gölge değişken yardımıyla aşağıdaki gibi kuralım:

$$Y_i = \beta_0 + \beta_1 D_i + u_i \quad (3.1)$$

Y_i sigara tüketim harcaması

$D_i = 1$ eğer kişi erkek ise

$D_i = 0$ eğer kişi kadın ise

⁷ Frank, R. C., Statistics and Econometrics, s.164

⁸ Gujarati, D. N., Essentials of Econometrics, s.251

⁹ Koutsoyiannis, A., Ekonometri Kuramı, Ekonometri Yöntemlerinin Tanıtımına Giriş, s.284

¹⁰ Gujarati, D. N., a.g.e, s.252

Gölge değişkenimiz modelde D harfi ile simgelenmektedir ve cinsiyeti ifade etmektedir. Kurulan model, cinsiyetin sigara tüketimi harcamasında rolü olup olmadığını, yani erkeklerle kadınlar arasında sigara tüketimi harcaması bakımından farklılık olup olmadığını anlamamızda yardımcı olmaktadır. Modeldeki β_0 parametresi kadınların ortalama sigara tüketim harcamasını göstermektedir.

$$E(Y_i / D_i = 0) = \beta_0 + \beta_1 \cdot (0) = \beta_0, \quad E(u_i) = 0 \quad (3.2)$$

Gölge değişken olan D_i 'nin katsayısı β_1 ise erkeklerin sigara tüketim harcamasının kadınlarınkinden farkını göstermektedir. β_1 pozitifse erkeklerin kadınlardan β_1 kadar fazla sigara harcaması yaptıklarını, negatif ise β_1 kadar daha az sigara harcaması yaptıklarını söyleriz.

Erkeklerin ortalama sigara tüketim harcaması ise şöyle bulunur:

$$E(Y_i / D_i = 1) = \beta_0 + \beta_1 \cdot (1) = \beta_0 + \beta_1 \quad (3.3)$$

β_1 ' in sıfırdan farklı olup olmadığı ($H_0 : \beta_1 = 0$) t testiyle test edilerek tüketim harcamasında kadın erkek farkının olup olmadığı ortaya konabilir.

n birimlik bir örnek söz konusu olsun. $D = 1$ olanların yani erkeklerin sayısı n_1 , $D = 0$ yani kadınların sayısı n_2 ve $n = n_1 + n_2$ ile gösterilsin. $Y_i = \beta_0 + \beta_1 D_i + u_i$ şeklinde kurulan modelimizde parametre tahminlerini yapmak için en küçük kareler yöntemindeki normal denklemleri kullanarak şu sonuçlara ulaşırız.

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 D \quad (3.4)$$

$$\left. \begin{aligned} I. \sum Y &= n\hat{\beta}_0 + \hat{\beta}_1 \sum D \\ II. \sum YD &= \hat{\beta}_0 \sum D + \hat{\beta}_1 \sum D^2 \end{aligned} \right\} \text{Normal Denklemler} \quad (3.5)$$

Gölge değişkenin 0 ya da 1 değerlerini almasından dolayı $\sum D = \sum D^2$ olur ve böylece II numaralı denklem $\sum YD = \hat{\beta}_0 \sum D + \hat{\beta}_1 \sum D$ olarak düzenlenip $\sum D = n_1$ ve $\sum YD$ 'nin de birinci şıktakilerin toplamını (erkeklerin sigara tüketim harcaması toplamını) verdiği göz önüne alınırsa şu sonuçlar kolaylıkla çıkartılabilir:

$$\sum YD = n_1(\hat{\beta}_0 + \hat{\beta}_1) \quad (3.6)$$

$$Y_{1n_1} + Y_{2n_1} + \dots + Y_{n_1n_1} = n_1(\hat{\beta}_0 + \hat{\beta}_1) \quad (3.7)$$

$$\bar{Y}_{n_1} = \hat{\beta}_0 + \hat{\beta}_1 \quad (3.8)$$

Y_{1n_1} , gözlem sayısı n_1 olan gruptaki 1.gözlemi, $\bar{Y}_{n_1}$ ise bağımlı değişkenin ortalama değerini, yani erkeklerin ortalama sigara tüketim harcamasını göstermektedir. Açıklamalar doğrultusunda normal denklemler yeniden düzenlenerek aşağıdaki biçimde tekrar yazılırsa:

$$\left. \begin{array}{l} I. \sum Y = n\hat{\beta}_0 + n_1\hat{\beta}_1 \\ II. \sum YD = n_1\hat{\beta}_0 + n_1\hat{\beta}_1 \end{array} \right\} \text{(I)-(II) yani I denkleminde II denklemini çıkarılırsa}$$

$$\sum Y - \sum YD = n\hat{\beta}_0 - n_1\hat{\beta}_0 = \hat{\beta}_0(n - n_1)$$

$$(\cancel{Y_{1n_1} + Y_{2n_1} + \dots + Y_{n_1n_1}} + Y_{1n_2} + Y_{2n_2} + \dots + Y_{n_2n_2}) - (\cancel{Y_{1n_1} + Y_{2n_1} + \dots + Y_{n_1n_1}}) = \hat{\beta}_0 n_2$$

$$\sum Y_{n_2} = n_2\hat{\beta}_0 \quad \Rightarrow \quad \bar{Y}_{n_2} = \hat{\beta}_0 \text{ olarak elde edilir.}$$

Sonuç olarak β_0 , $D = 0$ alınan şık için, $(\beta_0 + \beta_1)$ ise $D = 1$ alınan şık için ortalamayı göstermektedir.

Sıfır değerinin atandığı grup, kategori veya sınıfa temel, sabit nokta, kontrol, karşılaştırma, referans veya ihmal edilen kategori denilmektedir. Karşılaştırmalar bu kategoriye göre yapılmaktadır. Kurmuş olduğumuz modelde kadınlar taban kategori olarak ele alınmıştır.

Sadece gölge açıklayıcı değişkenli, yani açıklayıcı değişken(ler)in gölge değişken olduğu varyans analizi modelleri sosyoloji, psikoloji, eğitim ve pazar araştırması alanlarında kullanılmaktadır. Ekonometrik çalışmaların çoğunda, nitel ve nicel değişkenler birlikte yer almaktadır. Nitel ve nicel değişkenlerin bir karmasını içeren regresyon modellerine **kovaryans analizi modelleri** (ANCOVA-Analysis of Covariance) denir.¹¹ Çalışan kişilerin yıllık ücretlerinin yıl bazında iş deneyimine ve kişinin cinsiyetine göre bağıllığını incelemek için aşağıdaki modeli kuralım.

$$Y_i = \beta_0 + \beta_1 D_i + \beta_2 X_i + u_i \quad (3.9)$$

¹¹ Gujarati, D. N. , a.g.e , s.255

Y_i , çalışan kişinin aldığı yıllık ücret

X_i , yıl olarak kişinin iş deneyimi

$D_i = 1$ eğer kişi erkek ise

$D_i = 0$ eğer kişi kadın ise

Bir nicel bir de nitel bağımsız değişkenin yer aldığı bir model kurulmuştur. Kadın ve erkek çalışanların ortalama yıllık ücretleri hesaplanırsa aşağıdaki sonuçlar elde edilmektedir.

Kadın çalışanların ortalama yıllık ücreti;

$$E(Y_i / D_i = 0, X_i) = \beta_0 + \beta_2 X_i \quad (3.10)$$

Erkek çalışanların ortalama yıllık ücreti;

$$E(Y_i / D_i = 1, X_i) = \beta_0 + \beta_1 + \beta_2 X_i \quad (3.11)$$

Erkek çalışanların ortalama ücret düzeyinin kadın çalışanların ortalama ücret düzeyinden β_1 kadar farklı olduğu fakat deneyim yılları ile ortalama ücretteki değişim oranının (yani β_2) her iki cinsiyet için de aynı olduğu görülmektedir.

Eğer ortak eğim (β_2 ' lerin eşitliği) varsayımı geçerliyse erkek ve kadın çalışanlar için oluşturulan regresyon denklemlerinde sabit terimlerin aynı olduğu yani ücretlerde cinsiyet farkının olup olmadığı test edilebilir. β_1 ' in istatistiksel anlamlılığı t testi ile sınanır ve anlamlı olduğu bulunursa şu şekilde sonuca varılır: "Kadın ve erkek çalışanların ortalama yıllık ücretlerinin aynı olduğunu savunan $H_0 : \hat{\beta}_1 = 0$ hipotezini kabul etmemiz için yeterli bir neden yoktur, yıllık ücret cinsiyete göre değişim göstermektedir."

3.2 Gölge Değişken Tuzağı

Regresyon analizinde gölge değişken kullanımında çoklu doğrusal bağlantı (açıklayıcı değişkenler arasındaki doğrusal ilişkinin varlığı) sorunundan kaçınmak için uyulması gereken bir kural vardır. Kural gereğince sabit terimin yer aldığı bir regresyon modelinde nitel değişkenimize ait k adet kategori, sınıf varsa $k - 1$ adet gölge değişken kullanmak gerekir.¹² Aksi takdirde tam çoklu doğrusal bağlantı sorunu ortaya çıkar ve gölge değişken tuzağına düşülür. Kurmuş olduğumuz modellerde cinsiyetin 2 ($k = 2$) şıkkı olduğundan 1

¹² Watsham, T. J. ve Parromore, K., Quantitative Methods in Finance, s.210

$(k - 1 = 1)$ adet gölge değişken kullandık. Modelde sabit terim yer almadığında ise k adet gölge değişken kullanılır ve orijinden geçen regresyon yöntemi kullanılarak parametre tahminleri yapılır. Kurmuş olduğumuz ilk modeli sabit terim olmadan yeniden kurduğumuzu düşünelim.

$$\hat{Y}_i = \hat{\beta}_1 D_1 + \hat{\beta}_2 D_2 \quad (3.12)$$

$$D_1 = 1, \quad \text{erkek ise} \quad D_2 = 1, \quad \text{kadın ise}$$

$$D_1 = 0, \quad \text{diğer} \quad D_2 = 0, \quad \text{diğer}$$

Bu modelde $\hat{\beta}_1$ erkeklerin ortalama sigara tüketim harcamasını, $\hat{\beta}_2$ ise bayanların ortalama sigara tüketim harcamasını vermektedir.

Gölge değişken tuzağına açıklama getirmek açısından bir örnek verelim. Kurmuş olduğumuz ikinci modelde nitel değişkenimiz erkek ve kadın olarak 2 kategoriye sahiptir ve bu durumda 1 gölge değişken kullanmamız gerekirken (modelde sabit terim varken) 2 gölge değişken kullanarak modelimizi yeniden kuralım.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 D_1 + \beta_3 D_2 + u_i \quad (3.13)$$

$$D_1 = 1, \quad \text{erkek ise} \quad D_2 = 1, \quad \text{kadın ise}$$

$$D_1 = 0, \quad \text{diğer} \quad D_2 = 0, \quad \text{diğer}$$

4 gözlemimiz olduğunu ve ilk 2'sinin erkek diğerlerinin de kadın olduğunu düşünerek matris şeklindeki şu ifadeyi yazalım.

$$X = \begin{array}{c} \underline{X} \quad \underline{D_1} \quad \underline{D_2} \\ \left| \begin{array}{cccc} 1 & X_{11} & 1 & 0 \\ 1 & X_{21} & 1 & 0 \\ 1 & X_{31} & 0 & 1 \\ 1 & X_{41} & 0 & 1 \end{array} \right| \end{array}$$

İlk sütunun, son iki sütunun toplamı şeklinde olduğunu görmekteyiz ve bu yüzden de sütunlar arasında doğrusal bağımlılıktan söz edebiliriz.

$$X'X = \begin{vmatrix} 1 & 1 & 1 & 1 \\ X_{11} & X_{21} & X_{31} & X_{41} \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{vmatrix} \begin{vmatrix} 1 & X_{11} & 1 & 0 \\ 1 & X_{21} & 1 & 0 \\ 1 & X_{31} & 0 & 1 \\ 1 & X_{41} & 0 & 1 \end{vmatrix} = \begin{vmatrix} 4 & \sum_{i=1}^4 X_{i1} & 2 & 2 \\ \sum_{i=1}^4 X_{i1} & \sum_{i=1}^4 X_{i1}^2 & \sum_{i=1}^2 X_{i1} & \sum_{i=3}^4 X_{i1} \\ 2 & \sum_{i=1}^2 X_{i1} & 2 & 0 \\ 2 & \sum_{i=3}^4 X_{i1} & 0 & 2 \end{vmatrix}$$

1.sütunun 3. ve 4. sütunların toplamı şeklinde yazılabildiği ve böylece sütunlar arasında doğrusal bağımlılık olduğu görülmektedir. Matris bilgilerinden hatırlanacağı üzere böyle bir durumda matrisin determinantı 0'a eşit olmakta ve matrisin tersi alınamamaktadır. Sonuç olarak matris şeklinde gösterimde $\hat{\beta} = (X'X)^{-1}(X'Y)$ olarak hesaplandığından, $\hat{\beta}$ 'lar hesaplanamamakta ve çoklu doğrusal bağlantının yol açtığı sorun açıkça görülmektedir. Bu sorundan kaçınmak için modelimizde 1 adet gölge değişken kullanmalıydık.

3.3 Regresyon Modellerinin F Testleri Yardımıyla Karşılaştırılması

Buraya kadar incelediğimiz modellerde eğim katsayısının değişmediğini varsaydık. İki ya da daha çok regresyonun farklı olup olmadığını, yani regresyonların sabit terimlerinin, eğim katsayılarının ya da her ikisinin birden farklı olup olmadığını gölge değişken yardımıyla bulabiliriz. 2 regresyon doğrusunu düşündüğümüzde 4 farklı durumla karşılaşmaktadır.

- 1- Hem sabit hem de eğimler eşittir.(çakışan regresyonlar)
- 2- Yalnız sabit terimler farklıdır.(koşut,paralel regresyonlar)
- 3- Yalnız eğimler farklıdır.(uyumlu regresyonlar)
- 4- Hem sabit hem de eğimler farklıdır.(benzemez regresyonlar)

Bir tane nicel bağımsız değişken ve bir tane de 2 şıklı gölge değişkenle aşağıdaki modeli kuralım.

$$Y_i = \beta_0 + \beta_1 D_i + \beta_2 X_i + \beta_3 (D_i X_i) + u_i \quad (3.14)$$

Gölge değişkenin 0 ya da 1 değerlerini aldığını düşünürsek aşağıdaki ifadeleri elde ederiz.

$$E(Y_i / D_i = 0, X_i) = \beta_0 + \beta_2 X_i \quad (3.15)$$

$$E(Y_i / D_i = 1, X_i) = (\beta_0 + \beta_1) + (\beta_2 + \beta_3) X_i \quad (3.16)$$

β_1 sabit terim farkını, β_3 ise eğim farkı katsayısını göstermektedir. Regresyon modelinde bu katsayıların anlamlılıkları test edilerek gölge değişkenin her iki şıkkı için kurulan regresyon modellerinin birbirinden farklı olup olmadıkları, diğer bir ifadeyle farklılık varsa bu farklılığın sabit terimden mi, eğimden mi yoksa her ikisinden birden mi kaynaklandığını bulabiliriz. Bunun tespiti için aşağıda yer alan hipotezleri kurup F testi yardımıyla belirli anlamlılık düzeylerinde test etmemiz gerekmektedir. Öncelikle F testlerinin genel kullanımına bir göz atalım.

Kısmî F testi

Tam model : $Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + u_i$ biçiminde kurulsun.

$$\left. \begin{array}{l} H_0 : \beta_3 = 0 \\ H_1 : \beta_3 \neq 0 \end{array} \right\} \text{hipotezini test etmek isteyelim.}$$

İndirgenmiş model : $Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u_i$ biçiminde olur.

F testinde kullanacağımız değer ise şu şekilde bulunur ve tablo değeri ile karşılaştırılarak hipotezin reddine ya da kabulüne karar verilir. Öncelikle hesaplamalarda kullanılacak simgelerin açıklamaları yapılacaktır.

toplam değişkenlik = regresyonla açıklanan değişkenlik + açıklanamayan değişkenlik

$$(Y_i - \bar{Y}) = (\hat{Y}_i - \bar{Y}) + (Y_i - \hat{Y}_i) \quad (3.17)$$

$$\sum (Y_i - \bar{Y})^2 = \sum (\hat{Y}_i - \bar{Y})^2 + \sum (Y_i - \hat{Y}_i)^2 \quad (3.18)$$

$$SSTO = SSR + SSE \quad (3.19)$$

3 bağımsız değişkenin yer aldığı modeli düşünelim. Yalnızca X_1 değişkeni ile Y arasında regresyon çözümlemesi yapıldığında $SSR(X_1)$ elde edilir ve bu durumda hata kareleri toplamı $SSE(X_1)$ olur. Modele X_2 değişkeninin de eklenmesiyle elde edilecek olan $SSE(X_1, X_2)$, $SSE(X_1)$ değerinden düşük olacaktır. Aradaki fark $SSR(X_2 / X_1)$ olarak gösterilmekte ve X_1 değişkeni modeldeyken X_2 değişkeninin modele eklenmesiyle oluşacak etkiyi göstermektedir.¹³

¹³ Neter vd., Applied Linear Statistical Models, s.261

$$F^* = \frac{SSE(R) - SSE(F)}{df_R - df_F} \bigg/ \frac{SSE(F)}{df_F} \quad (3.20)$$

$$F^* = \frac{SSE(X_1, X_2) - SSE(X_1, X_2, X_3)}{(n-3) - (n-4)} \bigg/ \frac{SSE(X_1, X_2, X_3)}{(n-4)} \quad (3.21)$$

$SSE(X_1, X_2) - SSE(X_1, X_2, X_3) = SSR(X_3 / X_1, X_2)$ olduğu göz önüne alınır;

$$F^* = \frac{SSR(X_3 / X_1, X_2)}{1} \bigg/ \frac{SSE(X_1, X_2, X_3)}{(n-4)} \quad (3.22)$$

olarak elde edilir.

F testi

Tam model aynı şekilde kurulmuş olsun.

$$H_0 : \beta_2 = \beta_3 = 0$$

$$H_1 : \text{en azından biri 0'dan farklıdır.}$$

} hipotezini test etmek isteyelim.

İndirgenmiş model : $Y_i = \beta_0 + \beta_1 X_1 + u_i$ şeklinde olur. Kısmi F testinde kullanılan formülden hareket edecek olursak;

$$F^* = \frac{SSE(X_1) - SSE(X_1, X_2, X_3)}{(n-2) - (n-4)} \bigg/ \frac{SSE(X_1, X_2, X_3)}{(n-4)} \quad (3.23)$$

olarak bulunur.

$SSE(X_1) - SSE(X_1, X_2, X_3) = SSR(X_2, X_3 / X_1)$ olduğu göz önüne alınır;

$$F^* = \frac{SSR(X_2, X_3 / X_1)}{2} \bigg/ \frac{SSE(X_1, X_2, X_3)}{(n-4)} \quad (3.24)$$

formülü elde edilir.

Modelimizde bu testleri uygulayacak olursak;

$$\text{Tam model} : Y_i = \beta_0 + \beta_1 D + \beta_2 X_1 + \beta_3 DX_1 + u_i$$

$$H_0 : \beta_3 = 0$$

$$H_1 : \beta_3 \neq 0$$

İndirgenmiş model : $Y_i = \beta_0 + \beta_1 D + \beta_2 X_1 + u_i$ şeklinde kurulur.

$$F^* = \frac{SSR(DX_1 / X_1, D)}{1} \bigg/ \frac{SSE(D, X_1, DX_1)}{(n-4)} \quad (3.25)$$

Hesaplanan bu değer, 0,95 anlamlılık düzeyinde tabloda yer alan $F_{(0,95;1;n-4)}$ değeri ile karşılaştırılır ve $F^* < F_{TABLO} \Rightarrow H_0$ kabul edilir, diğer bir ifadeyle eğimlerin aynı olduğu söylenir.

Bu sefer de aşağıdaki hipotezleri test etmek isteyelim:

$$H_0 : \beta_1 = \beta_3 = 0$$

H_1 : en azından biri 0'dan farklıdır.

$$\text{Tam model} : Y_i = \beta_0 + \beta_1 D + \beta_2 X_1 + \beta_3 DX_1 + u_i$$

$$\text{İndirgenmiş model} : Y_i = \beta_0 + \beta_2 X_1 + u_i$$

$$F^* = \frac{SSR(D, DX_1 / X_1)}{2} \bigg/ \frac{SSE(D, X_1, DX_1)}{(n-4)} \quad (3.26)$$

Hesaplanan bu değer , 0,95 anlamlılık düzeyinde tabloda yer alan $F_{(0,95;2;n-4)}$ değeri ile karşılaştırılır ve $F^* < F_{TABLO} \Rightarrow H_0$ kabul edilir, diğer bir ifadeyle hem eğimlerin hem de gölge değişken katsayılarının aynı olduğu söylenir.

4. NİTEL TERCİH MODELLERİ

Y bağımlı değişkeninin nicel, bağımsız değişken veya değişkenlerin ise ya nicel, ya nitel ya da bunların karışımı yani bazısının nicel bazısının nitel değişken olarak kullanıldıkları modeller önceki bölümlerde anlatıldı. İki veya daha fazla değer alan gölge değişkenler bağımlı değişken olarak da regresyon modellerinde yer alabilirler. Bağımlı değişkenin nitel özellikte olduğu nitel tercih modelleri olarak da bilinen gölge bağımlı değişkenli modeller 2 grup halinde ayrıntılı bir biçimde incelenecektir. Bu tür modellerde bağımlı değişken iki değer alıyorsa değişken evet-hayır, başarılı-başarısız, olumlu-olumsuz, satın alma-almama, borsa ile ilgilenip ilgilenmeme gibi tercih veya karar belirtir. Bu nedenle bu modeller ikili tercih modelleri olarak adlandırılırlar. Bağımlı değişken ikiden fazla değer de alabilir. Bağımlı değişkenin evet-hayır-belki, marka tercihi, siyasi parti tercihi, çok zayıf-zayıf-orta-iyi-çok iyi gibi ikiden fazla değer aldığı modeller de çoklu tercih modelleri olarak adlandırılırlar. Her iki tip modelde de temel amaç tercih olasılığının belirlenmesidir.¹⁴

4.1 İkili Tercih Modelleri

İkili tercih modelleri, karar birimleri olarak bilinen bireyler, firmalar, hane halkları gibi birimlerin tercih yapmak üzere iki alternatifle karşı karşıya olduklarını ve sahip olduğu özelliklerin yapacakları tercihte etkili olduğunu varsaymaktadır. Bu modeller, bir karar biriminin sahip olduğu özellikleri ile belli bir tercihte bulunma olasılığı arasındaki ilişkiyi ortaya koymaktadır. Karar sürecinin olasılıklı yapısı hakkında geliştirilen varsayımlara dayalı olarak çeşitli tercih modelleri oluşturulmuştur.¹⁵ Tercih modelleri arasında en basit olanı *Doğrusal Olasılık Modeli*'dir.

4.1.1 Doğrusal Olasılık Modeli

Bağımlı değişkenin gölge değişken olduğu ve bağımsız değişkenlerin doğrusal bir fonksiyonu olarak gösterildiği modele *Doğrusal Olasılık Modeli* denir. Bu model, bir karar biriminin belli bir tercihi yapma olasılığının onun sahip olduğu özelliklerin katsayılar itibarıyla doğrusal bir fonksiyonu olduğunu varsaymaktadır.

¹⁴ Güriş, S. ve Çağlayan, E., *Ekonometri: Temel Kavramlar*, ss. 652-653

¹⁵ Pindyck, S. R. ve Rubinfeld, D. L., *Econometric Models and Economic Forecasts*, ss. 248-281

Bağımsız değişken X ve iki değer alan Y bağımlı değişkeninin oluşturduğu regresyon modeli şu şekilde olsun:

$$Y_i = \beta_0 + \beta_1 X_i + u_i \quad (4.1)$$

$Y_i = 1$, eğer 1.seçenek olursa (ev sahibi olma, evet oyu verme, araba alma vb.)

$Y_i = 0$, eğer 2.seçenek olursa (ev sahibi olmama, hayır oyu verme, araba almama vb.)

X_i bağımsız değişken (gelir, fiyat vb.)

Bağımlı değişken bir olayın meydana gelmesini veya gelmemesini gösterir. $Y_i = 1$ olma yani olayın gerçekleşme olasılığı P_i , $Y_i = 0$ olma yani olayın gerçekleşmeme olasılığı $1 - P_i$ olarak belirtilirse bağımlı değişkenin olasılık dağılımı aşağıdaki gibi olur.

Y_i	Olasılık
0	$1 - P_i$
1	P_i
Toplam Olasılık	1

Beklenen değer tanımına göre;

$Y = y$	$y_1 \quad y_2 \quad \dots \quad y_n$
$f(x) = P(Y = y)$	$f(y_1) \quad f(y_2) \quad \dots \quad f(y_n)$

$$E(Y) = y_1 \cdot f(y_1) + y_2 \cdot f(y_2) + \dots + y_n \cdot f(y_n) \quad (4.2)$$

olarak tanımlanmaktadır.

Bu tanımdan hareketle modelimizdeki bağımlı değişkenin beklenen değerini hesaplayacak olursak;

$$E(Y_i) = 0 \cdot (1 - P_i) + 1 \cdot P_i = P_i \quad (4.3)$$

şeklinde bir sonuca ulaşırız.

Diğer yandan $E(u_i) = 0$ varsayımı altında;

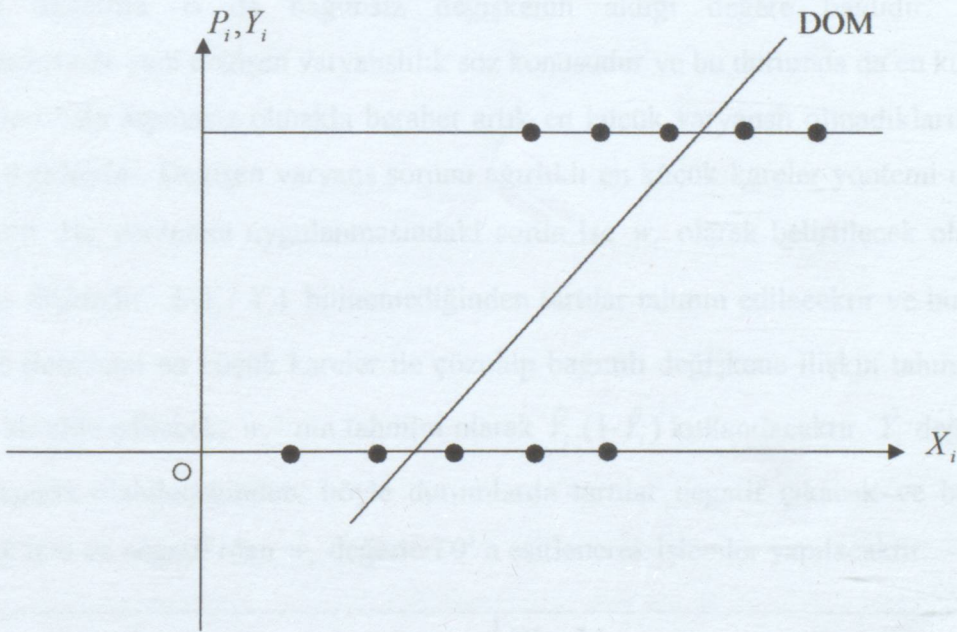
$$E(Y_i / X_i) = \beta_0 + \beta_1 X_i \quad (4.4)$$

olarak da yazılabilir.

X_i veri iken bağımlı değişkenin koşullu beklenen değeri olarak gösterilen $E(Y_i / X_i)$, X_i veri iken olayın gerçekleşmesinin olasılığı, $P(Y_i = 1 / X_i)$, olarak yorumlanabilmektedir.

$$E(Y_i / X_i) = \beta_0 + \beta_1 X_i = P_i \quad (4.5)$$

DOM, β_0 ve β_1 ' in tahmin değerlerine bağlı olarak bağımsız değişkenin çeşitli değerleri karşısında ilgilenilen olayın gerçekleşme olasılığını verirken, modeldeki eğim katsayısı olan β_1 , bağımsız değişkendeki bir birimlik değişimin olayın gerçekleşme olasılığı üzerindeki etkisini ifade etmektedir.



Şekil 4.1 Doğrusal olasılık modeli (Thomas, s.471)

Doğrusal olasılık modelinde ortaya çıkan sorunlar maddeler halinde açıklanacaktır.

1- u_i 'nin normal dağılmaması : Güven aralıkları ve hipotez testleri gibi istatistiksel çıkarımlar için hata terimlerinin normal dağıldığı varsayımı yapılmaktadır. Doğrusal olasılık modelinde bağımlı değişken gibi u_i 'ler de yalnızca 2 değer almaktadır. u_i 'nin normallik varsayımını bozmasına rağmen, regresyon katsayılarının tahmincileri yansız olmakta, fakat heteroskedastisiteden dolayı etkinlik özelliği bulunmamaktadır.

Ayrıca merkezî limit teoremine dayanarak, büyük örneklerde en küçük kareler tahmincileri normal dağılıma uyma eğilimi göstermektedir.

$$u_i = Y_i - \beta_0 - \beta_1 X_i \quad (4.6)$$

olduğu düşünülürse;

$$\left. \begin{array}{l} Y_i = 1 \Rightarrow u_i = 1 - \beta_0 - \beta_1 X_i \\ Y_i = 0 \Rightarrow u_i = -\beta_0 - \beta_1 X_i \end{array} \right\} \text{binom dağılımına uygunluk görülmektedir.} \quad (4.7)$$

2- Değişen varyans : Klasik doğrusal regresyon modelinin varsayımlarından biri olan hata terimlerinin sabit varyanslı olması, yani hata teriminin varyansının bağımsız değişkenin alacağı değerlere bağlı olarak değişmemesi varsayımı doğrusal olasılık modelinde gerçekleşmemektedir. Doğrusal olasılık modelinde varyans bağımlı değişkenin koşullu beklenen değerine, o da bağımsız değişkenin aldığı değere bağlıdır. Dolayısıyla heteroskedastisite yani değişen varyanslılık söz konusudur ve bu durumda da en küçük kareler tahmincileri hala sapmasız olmakla beraber artık en küçük varyanslı olmadıklarından dolayı etkin de değildirler. Değişen varyans sorunu ağırlıklı en küçük kareler yöntemi uygulanarak giderilebilir. Bu yöntemin uygulanmasındaki sorun ise w_i olarak belirtilecek olan tartıların tahminine ilişkindir. $E(Y_i / X_i)$ bilinemediğinden tartılar tahmin edilecektir ve bunun için de regresyon denklemi en küçük kareler ile çözülüp bağımlı değişkene ilişkin tahmin değerleri olan $\hat{Y}_i$ ' ler elde edilecek, w_i ' nin tahmini olarak $\hat{Y}_i \cdot (1 - \hat{Y}_i)$ kullanılacaktır. $\hat{Y}_i$ değerleri 0 – 1 aralığı dışında olabileceğinden, böyle durumlarda tartılar negatif çıkacak ve bu sorundan kurtulmak için de negatif olan w_i değerleri 0' a eşitlenerek işlemler yapılacaktır.

u_i	Olasılık
$-\beta_0 - \beta_1 X_i$	$1 - P_i$
$1 - \beta_0 - \beta_1 X_i$	P_i
Toplam Olasılık	1

$$E(u_i) = (1 - \beta_0 - \beta_1 X_i) \cdot P_i + (-\beta_0 - \beta_1 X_i) \cdot (1 - P_i) = 0 \quad (4.8)$$

$$\begin{aligned} &= P_i - P_i \beta_0 - P_i \beta_1 X_i - \beta_0 - \beta_1 X_i + P_i \beta_0 + P_i \beta_1 X_i \\ &= (\beta_0 + \beta_1 X_i) - \beta_0 - \beta_1 X_i \\ &= 0 \end{aligned}$$

$$Var(u_i) = \sigma_i^2 = E[u_i - E(u_i)]^2 = E(u_i)^2 \quad (4.9)$$

$$\begin{aligned} &= (1 - \beta_0 - \beta_1 X_i)^2 \cdot P_i + (-\beta_0 - \beta_1 X_i)^2 (1 - P_i) \\ &= (1 - \beta_0 - \beta_1 X_i)^2 (\beta_0 + \beta_1 X_i) + (-\beta_0 - \beta_1 X_i)^2 (1 - \beta_0 - \beta_1 X_i) \\ &= (1 - \beta_0 - \beta_1 X_i)^2 (\beta_0 + \beta_1 X_i) + [-(\beta_0 + \beta_1 X_i)]^2 (1 - \beta_0 - \beta_1 X_i) \\ &= (1 - \beta_0 - \beta_1 X_i)^2 (\beta_0 + \beta_1 X_i) + (\beta_0 + \beta_1 X_i)^2 (1 - \beta_0 - \beta_1 X_i) \\ &= (1 - \beta_0 - \beta_1 X_i)(\beta_0 + \beta_1 X_i)[(1 - \beta_0 - \beta_1 X_i) + (\beta_0 + \beta_1 X_i)] \\ &= (1 - \beta_0 - \beta_1 X_i)(\beta_0 + \beta_1 X_i) \\ &= (1 - P_i)P_i \\ &= [1 - E(Y_i / X_i)]E(Y_i / X_i) \end{aligned}$$

P_i 'nin 0 ve 1'e yaklaştığı durumlarda düşük varyans, $\frac{1}{2}$ 'ye yaklaştığı durumlarda ise daha yüksek varyanslar söz konusu olur.

3- $0 \leq E(Y_i / X_i) \leq 1$ varsayımının yerine gelmeyişi : Bu modelde $E(Y_i / X_i)$, X_i veri iken bağımlı değişkene ilişkin gerçekleşme olasılığını ölçtüğüne göre, bu değer 0 – 1 arasında olmalıdır. Önsel olarak doğru olan bu durum, uygulama aşamasında zaman zaman gerçekleşmeyebilmektedir. En küçük kareler yöntemi ile regresyon çözümlemesi yapıldığında, $E(Y_i / X_i)$ 'in tahmin edicileri olan $\hat{Y}_i$ 'lerin 0-1 arasında olup olmadığına bakılır ve eğer bu sınırları aşan değerler mevcutsa 0'dan küçük olanlara 0, 1'den büyük olanlara da 1 değeri verilir.¹⁶ Bu durumda tahmin süreci sapmasız tahminciler verebilmesine karşılık, modelin öngörü (prediction) amacıyla kullanılması halinde elde edilen öngörüler sapmalı olacaktır. Aşağıdaki ifadeden hareket edilebilir.

$$P_i = \begin{cases} \beta_0 + \beta_1 X_i & 0 < \beta_0 + \beta_1 X_i < 1 \\ 1 & \beta_0 + \beta_1 X_i \geq 1 \\ 0 & \beta_0 + \beta_1 X_i \leq 0 \end{cases}$$

¹⁶ Gujarati, D. N., Temel Ekonometri, s.544

4- Doğrusal olasılık modelinde $P(Y_i = 1 / X_i)$ değeri, X_i 'nin doğrusal bir fonksiyonu şeklinde ifade edilmekte ve dolayısıyla X_i 'nin marjinal etkisinin sabit kaldığı görülmektedir.

P_i 'nin bağımsız değişkene göre kısmi türevi alındığında, bağımsız değişkendeki değişmelerin olasılık üzerinde sabit bir etkiye sahip olduğunu görmekteyiz. Buna göre, bağımsız değişkendeki bir birimlik değişim olayın gerçekleşme olasılığını her defasında β_i kadar değiştirmektedir. Modelin zayıf noktalarından biri de bu şekilde ortaya konabilir.

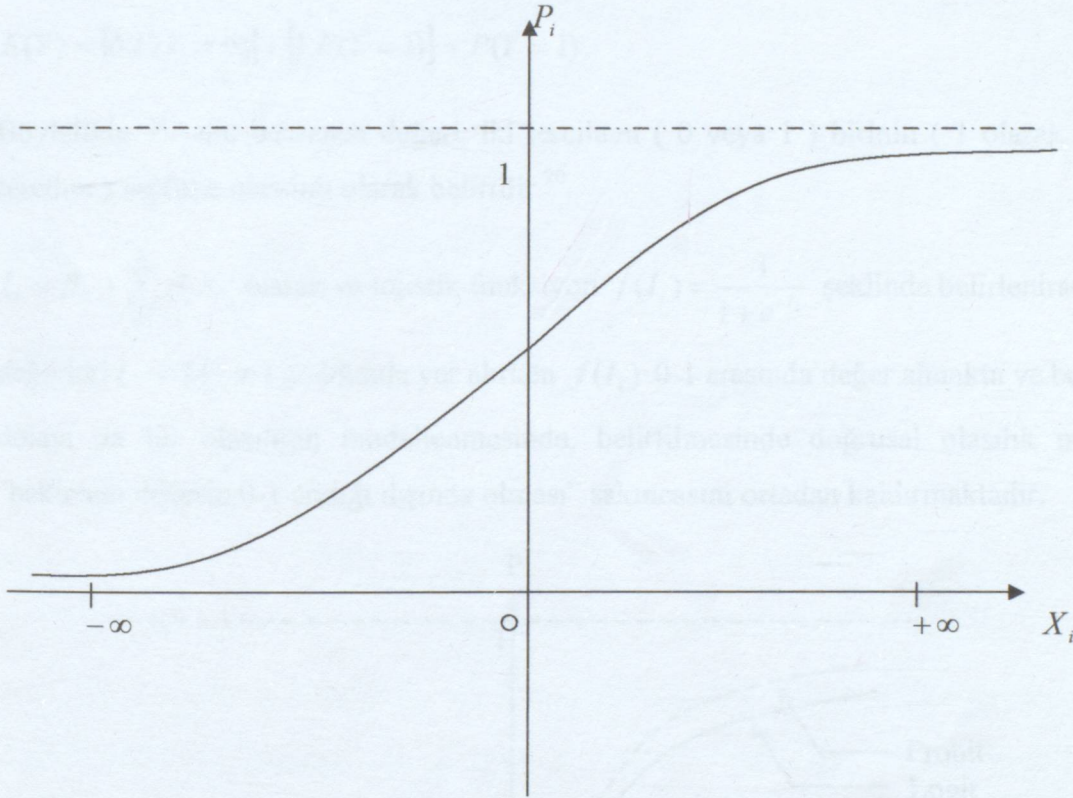
$$\frac{\partial P_i}{\partial X_i} = \beta_i \quad (4.10)$$

5- Tahmin edilen regresyon modelinin gerçek gözlem değerlerine uyumunun iyiliğinin ölçüsü olarak hesaplanan ve kullanılan belirlilik katsayısı gölge bağımlı değişkenli modellerde uyum iyiliğinin iyi bir ölçüsü olarak kullanılamamaktadır. Doğrusal olasılık modelinde belirlilik katsayısı değeri çoğu uygulamada 0,2 ile 0,6 arasında yer almaktadır. Bu nedenle Aldrich ve Nelson, gölge bağımlı değişkenli modellerde belirlilik katsayısı değerinin bir özetleme istatistiği olarak kullanılmasından kaçınılması gerektiğini ileri sürmüşlerdir.¹⁷

¹⁷ Aldrich, J. ve Nelson, F., Linear Probability, Logit and Probit Models, s. 15

4.1.2 Lojistik Regresyon ve Logit Model

Bir önceki bölümde doğrusal olasılık modeli ve bu modelin kullanımıyla ortaya çıkan birtakım sorunlardan bahsedildi. Özellikle 3. ve 4. maddelerde açıklanan sorunlardan kurtulmak için birikimli dağılım fonksiyonları yardımıyla kurulan modeller geliştirilmiştir. Şekilde görülen S biçimindeki eğri, tesadüfi bir değişkenin birikimli dağılım fonksiyonuna benzemektedir.¹⁸ Lojistik ve normal birikimli dağılım fonksiyonlarının kullanıldığı modeller logit ve probit modeller olarak bilinmektedir ve kullanılmaktadır.



Şekil 4.2 Birikimli dağılım fonksiyonu

4.1.2.1 Lojistik Regresyonda Genel İfadeler

Lojistik regresyon analizi bağımlı değişkenin gölge değişken olarak tanımlandığı durumlarda sıklıkla kullanılan ve açıklayıcı değişkenlere göre bağımlı değişkenin beklenen değerinin

¹⁸ Bir tesadüfi değişken X ' in BDF' si, bu değişkenin X_0 ' a eşit ya da ondan küçük bir değer alması olasılığıdır. X_0 , X ' in aldığı belli bir sayısal değeri ifade etmektedir. X ' in BDF' si $F(X = X_0) = P(X \leq X_0)$ olarak gösterilir. Tüm olasılıklar 0 - 1 arasında yer aldığından BDF' de 0 - 1 aralığında yer alacaktır.

olasılık olarak elde edildiği bir regresyon yöntemidir.¹⁹ Birikimli lojistik olasılık fonksiyonu üzerine kuruludur ve genel gösterimi aşağıdaki gibidir.

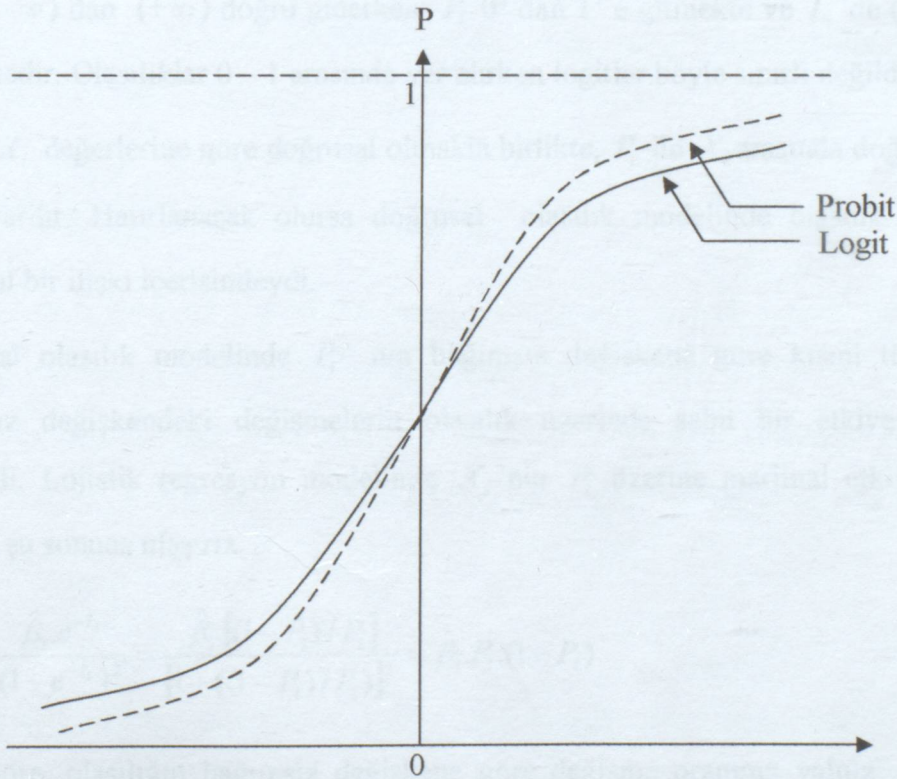
$$E(Y) = \frac{1}{1 + \exp\left[-(\beta_0 + \sum_{i=1}^k \beta_i X_i)\right]} \quad (4.11)$$

Bağımlı değişkenin 0, 1 değerlerini alan bir gölge değişken olduğu göz önüne alınırsa beklenen değerden hareketle şu sonuca varılır.

$$E(Y) = [0.P(Y = 0)] + [1.P(Y = 1)] = P(Y = 1) \quad (4.12)$$

Böylelikle Y 'nin beklenen değeri, iki tercihten (0 veya 1) birinin (1 olarak tanımlanan tercihin) seçilme olasılığı olarak belirtilir.²⁰

$I_i = \beta_0 + \sum_{i=1}^k \beta_i X_i$ olarak ve lojistik fonksiyon $f(I_i) = \frac{1}{1 + e^{-I_i}}$ şeklinde belirlenirse I_i 'nin değerleri $(-\infty) (+\infty)$ aralığında yer alırken $f(I_i)$ 0-1 arasında değer almakta ve bu özellikten dolayı da bir olasılığın modellenmesinde, belirtilmesinde doğrusal olasılık modelindeki “beklenen değer 0-1 aralığı dışında olması” sakıncasını ortadan kaldırmaktadır.



Şekil 4.3 Logit ve probit model birikimli dağılım eğrileri

¹⁹ Özdamar, K., Paket Programlar ile İstatistiksel Veri Analizi 1, s.623

²⁰ Kleinbaum vd., Applied Regression Analysis and Other Multivariable Methods, s.657

$$P_i = E(Y = 1 / X_i) = \frac{1}{1 + e^{-I_i}} \Rightarrow Q_i = 1 - P_i = \frac{1}{1 + e^{I_i}} \quad (4.13)$$

P_i/Q_i oranı bahis oranı (odds ratio) olarak adlandırılmakta ve gerekli işlemler yapılarak aşağıdaki sonuca ulaşılmaktadır.

$$\frac{P_i}{Q_i} = \frac{1}{1 + e^{-I_i}} \cdot \frac{1 + e^{I_i}}{1} = \frac{e^{I_i}}{1 + e^{I_i}} \cdot \frac{1 + e^{I_i}}{1} = e^{I_i} \quad (4.14)$$

P_i ile I_i arasındaki ilişkinin doğrusal olmadığını, diğer anlamda P_i ile hem bağımsız değişken X_i 'ler arasındaki hem de regresyon katsayıları olan β_i 'ler arasındaki ilişkinin doğrusal olmadığını görmekteyiz. Fakat uygulanacak bazı işlemlerle doğrusal biçime dönüştürülebilmektedir. Bahis oranının logaritması alınırsa *logit* olarak adlandırılan ifadeye ulaşılır. Logit hem bağımsız değişkene hem de regresyon katsayılarına göre doğrusaldır.

$$\ln \frac{P_i}{Q_i} = I_i = \beta_0 + \sum_{i=1}^k \beta_i X_i = L_i \quad (4.15)$$

P_i ve L_i aşağıda yer alan özelliklere sahiptir.

1- I_i ($-\infty$) dan $(+\infty)$ doğru giderken, P_i 0' dan 1' e gitmekte ve L_i 'de ($-\infty$) dan $(+\infty)$ a gitmektedir. Olasılıklar 0 – 1 arasında yer alırken logitler böyle sınırlı değildir.

2- L_i , X_i değerlerine göre doğrusal olmakla birlikte, P_i ile X_i arasında doğrusal olmayan bir ilişki vardır. Hatırlanacak olursa doğrusal olasılık modelinde olasılık değerleri X_i ile doğrusal bir ilişki içerisindeydi.

Doğrusal olasılık modelinde P_i ' nin bağımsız değişkene göre kısmi türevi alındığında, bağımsız değişkendeki değişmelerin olasılık üzerinde sabit bir etkiye sahip olduğuna değinildi. Lojistik regresyon modelinde X_i 'nin P_i üzerine marjinal etkisini hesaplayacak olursak şu sonuca ulaşırız.

$$\frac{\partial P_i}{\partial X_i} = \frac{\hat{\beta}_i \cdot e^{-I_i}}{(1 + e^{-I_i})^2} = \frac{\hat{\beta}_i \cdot [(1 - P_i) / P_i]}{[1 + ((1 - P_i) / P_i)]^2} = \hat{\beta}_i \cdot P_i \cdot (1 - P_i) \quad (4.16)$$

Buna göre, olasılığın bağımsız değişkene göre değişme oranının yalnız β_i ' e değil, aynı zamanda değişimin ölçüldüğü olasılığın düzeyine de bağlı olduğu sonucuna varılır. Birikimli lojistik dağılım fonksiyonu incelendiğinde, $P_i = 1/2$ olduğunda eğimin en büyük değerli

olduğu görülür. Bu da, bağımsız değişkenlerdeki değişimin, herhangi bir tercihin seçilme olasılığı üzerine en büyük etkisinin, dağılımın orta noktasında meydana geldiğini belirtmektedir.

Bu arada Amemiya doğrusal olasılık modeli ile logit modellerinin katsayılarının ilişkilendirilebileceğini aşağıdaki bağıntılarla göstermiştir.²¹

$$\beta_{DOM} \cong 0,25\beta_{Logit}; \text{ sabit terim bunun dışındadır.} \quad (4.17)$$

$$\beta_{DOM} \cong 0,25\beta_{Logit} + 0,5; \text{ sabit terim için} \quad (4.18)$$

4.1.2.2 Bahis Oranı (Odds Ratio) Hesaplanması

Lojistik regresyonda, regresyon katsayıları olan β_i 'lerin yorumlanmasında, bahis oranları olarak bilinen odds ratio değerleri sıklıkla kullanılmaktadır. X olayını kansere yakalanma olarak belirtecek olursak bu olaya ilişkin $odds(X)$ değeri, X olayının gerçekleşme olasılığının gerçekleşmeme olasılığına bölünmesiyle hesaplanmaktadır.²²

$$odds(X) = \frac{P(X)}{1 - P(X)} \quad (4.19)$$

OR simgesiyle gösterilecek olan bahis oranı (odds ratio) ise iki odds değerinin birbirine oranlanmasıyla bulunmaktadır.

$$OR_{A,B} = \frac{odds(X_A)}{odds(X_B)} = \frac{P(X_A)}{1 - P(X_A)} \bigg/ \frac{P(X_B)}{1 - P(X_B)} \quad (4.20)$$

A ve B karşılaştırılacak 2 grubu belirtmektedir. Örnekle konuya açıklık getirmeye çalışalım. A grubu sigara içenler, B grubu ise sigara içmeyenler olarak belirtilirse X_A sigara içenlerin kansere yakalanması, X_B ise sigara içmeyenlerin kansere yakalanması olarak ifade edilir. $P(X_A) = 0,25$ ve $P(X_B) = 0,10$ olarak ele alalım ve OR değerini hesaplayalım.

$$OR_{A,B} = \frac{0,25}{1 - 0,25} \bigg/ \frac{0,10}{1 - 0,10} = 3$$

Sigara içen bireylerin sigara içmeyenlere göre, kansere yakalanma olasılığı 3 kat daha fazla olmaktadır. Eğer bu oran 1'e eşit olmuş olsaydı, kansere yakalanmada sigara içmenin herhangi bir etkisinin olmadığı söylenirdi.

²¹ Gujarati, D. N., a.g.e., s.569

²² Kleinbaum vd., a.g.e, s.658

Lojistik modelin, logit şeklinde olan ifadesi daha yaygın olarak kullanılmaktadır. Bağımlı değişkeni, kişinin kanser olup olmamasına göre 2 değer alan bir gölge değişken ve bağımsız değişkeni de sigara alışkanlığı olarak ele alıp modelin logit formunu aşağıdaki şekilde yazabiliriz.

$$\text{Logit}[P(Y = 1)] = \log_e[\text{odds}(Y = 1)] = \log_e \left[\frac{P(Y = 1)}{1 - P(Y = 1)} \right] = I_i \quad (4.21)$$

$$\text{Logit}[P(Y = 1)] = \beta_0 + \beta_1 D_1 \quad (4.22)$$

$Y = 1$, kanser olma (kansere yakalanma)

$D_1 = 1$, sigara içme

$Y = 0$, kanser olmama (kansere yakalanmama)

$D_1 = 0$, sigara içmeme

Sigara içen grup A, sigara içmeyen grup ise B grubu olarak tanımlanır ve bu 2 gruba ilişkin OR değeri hesaplanır.

$$\log_e(\text{odds}(D_A)) = \beta_0 + \beta_1 \cdot 1 = \beta_0 + \beta_1 \quad (4.23)$$

$$\log_e(\text{odds}(D_B)) = \beta_0 + \beta_1 \cdot 0 = \beta_0 \quad (4.24)$$

$$OR_{A,B} = \frac{\text{odds}(D_A)}{\text{odds}(D_B)} = \frac{e^{\beta_0 + \beta_1}}{e^{\beta_0}} = e^{\beta_1} \quad (4.25)$$

Sigara içen ve içmeyen olmak üzere 2 grubun karşılaştırılmasında kullanılacak olan OR değeri böylece bulunur.

Bağımsız değişkenler $X_1, X_2, X_3, \dots, X_k$ olarak ifade edilir ve iki grup

$X_A = (X_{A1}, X_{A2}, X_{A3}, \dots, X_{Ak})$; $X_B = (X_{B1}, X_{B2}, X_{B3}, \dots, X_{Bk})$ şeklinde belirlenir.

Bağımsız değişken sayısını 3 ($k=3$) olarak ele alıp bu değişkenleri aşağıdaki gibi tanımlayalım.²³

X_1 = sigara alışkanlığı (1=sigara içen ; 0=sigara içmeyen)

X_2 = yaş

X_3 = renk (1=siyah ; 0=beyaz)

²³ Kleinbaum vd., a.g.e, s.659

Modelin logit formundaki genel biçimi şu şekilde olmaktadır.

$$\text{Logit}[P(Y = 1)] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \quad (4.26)$$

45 yaşında, rengi siyah olan, sigara içenlerin grubu $X_A = (1,45,1)$; 45 yaşında, siyah renkli, sigara içmeyenlerin grubu ise $X_B = (0,45,1)$ olarak tanımlanırsa bu 2 gruba ilişkin OR değeri hesaplanır.

$$OR_{X_A, X_B} = \frac{\text{odds}(X_A)}{\text{odds}(X_B)} = \frac{e^{\beta_0 + \sum \beta_j X_{Aj}}}{e^{\beta_0 + \sum \beta_j X_{Bj}}} = e^{\sum (X_{Aj} - X_{Bj}) \beta_j} \quad (4.27)$$

$$OR_{X_A, X_B} = e^{(1-0)\beta_1 + (45-45)\beta_2 + (1-1)\beta_3} = e^{\beta_1}$$

Bu oranın hesaplanmasında β_0 ' ın kullanılmadığı görülmektedir. Bu ifade anakütle bahis oranı olarak bilinir, çünkü regresyon katsayıları bilinmemektedir, bunların bir tahmini olarak en çok benzerlik yöntemi uygulanarak $\hat{\beta}_j$ 'ler hesaplanır ve bahis oranı böylelikle bulunmuş olur. Örneğimizde sigara içme durumu A grubunda 1 iken, B grubunda 0 olmakta ve sürekli değişken olarak tanımlanan yaş ile gölge değişken olarak tanımlanan renk değişkenleri sabit kalmaktadır. Genel olarak, sadece bir değişkenin değişiklik gösterdiği ve diğer değişken (lerin) sabit olarak kaldığı, değişiklik gösteren değişkene göre iki grubun bahis oranlarının karşılaştırıldığı bu oran *ayarlı bahis oranı* olarak adlandırılmaktadır.²⁴ Sigara içme *odak* değişken (exposure, study, treatment), diğer sabit kalanlar ise *kontrol* (confounder) değişkeni olarak belirlenir.

$\text{Exp}(\beta_p)$ değeri 1' e yakın değişkenler bağımlı değişkenin değişimine önemli katkıda bulunmazlar. Katsayı önemli olmak koşulu ile 1' den büyük $\text{Exp}(\beta_p)$ değerine sahip değişkenler ise önemli bir etkiye sahiptir. 0' a yakın $\text{Exp}(\beta_p)$ değerine sahip değişkenler ise bağımlı değişkenin düşük değerler almasına sebep olan negatif etkili değişkenlerdir.²⁵ (SPSS çıktılarında OR değerleri $\text{Exp}(\beta_p)$ olarak ayrı bir sütunda gösterilmektedir.)

4.1.2.3 Logit Modelinin Tahmin Edilmesi

Parametre tahminlerinde sıklıkla kullanılan yöntemler en küçük kareler ve en çok benzerlik yöntemleridir. En küçük kareler yönteminde bağımlı değişkenin gerçek değerleri ile tahmin değerleri arasındaki farkların karelerinin toplamını minimum yapan parametre değerleri elde

²⁴ Kleinbaum vd., a.g.e, s.660

²⁵ Özdamar, K., a.g.e, s.633

edilmektedir. En çok benzerlik yönteminde ise ana kütle ve bu ana kütlede çekilen örnek arasındaki benzerlik ilişkisinden yararlanılarak bu örneğin elde edilme olasılığını maksimum yapan parametre değerleri tahmin edilmektedir.²⁶ Benzerlik fonksiyonunun genel ifadesi şöyledir;

$$LL(X_1, X_2, \dots, X_n; Q_1, Q_2, \dots, Q_k)$$

$Q_1, Q_2, \dots, Q_k$, tahmin edilmek istenen fonksiyonun parametrelerini belirtir. X' in normal dağılıma sahip olması halinde benzerlik fonksiyonunun biçimi şöyle olur;

$$LL(X_1, X_2, \dots, X_n; \mu, \sigma^2)$$

Maksimum benzerlik yöntemi, benzerlik fonksiyonunun maksimizasyonundan oluşmaktadır ve bu yöntemin uygulanabilmesi için hata teriminin dağılımının bilinmesi gerekmektedir. Ayrıca, en çok benzerlik yöntemi kullanıldığında doğrusal olmayan parametreler tahmin edileceğinden, bu tahminler ancak iterasyonlarla yapılabilmektedir.²⁷

Bu yöntemle elde edilen en çok benzerlik tahmincileri küçük örneklerde sapmasız, etkin olmayan ve normal dağılmayan, büyük örneklerde ise tutarlı, asimptotik etkin ve asimptotik normal tahmincilerdir.²⁸

$$\ln \frac{P_i}{Q_i} = L_i = \beta_0 + \sum_{i=1}^k \beta_i X_i = L_i$$

Logit modelin tahmin edilmesi için L_i değerlerinin bilinmesi gerekmektedir. Mikro ya da kişisel, bireysel düzeyde veriler olduğunda, diğer bir ifadeyle gruplandırılmamış veriler söz olduğunda ve bu verilerle logit model tahmininde bulunmak istenildiğinde bu noktadaki sorunun görülebilmesi için, P_i değerinin 1 ya da 0 olduğunu hatırlamak ve bu değerlerin L_i 'de yerine konulması gerekmektedir.

$$P_i = 1 \Rightarrow L_i = \ln \left(\frac{1}{0} \right) \quad P_i = 0 \Rightarrow L_i = \ln \left(\frac{0}{1} \right)$$

Bulunan bu ifadelerin anlamsız olmasından dolayı bu durumda en küçük kareler yöntemi kullanılamamakta ve regresyon katsayılarının tahmin edilmesinde en çok benzerlik yöntemi kullanılmaktadır.

²⁶ Akın, F., Ekonometri, s.113

²⁷ Gürış, S. ve Çağlayan, E., a.g.e, s.661

²⁸ Gürış, S. ve Çağlayan, E., a.g.e, s.661

Modelimize ilişkin benzerlik fonksiyonu şöyle ifade edilir;

$$LL(\beta / X_i) = \prod_{i=1}^n P(Y_i / X_i; \beta) = \prod P(Y_i = 0 / X_i; \beta)^{1-Y_i} \cdot P(Y_i = 1 / X_i; \beta)^{Y_i} \quad (4.28)$$

Bu ifadenin logaritması alınacak olursa;

$$\ln LL(\beta / X_i) = \sum_{i=1}^n [(1 - Y_i) \cdot \ln P(Y_i = 0 / X_i; \beta) + Y_i \cdot \ln P(Y_i = 1 / X_i; \beta)] \quad (4.29)$$

$$\ln LL(\beta / X_i) = \sum_{i=1}^n (1 - Y_i) \ln(1 - P_i) + Y_i \ln P_i$$

$$\ln LL = \sum_{i=1}^n Y_i [\ln P_i - \ln(1 - P_i)] + \ln(1 - P_i)$$

$$= \sum_{i=1}^n Y_i \ln [P_i / (1 - P_i)] + \ln(1 - P_i)$$

$$= \sum_{i=1}^n Y_i \ln [P_i / Q_i] + \sum_{i=1}^n \ln \left[1 - \frac{1}{1 + e^{-I_i}} \right]$$

$$= \sum_{i=1}^n Y_i \cdot I_i + \sum_{i=1}^n \ln \left[1 - \frac{e^{I_i}}{1 + e^{I_i}} \right]$$

$$= \sum_{i=1}^n Y_i I_i + \sum_{i=1}^n \ln \left[\frac{1}{1 + e^{I_i}} \right]$$

$$= \sum_{i=1}^n Y_i I_i + \sum_{i=1}^n \ln(1 + e^{I_i})^{-1} \quad \text{sonucuna ulaşılır.}$$

Bu fonksiyonun 0' a eşitlenen kısmî türevleri çözülerek bilinmeyen regresyon katsayılarının en yüksek olabilirlik tahmin edicileri bulunabilir.

$$\frac{\partial \ln LL}{\partial \beta} = \sum_{i=1}^n Y_i X_i - \sum_{i=1}^n \frac{e^{I_i}}{1 + e^{I_i}} X_i = \sum_{i=1}^n (Y_i - P_i) X_i = 0 \quad (4.30)$$

Gruplanmış veriler olduğunda logit modelinin tahmin edilmesinde en küçük kareler yöntemi kullanılabilmektedir. Belirli bir gelir düzeyindeki birey sayısını N_i , bu gelir düzeyindeki istenilen özelliğe (1.tercihi seçenlerin sayısı) sahip kişilerin sayısını da n_i olarak simgeleyelim. Örneğin 1000 YTL gelire sahip 70 birey olduğunu ($N_i = 70$) ve bu bireylerden

de 17 tanesinin ($n_i = 17$) ev sahibi olduğunu düşünelim. Gruplanmış verilerde P_i 'nin tahmini olarak n_i / N_i değeri kullanılmaktadır. Bu durumda 1000 YTL gelir düzeyi için tahmin edilen değer 0,24 olarak hesaplanmaktadır.

$$\hat{P}_i = \frac{n_i}{N_i} \quad (4.31)$$

Tahmin edilmiş P_i değerlerini kullanarak logitler de şöyle bulunmaktadır.

$$\hat{L}_i = \ln\left(\frac{\hat{P}_i}{1 - \hat{P}_i}\right) = \ln\left(\frac{n_i / N_i}{1 - (n_i / N_i)}\right) = \ln\left(\frac{n_i}{N_i - n_i}\right) \quad (4.32)$$

$\hat{P}_i$, ortalaması P_i 'ye, varyansı da $P_i(1 - P_i)/N_i$ 'ye eşit binom dağılımına uymakta ve N_i sonsuza doğru büyüdükçe binom dağılımı da normal dağılıma yakınsamaktadır. Logit modelde hata terimleri 0 ortalamalı ve $1/N_i P_i(1 - P_i)$ varyanslı normal dağılıma uymaktadır. Bu nedenle, en küçük kareler yöntemi yerine, bu sorunu çözen ve genelleştirilmiş en küçük kareler yönteminin bir ifadesi olan, ağırlıklı en küçük kareler yöntemi kullanılmaktadır.

4.1.2.4 Uyum İyiliği Ölçüleri

Uyum iyiliği ölçüleri, gözlenen veriler ile tahmin edilen modelin doğruluğunu, tamlığını ölçmede kullanılan ve yararlanılan istatistiklerdir. Belirli verilere dayalı olarak tahmin edilen bir örneklem regresyon doğrusu için uyumun iyiliği, söz konusu regresyon doğrusunun gerçek gözlem değerlerine ne kadar iyi uyduğunu, diğer bir ifadeyle onları ne kadar iyi temsil ettiğini ifade etmektedir. Doğrusal regresyon modellerinde kullanılan belirlilik katsayısı bir örnek olarak verilebilir.²⁹ Bağımlı değişkenin nitel olduğu tercih modellerinde belirlilik katsayısının değeri oldukça düşük olduğundan bu modellerde uyum iyiliğinin ölçütü olarak kabul edilmemektedir. Bağımlı değişkeni gölge değişken olan regresyon modellerinde uyum iyiliği ölçütü olarak aşağıdaki istatistikler kullanılmaktadır.

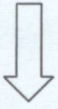
$$R_{PSD}^2 = 1 - \frac{1}{1 + 2(\log LL_1 - \log LL_0) / N} \quad (4.33)$$

$$R_{MF}^2 = 1 - \frac{\log LL_1}{\log LL_0} \quad (4.34)$$

²⁹ Verbeek, M., A Guide to Modern Econometrics, s.182

LL_1 benzerlik fonksiyonunun maksimum değerini, LL_0 sabit terim dışındaki tüm eğim katsayılarının sıfıra eşit olduğu kısıtı dikkate alındığında elde edilen benzerlik fonksiyonunun maksimum değerini ve N ise gözlem sayısını göstermektedir. $McFaddenR^2$ değeri benzerlik oran endeksi (likelihood ratio index) olarak da bilinmektedir.³⁰ Tahmin edilen parametre değerlerinin tümünün 0 değerine eşit olması durumunda $\log LL_0 = \log LL_1$ olacak ve böylece uyum iyiliğinde kullanacağımız her iki değer de 0' a eşit olacaktır. LL_1 ve LL_0 değerleri sırasıyla (4.36) ve (4.35) için olabilirlik fonksiyonlarının maksimum değerlerini belirtmektedir.³¹

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$



$$\ln\left(\frac{P_i}{1-P_i}\right) = \beta_0 + \varepsilon_i \quad (4.35)$$

$$\ln\left(\frac{P_i}{1-P_i}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon_i \quad (4.36)$$

Modele yeni değişkenler eklendiğinde $McFaddenR^2$ değeri artmakta ya da en azından sabit kalmaktadır. Bu yüzden dolayı, daha iyi bir karşılaştırma imkanı sunan düzeltilmiş $McFaddenR^2$ değeri kullanılmaktadır. Horowitz (1982,1983a) tarafından bulunan ve kullanılan formülde ise K yerine $K/2$ değeri yer almaktadır.³²

$$\overline{R_{MF}^2} = 1 - \frac{LL_1 - K}{LL_0} \quad (4.37)$$

Katsayıların ayrı ayrı anlamlılıklarının testinde t testi kullanılmaktadır. Hipotezler aşağıdaki gibi kurulur ve hesaplanan t değerinin tablo değeriyle karşılaştırılması sonucunda bir karara varılır.

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0$$

³⁰ Ramanathan, R., Introductory Econometrics with Applications, s.284

³¹ Thomas, R. L., Modern Econometrics-An Introduction, s.475

³² Lerman, R. S. ve Akiva, B. M., Discrete Choice Analysis-Theory and Application to Travel Demand, s.168

Katsayıların birlikte anlamlılıklarının testinde ise olabilirlik oranı testi kullanılmaktadır.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

H_1 : en az biri sıfırdan farklıdır.

Olabilirlik oranı test istatistiğinin hesaplanmasında kısıtlı ve kısıtsız modeller en çok benzerlik oranı yöntemiyle tahmin edilerek maksimum olabilirlik değerleri bulunur. Bu test istatistiği m (kısıt sayısı) serbestlik derecesi ile χ^2 dağılımına uyar. Belirlenen önem düzeyinde olabilirlik oranı test istatistiği tablo χ^2 değerinden büyük olduğunda sıfır hipotezinin kabulü için yeterli bir neden bulunamamaktadır şeklinde bir karara varılır.

$$OO = -2(\text{Log}LL_0 - \text{Log}LL_1) \quad (4.38)$$

Nitel tercih modellerinde kullanılan uyum iyiliği ölçütleri 3 kategori bazında verilecektir.³³

a) Doğru ve yanlış tahminlerin sayısına dayalı olan ölçütler;

- Yanlış tahminlerin sayısı

$$\hat{Y}_i = 1 \quad \text{eğer } \hat{P}_i \geq 0,5$$

$$\hat{Y}_i = 0 \quad \text{eğer } \hat{P}_i < 0,5$$

$$NOWP = \sum_{i=1}^N (Y_i - \hat{Y}_i)^2 \quad (4.39)$$

(4.39)' da elde edilen değer, diskriminant analizinde daha çok kullanılmaktadır. İkili tercih modellerinde ise doğru tahminlerin sayısının toplam gözlem sayısına bölünmesiyle elde edilen Count R^2 değeri daha sık kullanım alanına sahiptir.³⁴

- Doğru tahminlerin oranı

Formülde yer alan n_c , doğru tahminlerin sayısını, n_1 , değeri 1 olan gözlemlerin sayısını, n_0 , değeri 0 olan gözlemlerin sayısını göstermektedir.

$$PCP = \frac{n_c - n_m}{N - n_m}, \quad n_m = \max(n_1, n_0) \quad (4.40)$$

³³ Klein, I. ve Mitnik, S., Contributions to Modern Econometrics - From Data Analysis to Economic Policy, ss. 40-41-42

³⁴ Akın, F., Kalitatif Tercih Modelleri Analizi, s.58

b) Hata kareleri toplamına dayalı olan ölçütler;

- Hata kareleri toplamı

$$P(Y_i = 1) = F(X_i' \beta) = F_i \quad (4.41)$$

$$SSR = \sum_{i=1}^N (Y_i - \hat{F}_i)^2 \quad (4.42)$$

- Ağırlıklı hata kareleri toplamı

$$WSSR = \sum_{i=1}^N \frac{(Y_i - \hat{F}_i)^2}{\hat{F}_i (1 - \hat{F}_i)} \quad (4.43)$$

- Efron' s R^2

Efron' un uyum iyiliği ölçüsü, iki değerli bağımlı değişkenlerin toplam değişimini açıklanan ve açıklanmayan değişim olarak ikiye ayırmaktadır. Sadece $Y_i = \hat{F}_i$ olduğunda uyum iyiliği ölçüsü değeri 1' e eşit olmaktadır.

$$R_{EF}^2 = 1 - \frac{\sum_{i=1}^N (Y_i - \hat{F}_i)^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2} \quad (4.44)$$

- Buse' s R^2

Formülde yer alan $WSSR_0$, açıklayıcı değişken olarak sadece sabit terimin yer aldığı modele ilişkin ağırlıklı hata kareleri toplamını ifade etmektedir.

$$R_{Buse}^2 = \frac{(WSSR_0 - WSSR)}{WSSR_0} \quad (4.45)$$

- Kareli korelasyon katsayısı(squared correlation coefficient)

$$S_{qcc} = \frac{\left(\sum_{i=1}^N (Y_i - \bar{Y})(\hat{F}_i - \bar{F}) \right)^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2 \sum_{i=1}^N (\hat{F}_i - \bar{F})^2} \quad (4.46)$$

c) Benzerlik fonksiyonuna dayalı olan ölçütler;

- McFadden' s R^2

$$R_{MF}^2 = 1 - \frac{\ln LL_1}{\ln LL_0}$$

- Maddala' s R^2

$$R_{MD}^2 = 1 - \left(\frac{\ln LL_1}{\ln LL_0} \right)^{2/N} \quad (4.47)$$

- Aldrich Nelson R^2

$$R_{AN}^2 = \frac{2(\ln LL_1 - \ln LL_0)}{2(\ln LL_1 - \ln LL_0) + N} \quad (4.48)$$

- Normalised AN R^2

$$R_{ANN}^2 = \frac{(2(\ln LL_1 - \ln LL_0)) / (2(\ln LL_1 - \ln LL_0) + N)}{(-2 \ln LL_0) / (N - 2 \ln LL_0)} \quad (4.49)$$

- Pseudo R^2

$$R_{PSD}^2 = \frac{e^{\frac{2}{N} \ln LL_1} - e^{\frac{2}{N} \ln LL_0}}{1 - e^{\frac{2}{N} \ln LL_0}} \quad (4.50)$$

- McKelvey ve Zovoina' s R^2

Probit model yardımıyla parametre tahminleri yapılacak olursa;

$\hat{Y}_i^* = X_i' \hat{\beta}$ elde edilir.

$$R_{MZ}^2 = \frac{\sum_{i=1}^N (\hat{Y}_i^* - \bar{Y}^*)^2}{\sum_{i=1}^N (\hat{Y}_i^* - \bar{Y}^*)^2 + N} \quad (4.51)$$

- Akaike bilgi kriteri

$$AIC = -2(\log(LL_1)/n) + 2(k/n) \quad (4.52)$$

- Veall-Zimmerman' ın uyum iyiliği ölçüsü

Modelin logaritmik benzerlik fonksiyonuna dayalıdır ve aşağıdaki gibi tanımlanmaktadır.

$$R_{VZ}^2 = \frac{2[InLL_1 - InLL_0]}{2[InLL_1 - InLL_0] + n} \frac{2InLL_0 - N}{2InLL_0} \quad (4.53)$$

$\alpha = N/(2InLL_0)$ olarak tanımlanırsa R_{VZ}^2 ile R_{MF}^2 arasındaki bağıntı yazılabilir.

$$R_{VZ}^2 = \frac{\alpha - 1}{\alpha - R_{MF}^2} \cdot R_{MF}^2 \quad (4.54)$$

Böylece, $R_{MF}^2 = 1$ veya $R_{MF}^2 = 0$ ise $R_{MF}^2 = R_{VZ}^2$ ve $0 < R_{MF}^2 < 1$ ise $R_{VZ}^2 > R_{MF}^2$ olur. α artarsa iki uygunluk ölçüsü arasındaki fark küçülür.³⁵

³⁵ Akın, F., Kalitatif Tercih Modelleri Analizi, ss. 56-57

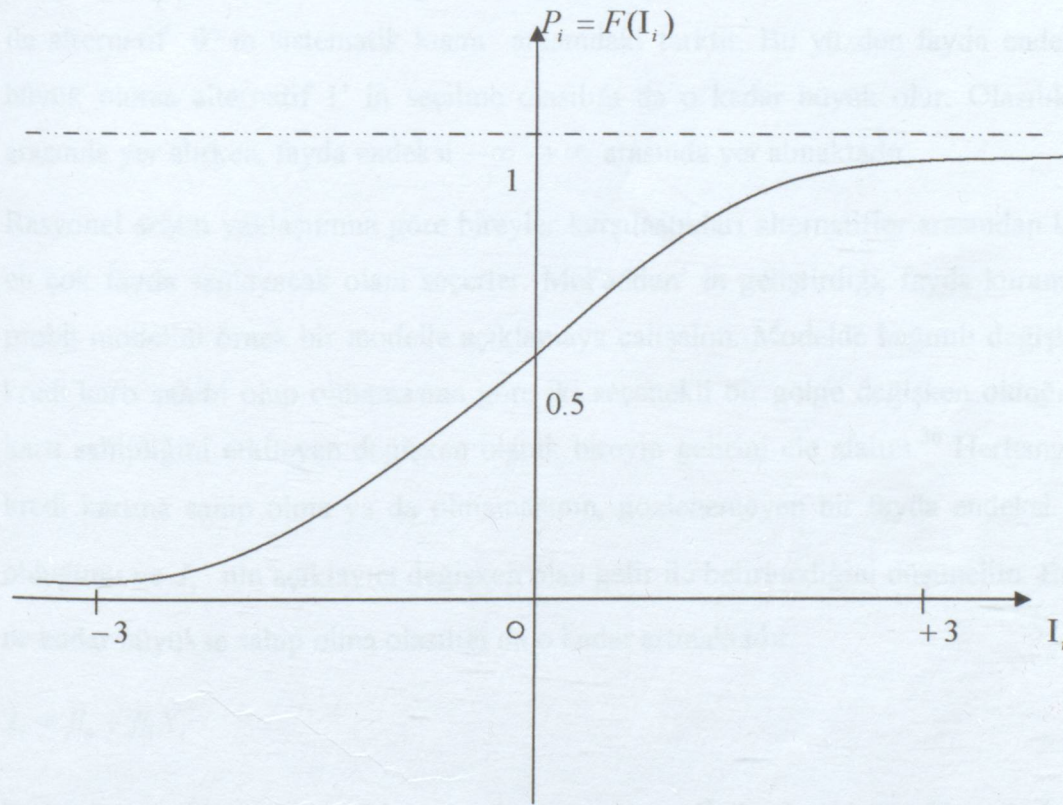
4.1.3 Probit Model

Doğrusal olasılık modelindeki en belirgin sorun olan “tahmin edilen olasılık değerlerinin 0 – 1 aralığı dışına çıkması” sorunundan kurtulmak için geliştirilen modellerden birisi de probit modelidir. Logit modelde birikimli lojistik dağılım fonksiyonu kullanılırken probit modelde birikimli normal dağılım fonksiyonu kullanılmaktadır.

X değişkeni μ ortalama ve σ^2 varyansla normal dağılıyorsa X değişkenine ait olasılık yoğunluk ve birikimli dağılım fonksiyonları şu şekilde olmaktadır:

$$f(X) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(X-\mu)^2 / 2\sigma^2}, -\infty < X < \infty \longrightarrow \text{OYF} \quad (4.55)$$

$$F(X) = \int_{-\infty}^{X_0} \frac{1}{\sqrt{2\pi}\sigma} e^{-(X-\mu)^2 / 2\sigma^2} \longrightarrow \text{BDF} \quad (4.56)$$



Şekil 4.4 Probit model

Olasılığın 0 – 1 arasında kalmasını sağlayan ve katsayılar itibariyle doğrusal olmayan probit modeli, genellikle gözlenemeyen bir fayda endeksi yardımıyla oluşturulduğundan, fayda endeksi hakkında bilgi verilmelidir. i . bireyin j . alternatiften elde ettiği fayda olan U_{ij} , sistematik (V_{ij}) ve tesadüfi (ε_{ij}) olmak üzere 2 parçaya ayrılarak gösterilebilir.

$$U_{ij} = V_{ij} + \varepsilon_{ij} \quad i = 1, 2, \dots, n \text{ ve } j = 1, 0 \quad (4.57)$$

Sistematik kısım olan V_{ij} , hem alternatiflerin fiyat, kalite, dayanıklılık, boyut gibi özelliklerinin, hem de bireylerin gelir, yaş, eğitim durumu, cinsiyet gibi özelliklerinin katsayılar itibarıyla doğrusal bir fonksiyonu şeklinde ifade edilmektedir. Hata terimini gösteren ε_{ij} ise ortalaması 0 ve varyansı da σ_j^2 olan bağımsız normal dağılımlı bir değişkendir. Sistematik kısım bilinse bile tesadüfi kısım gözlenemediğinden yapılacak tercih ile ilgili olasılıklı ifadelerde bulunulur ve i . bireyin 1. alternatifi seçme olasılığı aşağıdaki gibi belirtilir.

$$P_i = P(Y_i = 1) = P(U_{i1} \geq U_{i0}) = P(V_{i1} + \varepsilon_{i1} \geq V_{i0} + \varepsilon_{i0}) = P(\varepsilon_{i0} - \varepsilon_{i1} \leq V_{i1} - V_{i0}) \quad (4.58)$$

Olasılık ifadesinde yer alan $(V_{i1} - V_{i0})$ terimi fayda endeksi olarak tanımlanmakta ve I_i simgesiyle gösterilmektedir. Diğer bir ifadeyle fayda endeksi, alternatif 1' in sistematik kısmı ile alternatif 0' in sistematik kısmı arasındaki farktır. Bu yüzden fayda endeksi ne kadar büyük olursa alternatif 1' in seçilme olasılığı da o kadar büyük olur. Olasılık değeri 0-1 arasında yer alırken, fayda endeksi $-\infty + \infty$ arasında yer almaktadır.

Rasyonel seçim yaklaşımına göre bireyler karşılaştıkları alternatifler arasından kendileri için en çok fayda sağlayacak olanı seçerler. McFadden' in geliştirdiği, fayda kuramına dayanan probit modelini örnek bir modelle açıklamaya çalışalım. Modelde bağımlı değişkenin kişinin kredi kartı sahibi olup olmamasına göre iki seçeneqli bir gölge değişken olduğunu ve kredi kartı sahipliğini etkileyen değişken olarak bireyin gelirini ele alalım.³⁶ Herhangi bir bireyin kredi kartına sahip olma ya da olmamasının, gözlenemeyen bir fayda endeksi I_i ' ye bağlı olduğunu ve I_i ' nin açıklayıcı değişken olan gelir ile belirlendiğini düşünelim. Endeks değeri ne kadar büyükse sahip olma olasılığı da o kadar artmaktadır.

$$I_i = \beta_0 + \beta_1 X_i \quad (4.59)$$

Endeks değerinin kendisi gibi gözlemlenemeyen ve I_i^* ile simgelenen bir eşik değerine sahip olduğunu düşünecek olursak, eğer I_i değeri I_i^* ' yi aşarsa kişinin kredi kartına sahip olması, aşmazsa sahip olmaması beklenir. Normallik varsayımı altında, I_i^* değerinin I_i ' den küçük ya da I_i ' ye eşit olma olasılığı, standartlaştırılmış birikimli dağılım fonksiyonundan hareketle hesaplanmaktadır.

³⁶ Ertek, T., Ekonometriye Giriş, s.293

$$P_i = P(Y = 1) = P(I_i^* \leq I_i) = F(I_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{I_i} e^{-z^2/2} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta_0 + \beta_1 X_i} e^{-z^2/2} dz \quad (4.60)$$

Z değişkeni 0 ortalama ve 1 varyansa sahip standartlaştırılmış normal değişkendir.

Bireyin kredi kartı sahipliğini ölçen I_i endeksi X bağımsız değişkenlerinin bir fonksiyonudur ve endeks değerinin yüksek olması olasılığın fazla olması, endeks değerinin düşük olması ise olasılığın az olması anlamına gelebilmektedir. I_i endeks değeri ile I_i^* kritik değeri arasındaki ilişkiden yararlanılarak olasılık formülleri ve benzerlik fonksiyonu yazılır.

$$I_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

$$Y_i = \begin{cases} 1 & \text{eğer ; } I_i \geq I_i^* \\ 0 & \text{eğer ; } I_i < I_i^* \end{cases}$$

I_i^* stokastik bir değişkendir ve aşağıdaki olasılık formülünde yer almaktadır.

$$P(Y_i = 1 / I_i) = P(I_i \geq I_i^*) \longrightarrow \text{kredi kartı sahibi olma}$$

$$P(Y_i = 0 / I_i) = P(I_i < I_i^*) \longrightarrow \text{kredi kartı sahibi olmama}$$

Çekilen örnekte n_1 adet bireyin kredi kartı sahibi olduğunu, n_2 adet bireyin ise sahip olmadığını farzederek örneğe ilişkin benzerlik fonksiyonunu yazalım.

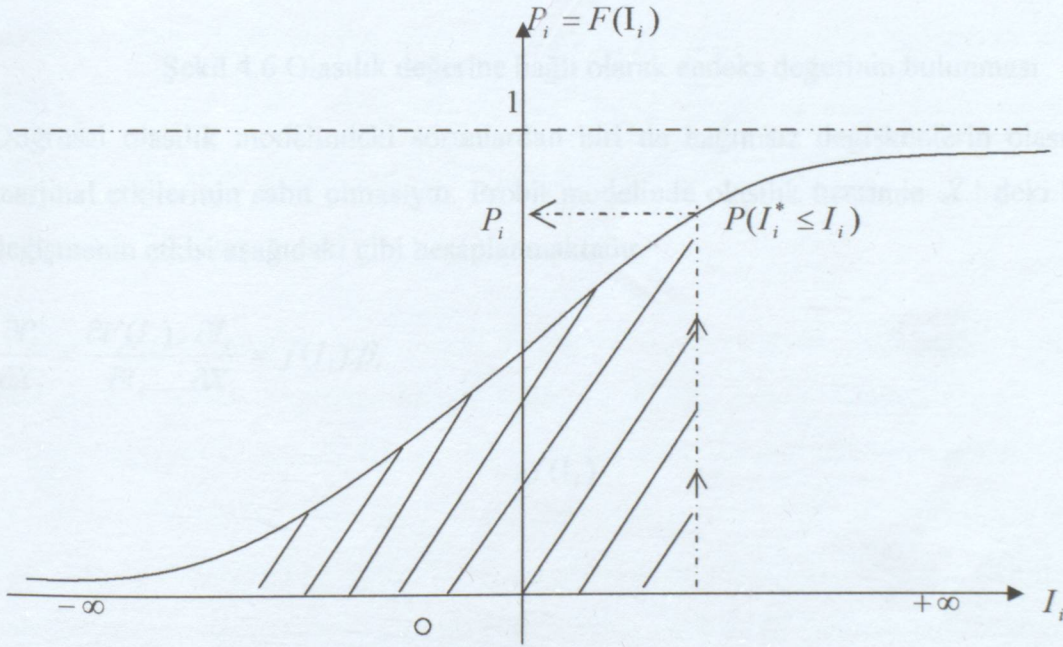
$$LL = P(I_1 \geq I_1^*) . P(I_2 \geq I_2^*) . \dots . P(I_{n_1} \geq I_{n_1}^*) . P(I_{n_1+1} < I_{n_1+1}^*) . \dots . P(I_{n_1+n_2} < I_{n_1+n_2}^*) \quad (4.61)$$

$F(I_i)$ simgesinin birikimli dağılım fonksiyonunu gösterdiğini hatırlayarak benzerlik fonksiyonunu düzenleyecek olursak aşağıdaki ifadeyi elde ederiz.

$$LL = F(I_1) . F(I_2) . \dots . F(I_{n_1}) . [1 - F(I_{n_1+1}) . \dots . 1 - F(I_{n_1+n_2})] \quad (4.62)$$

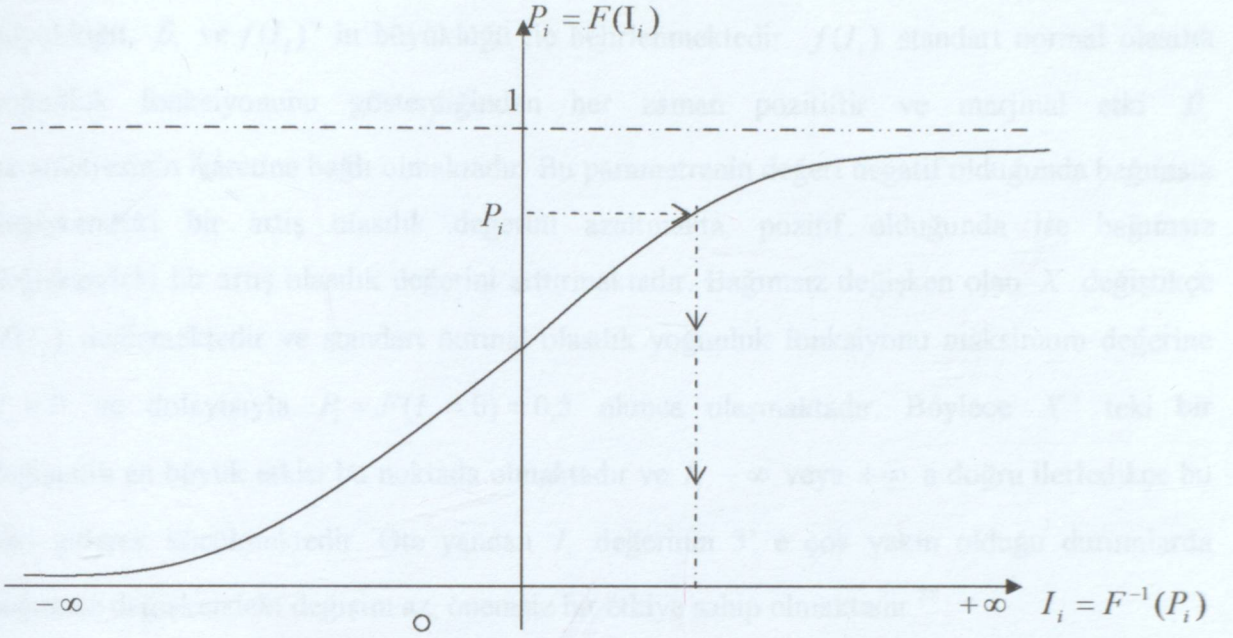
Benzerlik fonksiyonu, endeks değerinin hesaplanmasında kullanılan regresyon katsayılarının ve bağımsız değişkenlerin bir fonksiyonu olmaktadır. Regresyon katsayıları benzerlik fonksiyonunu maksimum yapacak şekilde belirlenmelidir. LL 'yi maksimum yapacak katsayı değerlerinin tahmini için de nümerik iterasyon yöntemleri kullanılmaktadır.³⁷

P_i kredi kartına sahip olma olasılığını göstermekte ve aşağıdaki çizimlerde gösterildiği gibi, standart normal eğrinin altında, $-\infty$ ' dan I_i ' ye kadar olan alanla ölçülmektedir. Şekil (4.5)' de görüldüğü üzere, endeks değeri arttıkça olasılık değeri de artmaktadır fakat bu artış doğrusal olasılık modelindeki gibi sabit değildir. Olasılığın 0 ve 1' e yakın olduğu bölgelerde endeksteeki eşit miktardaki değişimler olasılık değerinde giderek daha küçük değişimlere yol açmaktadır. Şekil (4.6)' da P_i belli iken endeks değerinin bulunması gösterilmektedir.



Şekil 4.5 Endeks değerine bağlı olarak olasılık bulunması

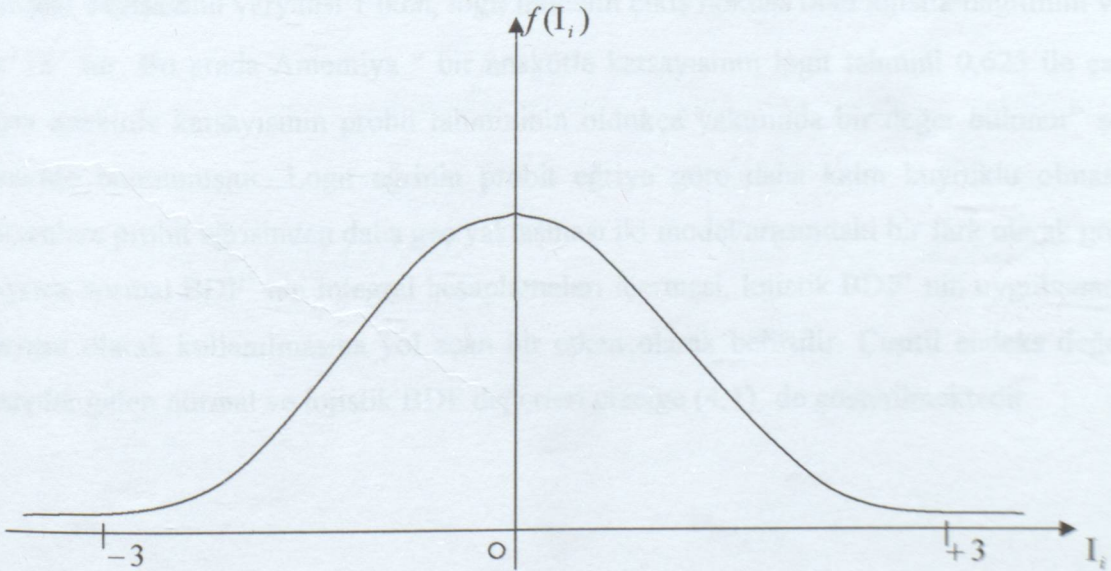
³⁷ Frank, R. C., a.g.e, ss. 345-346



Şekil 4.6 Olasılık değerine bağlı olarak endeks değerinin bulunması

Doğrusal olasılık modelindeki sorunlardan biri de bağımsız değişkenlerin olasılık üzerine marjinal etkilerinin sabit olmasıydı. Probit modelinde olasılık üzerinde X 'deki bir birimlik değişimin etkisi aşağıdaki gibi hesaplanmaktadır.

$$\frac{\partial P_i}{\partial X_i} = \frac{\partial F(I_i)}{\partial I_i} \cdot \frac{\partial I_i}{\partial X_i} = f(I_i) \cdot \beta_i \quad (4.63)$$



Şekil 4.7 Standart normal olasılık yoğunluk fonksiyonu

Bağımsız değişkendeki bir değişimin olasılık değerinde meydana getireceği değişimin büyüklüğü, β_i ve $f(I_i)$ ' in büyüklüğü ile belirlenmektedir. $f(I_i)$ standart normal olasılık yoğunluk fonksiyonunu gösterdiğinden her zaman pozitifdir ve marjinal etki β_i parametresinin işaretine bağlı olmaktadır. Bu parametrenin değeri negatif olduğunda bağımsız değişkendeki bir artış olasılık değerini azaltmakta, pozitif olduğunda ise bağımsız değişkendeki bir artış olasılık değerini arttırmaktadır. Bağımsız değişken olan X değiştiğinde $f(I_i)$ değişmektedir ve standart normal olasılık yoğunluk fonksiyonu maksimum değerine $I_i = 0$ ve dolayısıyla $P_i = F(I_i = 0) = 0,5$ olunca ulaşmaktadır. Böylece X ' teki bir değişimin en büyük etkisi bu noktada olmaktadır ve $I_i - \infty$ veya $+\infty$ a doğru ilerledikçe bu etki giderek küçülmektedir. Öte yandan I_i değerinin 3' e çok yakın olduğu durumlarda bağımsız değişkendeki değişim az, önemsiz bir etkiye sahip olmaktadır.³⁸

Gözlenemeyen fayda endeksi olan I_i , normal eşdeğeri sapma (n.e.s) ya da normit olarak bilinmektedir. $P_i < 0,5$ olduğunda I_i eksi değer aldığından uygulamada n.e.s değerine 5 eklenmektedir ve bulunan sonuca da probit denilmektedir.

$$Probit = n.e.s. + 5 = I_i + 5 \quad (4.64)$$

Logit ve probit modeller nitel olarak benzer sonuçlar verse de, iki modelin tahmin edilen parametre değerleri doğrudan karşılaştırılmaz. Çünkü probit modelin temeli olan standart normal değişkenin varyansı 1 iken, logit modelin çıkış noktası olan lojistik dağılımın varyansı $\pi^2/3$ tür. Bu arada Amemiya “ bir anakütle katsayısının logit tahmini 0,625 ile çarpılırsa aynı anakütle katsayısının probit tahmininin oldukça yakınında bir değer bulunur” şeklinde öneride bulunmuştur. Logit eğrinin probit eğriye göre daha kalın kuyruklu olması, yani eksenlere probit eğrisinden daha geç yaklaşması iki model arasındaki bir fark olarak gösterilir. Ayrıca normal BDF' nin integral hesaplamaları içermesi, lojistik BDF' nin uygulamada daha yaygın olarak kullanılmasına yol açan bir etken olarak belirtilir. Çeşitli endeks değerlerine karşılık gelen normal ve lojistik BDF değerleri çizelge (4.1)' de gösterilmektedir.

³⁸ Griffiths vd., Undergraduate Econometrics, ss. 202-203

Çizelge 4.1 Endeks değerine bağlı olarak normal ve lojistik bdf değerleri
(Pindyck ve Rubinfeld , s. 258)

I_i	$P_i = F(I_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{I_i} e^{-Z^2/2} .dZ$	$P_i = F(I_i) = \frac{1}{1+e^{-I_i}}$
-3.0	0.0013	0.0474
-2.0	0.0228	0.1192
-1.5	0.0668	0.1824
-1.0	0.1587	0.2689
-0.5	0.3085	0.3775
0.0	0.5000	0.5000
0.5	0.6915	0.6225
1.0	0.8413	0.7311
1.5	0.9332	0.8176
2.0	0.9772	0.8808
3.0	0.9987	0.9526

İkili tercih modellerinde katsayı tahminçileri arasındaki ilişkiler aşağıda gösterilmektedir.

Çizelge 4.2 İkili tercih modellerinde tahminçiler arasındaki ilişkiler (Intriligator vd., s.163)

Sabit Terimler Arasında	Eğim Katsayıları Arasında
$\hat{\beta}_{Pr\ obit} \cong 0.625\hat{\beta}_{Logit}$	$\hat{\beta}_{Pr\ obit} \cong 0.625\hat{\beta}_{Logit}$
$\hat{\beta}_{DOM} \cong 0.4\hat{\beta}_{Pr\ obit} + 0.5$	$\hat{\beta}_{DOM} \cong 0.4\hat{\beta}_{Pr\ obit}$
$\hat{\beta}_{DOM} \cong 0.25\hat{\beta}_{Logit} + 0.5$	$\hat{\beta}_{DOM} \cong 0.25\hat{\beta}_{Logit}$

Probit ve logit modelde değişen varyans sorununa çözüm olarak, hata teriminin varyansının karekökü tartı olarak kullanılır.

$$\sigma_u^2 = \frac{1}{N_i P_i (1 - P_i)} \quad \text{Logit model} \quad (4.65)$$

$$\sigma_u^2 = \frac{P_i (1 - P_i)}{N_i f_i^2} \quad \text{Probit model} \quad (4.66)$$

Her birinin hata terimlerinin varyansları dikkate alınarak bu modellere tartılı en küçük kareler uygulanması, minimum χ^2 yöntemi olarak bilinmektedir.

N adet gözlemin yer aldığı bir örneklem ele alındığında, N ' in büyümesi durumunda parametre tahmincileri yaklaşık olarak normal dağılmaktadır. Tahmincilerin ortalaması β ve kovaryans matrisi $\text{cov}(\hat{\beta})$ olmaktadır.³⁹

$$\text{cov}(\hat{\beta}) = (X' K X)^{-1} \quad (4.67)$$

M adet bağımsız değişkenin yer aldığı bir modelde X matrisi $(N * M)$ boyutlu veri matrisi ve $K = \text{diag}(d_1, d_2, \dots, d_N)$ olan diagonal bir matristir.

$$d_i = \frac{[f(X_i' \beta)]^2}{F(X_i' \beta) [1 - F(X_i' \beta)]} \quad (4.68)$$

Parametre tahmincilerinin tek tek anlamlılıkları test edilmek istendiğinde, aşağıda belirtilen t değerleri elde edilmekte ve elde edilen bu değerler istenilen anlamlılık düzeyinde tablo değerleriyle karşılaştırılarak hipotezlerin reddine veya kabulüne karar verilmektedir.

$$H_0 : \hat{\beta}_k = 0$$

$$H_1 : \hat{\beta}_k \neq 0$$

t değerlerinin elde edilmesinde ilgili katsayının değeri ve standart hatası kullanılmaktadır.⁴⁰

$$t = \frac{\hat{\beta}_k}{se(\hat{\beta}_k)} \quad (4.69)$$

³⁹ Griffiths vd., Learning and Practicing Econometrics, s.746

⁴⁰ Griffiths vd., a.g.e, s.748

Katsayıların birlikte anlamlılıkları test edilmek istendiğinde ise hesaplanan χ^2 değerleri istenilen anlamlılık düzeyinde tablo değerleriyle karşılaştırılarak hipotezlerin reddine veya kabulüne karar verilmektedir.

$$H_0 : R\hat{\beta} = r$$

$$H_1 : R\hat{\beta} \neq r$$

R , diagonalinde 1'lerin yer aldığı ve diğer elemanları 0 olan bir kare matris, r ise anlamlılığı test edilmek istenen parametre sayısı (a) kadar 0 içeren bir sütun matrisi ifade etmektedir.⁴¹

$$(R\hat{\beta} - r)'[R \text{cov}(\hat{\beta})R']^{-1}(R\hat{\beta} - r) \approx \chi^2_{(a)} \quad (4.70)$$

⁴¹ Griffiths vd., a.g.e, s.749

4.2 ÇOKLU TERCİH MODELLERİ

Nitel tercih modelleri arasında en basit olan, bağımlı değişkenin iki değer aldığı ikili tercih modelleridir. Daha karmaşık olan çoklu tercih modellerinde karar birimleri ikiden fazla tercihle karşı karşıyadır ve dolayısıyla bağımlı değişken ikiden fazla değer almaktadır.

Bağımlı değişkenin kesikli değer alması onun nitel bir değişken olduğunu göstermez, dolayısıyla, özellikle çoklu tercih modellerinde bir karışıklığa fırsat vermemek bakımından değişkenler, kategorik ve kategorik olmayan değişkenler biçiminde sınıflandırılabilir. ⁴² Bireylerin gelir, harcama, meslek gibi özelliklerine göre gruplandırılmaları halinde kullanılan ve 1, 2, 3 gibi kesikli değerler alan değişkenler kategorik değişkendir. Buna karşılık belli bir dönemde bir öğrencinin okula gitmeme sayısı, çağrı merkezine gelen çağrı sayısı veya bir öğretim elemanının yayınlattığı bilimsel makale sayısı gibi belirli bir zaman aralığında meydana gelen olayları temsilen kullanılan ve yine 0, 1, 2, 3 gibi kesikli değerler alan değişkenler kategorik olmayan değişkenler olup, Poisson regresyon modeli ile analiz edilebilmektedir.

Kategorik değişkenler de ayrıca sıralı, sıralı olmayan ve ardışık değişkenler olarak sınıflandırılabilir. ⁴³ Bu değişkenler kullanılarak sıralı, sıralı olmayan ve ardışık tercih modelleri oluşturulmaktadır.

Bireylerin gelirlerine göre, aşağıdaki gibi belirlenen değişken sıralı bir değişkendir.

$Y = 1$ geliri 500 YTL' den az ise

$Y = 2$ geliri 500 YTL ile 749 YTL arasında ise

$Y = 3$ geliri 750 YTL' den fazla ise

Bireyler mesleklerine göre gruplandırılırsa, sıralı olmayan bir değişken elde edilir.

$Y = 1$ birey doktor ise

$Y = 2$ birey işçi ise

$Y = 3$ birey serbest meslek sahibi ise

$Y = 4$ birey gazeteci ise

⁴² Maddala, G. S., Limited Dependent and Qualitative Variables in Econometrics, s.14

⁴³ Cox, D. R., Analysis Of Binary Data, Bölüm 7

Bireyler eğitim durumlarına göre kategorilere ayrılırsa, ardışık bir değişken elde edilir.

$Y = 1$ ilkokul mezunu ise

$Y = 2$ ortaokul mezunu ise

$Y = 3$ lise mezunu ise

$Y = 4$ üniversite mezunu ise

$Y = 5$ yüksek lisans mezunu ise

$Y = 6$ doktora yapmış ise

4.2.1 Sıralı Olmayan Tercih Modelleri

4.2.1.1 Çok Durumlu Doğrusal Olasılık Modeli

Çok durumlu DOM, iki alternatif için söz konusu olan doğrusal olasılık modelinin üç veya daha fazla karşılıklı dışlayıcı alternatif için genelleştirilmesinden oluşmaktadır. Karşılıklı dışlayıcı terimiyle anlatılmak istenen, bireyin alternatifler arasından sadece birini tercih edebilmesidir. Üç alternatifin yer aldığı bir modeli ele alalım;

P_{ji} , i . bireyin j . alternatifi seçme olasılığını göstermektedir. Her bir alternatife ilişkin seçim olasılıkları yazılabilmektedir.

$$\begin{aligned} P_{1i} &= \beta_0 + \beta_1 X_i \\ P_{2i} &= \beta_2 + \beta_3 X_i \\ P_{3i} &= \beta_4 + \beta_5 X_i \end{aligned} \quad (4.71)$$

Olasılıklara ilişkin kısıtlar dikkate alınarak denklemlerden sadece ikisine tahmin yönteminin uygulanması yeterli olacaktır. Tahmin edilen olasılıkların toplamının 1' e eşit olması gerektiğinden dolayı tahmin edilen sabit terimlerin toplamı 1' e ve eğim katsayılarının toplamı da 0' a eşit olur. Olasılıklar P_{ji} , olasılıkların ortalamaları $\bar{P}_j$ ve olasılıkların ortalamalarından farkları da p_{ji} olarak simgelenirse aşağıdaki kısıtlamalar elde edilir.

$$\begin{aligned} P_{1i} + P_{2i} + P_{3i} &= 1 \\ \bar{P}_1 + \bar{P}_2 + \bar{P}_3 &= 1 \\ p_{1i} + p_{2i} + p_{3i} &= 0 \end{aligned} \quad (4.72)$$

Bu kısıtlamalardan hareketle, modelin en küçük kareler yöntemiyle tahmin edilmesi halinde eğim katsayılarının tahmincilerinin toplamının 0' a ve sabit terim tahmincilerinin toplamının da 1' e eşit olduğu gösterilebilmektedir.

$$\hat{\beta}_1 + \hat{\beta}_3 + \hat{\beta}_5 = \frac{\sum p_{1i}x_i}{\sum x_i^2} + \frac{\sum p_{2i}x_i}{\sum x_i^2} + \frac{\sum p_{3i}x_i}{\sum x_i^2} = \frac{\sum (p_{1i} + p_{2i} + p_{3i})x_i}{\sum x_i^2} = 0 \quad (4.73)$$

$$\begin{aligned} \hat{\beta}_0 + \hat{\beta}_2 + \hat{\beta}_4 &= (\bar{P}_1 - \hat{\beta}_1\bar{X}) + (\bar{P}_2 - \hat{\beta}_3\bar{X}) + (\bar{P}_3 - \hat{\beta}_5\bar{X}) \\ &= (\bar{P}_1 + \bar{P}_2 + \bar{P}_3) - (\hat{\beta}_1 + \hat{\beta}_3 + \hat{\beta}_5)\bar{X} = 1 \end{aligned} \quad (4.74)$$

Sonuç olarak;

$$\begin{aligned} \hat{P}_{1i} + \hat{P}_{2i} + \hat{P}_{3i} &= (\hat{\beta}_0 + \hat{\beta}_1X_i) + (\hat{\beta}_2 + \hat{\beta}_3X_i) + (\hat{\beta}_4 + \hat{\beta}_5X_i) \\ &= (\hat{\beta}_0 + \hat{\beta}_2 + \hat{\beta}_4) + (\hat{\beta}_1 + \hat{\beta}_3 + \hat{\beta}_5)X_i = 1 \end{aligned} \quad (4.75)$$

elde edilir. Çok durumlu doğrusal olasılık modeli tahmin edilirken denklemlerden ikisi en küçük kareler yöntemi ile tahmin edilir ve elde edilen tahmin sonuçları kullanılarak üçüncü denklemin katsayıları bulunur.

4.2.1.2 Çok Durumlu Probit Model

Çok durumlu probit modeli tesadüfî fayda modelindeki hata terimlerinin çok değişkenli normal dağılıma sahip olduğu bir model olarak ortaya çıkmaktadır. İlk olarak Thurstone (1927) tarafından ileri sürülen bu model Bock ve Jones (1968) tarafından psikolojik tercih verilerine, Haussman ve Wise (1978) tarafından ulaşım tercih verilerine uygulanmıştır. Bu model çoklu integrallerin hesaplanmasını gerektirdiğinden uygulamada pek fazla tercih edilmemektedir. 3 alternatifin söz konusu olduğu bir durumu ele alacak olursak;

$$\begin{aligned} Y_1^* &= \beta_1'X + \varepsilon_1 \\ Y_2^* &= \beta_2'X + \varepsilon_2 \\ Y_3^* &= \beta_3'X + \varepsilon_3 \end{aligned} \quad (4.76)$$

elde edilir.

Hata terimlerinin ortalama vektörü 0 ve varyans - kovaryans matrisi de Σ olmaktadır.

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{bmatrix} \quad (4.77)$$

1. alternatifin seçilme olasılığı,⁴⁴

$$P_1 = \int_{-\infty}^{V_{12}} \int_{-\infty}^{V_{13}} f(\eta_{21}, \eta_{31}) d\eta_{21} d\eta_{31} \quad (4.78)$$

$f(\eta_{21}, \eta_{31})$ ortalama vektörü 0 ve kovaryans matrisi Ω_1 olan normal dağılıma sahiptir. Diğer bir ifadeyle $f(\eta_{21}, \eta_{31})$, sıfır ortalama vektörü ve Ω_1 kovaryans matrisli iki değişkenli normal dağılıma sahip bileşik olasılık fonksiyonudur.

$$\Omega_1 = \begin{bmatrix} \sigma_1^2 + \sigma_2^2 - 2\sigma_{12} & \sigma_1^2 - \sigma_{13} - \sigma_{12} + \sigma_{23} \\ \sigma_1^2 - \sigma_{13} - \sigma_{12} + \sigma_{23} & \sigma_1^2 + \sigma_3^2 - 2\sigma_{13} \end{bmatrix} \quad (4.79)$$

2. ve 3. alternatiflere ilişkin seçilme olasılıkları aşağıdaki gibi hesaplanabilmektedir.

$$P_2 = \int_{-\infty}^{V_{21}} \int_{-\infty}^{V_{23}} f(\eta_{12}, \eta_{32}) d\eta_{12} d\eta_{32} \quad (4.80)$$

$$P_3 = \int_{-\infty}^{V_{31}} \int_{-\infty}^{V_{32}} f(\eta_{13}, \eta_{23}) d\eta_{13} d\eta_{23} \quad (4.81)$$

Dört alternatifli bir durumda üçlü integrallerin hesaplanması gerektiği ve bunun da çok zor olabileceği kolayca görülebilir. Ayrıca, dörtten fazla alternatif için hesaplamalar hemen hemen elverişsizdir.⁴⁵

4.2.1.3 Çok Durumlu Logit Model

İki alternatifin söz konusu olduğu ikili logit modelin, üç veya daha fazla alternatifli durum için genelleştirilmesinden meydana gelen bu modelde rasgele fayda modelinden yararlanılmaktadır. Bir alternatifle ilişkin rasgele fayda gözlenen veya sistematik bileşen ile gözlenemeyen bileşen olmak üzere 2 kısımda incelenmektedir. V_{in} ve V_{jn} , i . ve

⁴⁴ $V_{12} = V_1 - V_2, V_{13} = V_1 - V_3, \eta_{21} = \varepsilon_2 - \varepsilon_1, \eta_{31} = \varepsilon_3 - \varepsilon_1$ olarak simgelenmekte ve ilerleyen bölümlerde ayrıntılı olarak ele alınmaktadır

⁴⁵ Maddala, G. S., a.g.e., ss. 62-64

j . alternatiflere ilişkin faydanın sistematik kısmını, ε_{in} ve ε_{jn} ise gözlenemeyen yani rasgele kısımları belirtmektedir.

$$U_{in} = V(z_{in}, S_n) + \varepsilon(z_{in}, S_n) = V_{in} + \varepsilon_{in} \quad (4.82)$$

$$P_n(j) = P(U_{jn} \geq U_{in}) = P(V_{jn} + \varepsilon_{jn} \geq V_{in} + \varepsilon_{in}) = P(\varepsilon_{in} - \varepsilon_{jn} \leq V_{jn} - V_{in}) \quad (4.83)$$

(4.82) ve (4.83) numaralı formüllerde n . bireyin i . alternatifi seçmesiyle elde edeceği fayda ve aynı bireyin j . alternatifi seçme olasılığı gösterilmektedir. J adet alternatifle karşılaşan bir bireyin j . alternatifi seçme olasılığı, çok durumlu logit modelde aşağıdaki gibi gösterilmektedir.

$$P_n(j) = \frac{e^{V_{jn}}}{\sum_{k=1}^J e^{V_{kn}}} = \frac{e^{\beta'X_{jn}}}{\sum_{k=1}^J e^{\beta'X_{kn}}} \quad (4.84)$$

Çok durumlu logit modelin daha kolay anlaşılması açısından 3 alternatifin söz konusu olduğu ve 3 adet bağımsız değişkenin yer aldığı çoklu seçim modeli incelenecektir. Bu modelde n . bireyin alternatifleri seçmesiyle elde edeceği fayda fonksiyonları ve fayda fonksiyonlarının sistematik kısımlarının gösterilmesi gerekmektedir.

$$\begin{aligned} V_{jn} &= \beta_{j0} + \beta_{j1}X_{1n} + \beta_{j2}X_{2n} + \beta_{j3}X_{3n} \\ V_{1n} &= \beta_{10} + \beta_{11}X_{1n} + \beta_{12}X_{2n} + \beta_{13}X_{3n} \\ V_{2n} &= \beta_{20} + \beta_{21}X_{1n} + \beta_{22}X_{2n} + \beta_{23}X_{3n} \\ V_{3n} &= \beta_{30} + \beta_{31}X_{1n} + \beta_{32}X_{2n} + \beta_{33}X_{3n} \end{aligned} \quad (4.85)$$

$$\begin{aligned} U_{jn} &= V_{jn} + \varepsilon_{jn} \\ U_{1n} &= V_{1n} + \varepsilon_{1n} \\ U_{2n} &= V_{2n} + \varepsilon_{2n} \\ U_{3n} &= V_{3n} + \varepsilon_{3n} \end{aligned} \quad (4.86)$$

Yapılan tanımlamalar doğrultusunda n . bireyin 1. alternatifi seçme olasılığı aşağıda gösterildiği gibi hesaplanmaktadır.

$$P_n(1) = \frac{e^{V_{1n}}}{\sum_{k=1}^3 e^{V_{kn}}} \quad (4.87)$$

4.2.1.4 Koşullu Logit Model (Conditional Logit Model)

Karar biriminin J alternatifle karşı karşıya olduğunu, j . alternatifi tercih etmenin faydasının aşağıdaki gibi belirtildiğini ve bu faydanın diğer alternatifler arasında maksimum olduğunu varsayalım.

$$U_{jn} = V(z_{jn}, S_n) + \varepsilon(z_{jn}, S_n) = V_{jn} + \varepsilon_{jn} \quad (4.87)$$

Fayda fonksiyonu tercihin kendisi ile ilgili olan ve birey tarafından değerlendirilen özelliklerle ilgili vektör (kişinin yolculuk tercihinde seçtiği alternatife ilişkin yolculuk süresi, yolculuk giderleri, yolculuğun rahatlık ve güvenlik düzeyi gibi değişkenler) ile bireyin karakteristik özellikleriyle (gelir, yaş, cinsiyet, eğitim durumu gibi değişkenler) ilgili vektör olmak üzere 2 kısımdan oluşmaktadır.⁴⁶ Alternatiflerin niteliklerinin, alternatifler arasında farklı olmasının yanı sıra bu nitelikler bireyden bireye de değişebilecektir. Buna karşılık, bireylerin özellikleri bireyler arasında farklı ancak tüm alternatifler için aynı olacaktır.⁴⁷

$$U_{jn} = U(z_{jn}, S_n) \quad (4.88)$$

$$U_{jn} = \alpha'_j S_n + \beta' z_{jn} + \varepsilon_{jn} \quad (4.89)$$

Mc Fadden (1973) hata terimlerinin birbirinden bağımsız ve Weibull dağılımına uymaları durumunda olasılığı şu şekilde ifade etmiştir.

$$P(Y_i = j) = \frac{e^{\beta' z_{jn} + \alpha'_j S_n}}{\sum_{j=1}^J e^{\beta' z_{kn} + \alpha'_k S_n}} \quad (4.90)$$

Weibull dağılımının ζ ile simgelenen yalnızca 1 parametresi vardır. Birikimli Weibull dağılım fonksiyonu aşağıdaki gibidir.

$$F(\varepsilon_j) = P(\varepsilon_j < \varepsilon) = \exp\left(-e^{-(\varepsilon + \zeta_j)}\right) \quad (4.91)$$

Yoğunluk fonksiyonu da şu şekilde yazılmaktadır.

$$f(\varepsilon) = e^{-(\varepsilon + \zeta_j)} \exp\left(-e^{-(\varepsilon + \zeta_j)}\right) \quad (4.92)$$

Parametre değeri 0' a eşit olduğunda, dağılımın mod değeri 0, ortalaması 0,575 ve varyansı 1,622 olmaktadır. Standart normal dağılım yoğunluk fonksiyonunun 0 ortalama ve birim

⁴⁶ Lerman, R. S. ve Akiva, B. M., a.g.e, s. 60

⁴⁷ Greene, W. H. , Econometric Analysis (3rd Edition), s. 914

varyansa sahip olduğu göz önüne alınırsa, $\zeta = 0$ durumunda weibull yoğunluk fonksiyonunun daha çarpık ve daha kalın kuyruklu olduğu anlaşılmaktadır.⁴⁸

McFadden, bağımsız ε_j ($j = 1, 2, \dots, J$) için birikimli Weibull dağılım fonksiyonunu şöyle geliştirmiştir;

$$F(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_j) = \exp \left[-G(\exp(-\varepsilon_1), \exp(-\varepsilon_2), \dots, \exp(-\varepsilon_j)) \right]$$

G ile belirlenen, genelleştirilmiş birikimli Weibull dağılım fonksiyonundan türetilen modeller sınıfına Genelleştirilmiş Uç Değer Modeli (GEV) denir.⁴⁹

Koşullu logit model, ε_j ' lerin uç değer yoğunluk fonksiyonlarına sahip bağımsız ve tanımlı olarak dağıldığı varsayımına dayanır. Bu dağılıma sahip iki tesadüfi değişken arasındaki fark, iki durumlu logit modeli veren lojistik bir dağılım fonksiyonuna sahiptir.⁵⁰

4.2.1.5 Yuvalanmış Logit Model

McFadden' in literatüre kazandırdığı genelleştirilmiş uç değer dağılımının özel bir durumu olan yuvalanmış logit modelde alternatifler alt gruplara ayrılmakta ve tercih yapılırken önce grup seçimi yapılmakta ve daha sonra bu gruptaki belli bir alternatif tercih edilmektedir. İlk aşamada genelleştirilmiş uç değer modelinin özellikleri üzerinde durulacak ve daha sonraki aşamada da yuvalanmış model incelenecektir.

4.2.1.5.1 Genelleştirilmiş Uç Değer Modeli (Generalized Extreme Value Model)

Model tanımlaması aşağıdaki gibi yapılmaktadır.

$y_1, y_2, y_3, \dots, y_{Jn} \geq 0$ için $G(y_1, y_2, y_3, \dots, y_{Jn})$ aşağıdaki özelliklere sahip bir fonksiyon olsun:

1) G pozitifdir.

$$2) G(\alpha y_1, \alpha y_2, \alpha y_3, \dots, y_{Jn}) = \alpha^\mu G(y_1, y_2, y_3, \dots, y_{Jn}) \quad (4.93)$$

$$3) \lim_{y_i \rightarrow \infty} G(y_1, y_2, y_3, \dots, y_{Jn}) = \infty; i = 1, 2, 3, \dots, J_n \quad (4.94)$$

⁴⁸ Chow, C. G., Econometrics, s.258

⁴⁹ Akin, F., a.g.e, s.89

⁵⁰ Hausman, J. A. ve Wise, D. A., "A Conditional Probit Model for Qualitative Choice Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences", Vol.46

$G_i(y_1, y_2, y_3, \dots, y_{J_n}) = \partial G / \partial y_i$ olarak belirtilirse GEV modeline ilişkin olasılık şu şekilde hesaplanır.

$$P_n(i) = \frac{e^{V_{in}} G_i(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{J_n n}})}{\mu G(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{J_n n}})} \quad (4.95)$$

$G = \sum_{j=1}^{J_n} y_j^\mu$ olarak tanımlanan modelin yukarıda bahsedilen 3 özelliği de sağladığını gösterecek olursak;

1) $y_1, y_2, y_3, \dots, y_{J_n} \geq 0$ ise $G \geq 0$ olmaktadır.

$$2) G(\alpha y_1, \alpha y_2, \alpha y_3, \dots, \alpha y_{J_n}) = \sum_{j=1}^{J_n} (\alpha y_j)^\mu = \alpha^\mu \sum_{j=1}^{J_n} y_j^\mu = \alpha^\mu G(y_1, y_2, \dots, y_{J_n})$$

$$3) \lim_{y_i \rightarrow \infty} G = \infty; i = 1, 2, 3, \dots, J_n$$

$$G_i(y_1, y_2, y_3, \dots, y_{J_n}) = \partial G / \partial y_i = \mu y_i^{\mu-1}$$

Fonksiyona ilişkin olasılık şu şekilde tanımlanmaktadır ve formülden de görüldüğü gibi çok durumlu logit modele ulaşılmış olur.

$$P_n(i) = \frac{\mu e^{V_{in}(\mu-1)} e^{V_{in}}}{\mu \sum_{j=1}^{J_n} e^{\mu V_{jn}}} = \frac{e^{\mu V_{in}}}{\sum_{j=1}^{J_n} e^{\mu V_{jn}}}$$

4.2.1.5.2 Yuvalanmış Model (Nested Model)

Çoklu logit modelde belirtilen, alternatifler arasındaki bağımsızlık özelliğinin sağlanmaması durumunda yuvalanmış logit model kullanılabilir. Yuvalanmış logit modelin kullanılma gereksinimini ve yararlılığını bir örnekle açıklamaya çalışalım. Çalışan bir kişi evinden işyerine giderken ulaşım aracı olarak kendi otomobilini kullanmayı ya da belediye otobüsüne binmeyi tercih edebilmektedir. Böylece kişinin tercih kümesi otomobil ve belediye otobüsünden oluşmaktadır. Kişinin bu iki alternatifi seçme olasılığını eşit olarak düşünürsek; iki tercihin seçilme olasılıklarının oranı 1 olacaktır. Şimdi de kişinin yeni bir alternatif olarak özel halk otobüslerine binmeyi tercih ettiğini düşünelim. Birey otobüse binmeyi tercih

ettikten sonra belediye otobüsü ile özel halk otobüsü arasında bir fark gözetmeyebilir ve bu durumda da olasılıklar aşağıdaki gibi yazılabilir.⁵¹

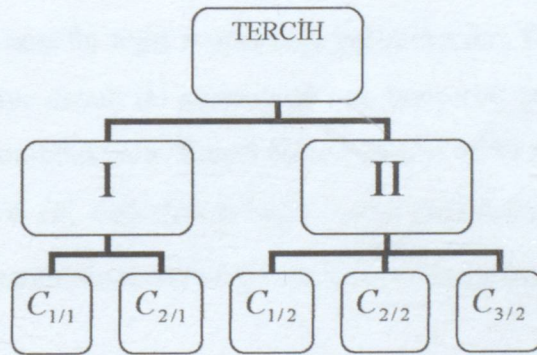
$$P(\text{otomobil}) = 1/2; \quad P(\text{belediyeotobüsü}) = 1/4; \quad P(\text{halkotobüsü}) = 1/4$$

Yeni tercih kümesi ve olasılıklara bakıldığında IIA varsayımının artık gerçekleşmediği görülmektedir. İlk iki tercihin seçilme olasılıklarının oranı artık 1' den farklı bir değerdedir. Bu sorundan kurtulabilmek için alternatiflerin aşamalı olarak düzenlendiği bir yöntem olan yuvalanmış model kullanılmaktadır.

J alternatif L adet alt gruba ayrıldığında seçim seti aşağıdaki gibi yazılabilir:

$$[C_1, C_2, \dots, C_J] = (C_{1/1}, \dots, C_{J_1/1}) \dots (C_{1/L}, \dots, C_{J_L/L})$$

5 alternatifin 2 alt gruba ayrıldığı iki seviyeli yuvalanmış logit model üzerinde çalışmalar yapılacaktır.



Şekil 4.8 İki seviyeli yuvalanmış logit model

Veri setinin alternatiflerin özelliklerine ($X_{j/l}$) ve alternatif setlerinin özelliklerine (Z_l) ilişkin gözlemlerden oluşması halinde yuvalanmış logit modelin matematiksel biçimini elde etmek için l . alt gruptaki j . alternatifin tercih edilmesinin koşulsuz olasılığından hareket edilmektedir.

$$P(\text{twig}_j, \text{branch}_l) = P_{jl} = \frac{e^{\beta'X_{j/l} + \gamma'Z_l}}{\sum_{l=1}^L \sum_{j=1}^{j_l} e^{\beta'X_{j/l} + \gamma'Z_l}} \quad (4.96)$$

⁵¹ Kalkan, E., "Yatay Birleşmelerin İncelenmesinde İktisadi Tekniklerin Kullanılması", s.55

$$P_{jl} = P_{j|l} \cdot P_l = \left(\frac{e^{\beta'X_{j|l}}}{\sum_{j=1}^{Jl} e^{\beta'X_{j|l}}} \right) \left(\frac{e^{\gamma'Z_l}}{\sum_{l=1}^L e^{\gamma'Z_l}} \right) \frac{\left(\sum_{j=1}^{Jl} e^{\beta'X_{j|l}} \right) \left(\sum_{l=1}^L e^{\gamma'Z_l} \right)}{\left(\sum_{l=1}^L \sum_{j=1}^{Jl} e^{\beta'X_{j|l} + \gamma'Z_l} \right)} \quad (4.97)$$

l . alt grup için dahili değer (inclusive value) tanımlanarak olasılıklar yeniden düzenlenebilir.

$$IV_l = \ln \sum_{j=1}^{Jl} e^{\beta'X_{j|l}} \quad (4.98)$$

$$P_{j|l} = \frac{e^{\beta'X_{j|l}}}{\sum_{j=1}^{Jl} e^{\beta'X_{j|l}}} \quad (4.99)$$

$$P_l = \frac{e^{\gamma'Z_l + \tau_l IV_l}}{\sum_{l=1}^L e^{\gamma'Z_l + \tau_l IV_l}} \quad (4.100)$$

$\tau = 1$ olması durumunda koşullu logit model elde edilmektedir. Yuvalanmış modelin katsayı tahminlerinde sınırlı bilgiye dayalı iki aşamalı en çok benzerlik ve tam bilgiye dayalı en çok benzerlik yöntemleri kullanılmaktadır. Sınırlı bilgiye dayalı EYO tahmininde her alt grup için IV değerleri bulunur, z_l ve IV_l değerlerine bağlı olarak parametre tahminleri gerçekleştirilir. Tam bilgiye dayalı EYO tahmininde benzerlik fonksiyonu aşağıdaki gibi tanımlanmaktadır.

$$\ln LL = \sum_{i=1}^n \ln [P_{j|l} \cdot P_l] \quad (4.101)$$

Yuvalanmış logit modeli, Coslet (1978) ulaştırma biçiminin seçiminde, Brownstone (1980) dayanıklı tüketim mallarının satın alınmasında, Goett (1979) enerji talebinde ve Small – Brownstone (1982) trip timing üzerine karar almada kullanmıştır. Small-Brownstone, farklı varış sürelerine karşılık 12 değer alan bağımlı değişkenli model üzerine çalışmışlardır. Bu değerler şöyledir; $B_1=(1,2, \dots, 8)$ ve $B_2=(9,10,11,12)$ veya $B_1=(1,2, \dots, 8)$, $B_2=(9)$ ve $B_3=(10,11,12)$. McFadden (1978), bireylerin yaşamak için bir topluluk ve mesken tipini seçme konusunda araştırma yapmıştır.⁵²

⁵² Akın, F., "Kadınların İş Gücüne Katılımı ve İşteki Durum Tercihinin Nested Logit Model ile Analizi", s.2

4.2.2 Sıralı (Ordered) Tercih Modelleri

Nitel bağımlı değişkenin sıralı olması durumunda sıralı tercih modelleri kullanılmaktadır. Y^* sürekli tesadüfi değişken ve γ tahmin edilecek limit değerleri göstermek üzere;

$$Y = j+1 \text{ eğer } \gamma_j < Y^* \leq \gamma_{j+1}, \quad j = 0,1,2,\dots,J \quad (4.102)$$

$$Y = 0 \text{ eğer } Y^* \leq 0$$

$$Y = 1 \text{ eğer } 0 < Y^* \leq \gamma_1$$

$$Y = 2 \text{ eğer } \gamma_1 < Y^* \leq \gamma_2$$

..

..

..

$$Y = j \text{ eğer } \gamma_{j-1} \leq Y^* \quad (4.103)$$

Bağımlı değişkenin bu değerleri alma olasılıkları probit model ele alınarak hesaplanmaktadır. Sıralı probit modelde kalıntıların normal dağıldığı varsayılmaktadır.

$$P(Y = 0) = \Phi(-\beta'X)$$

$$P(Y = 1) = \Phi(\gamma_1 - \beta'X) - \Phi(-\beta'X)$$

$$P(Y = 2) = \Phi(\gamma_2 - \beta'X) - \Phi(\gamma_1 - \beta'X)$$

..

..

..

$$P(Y = j) = 1 - \Phi(\gamma_{j-1} - \beta'X) \quad (4.104)$$

Tüm olasılıkların pozitif olması için, $0 < \gamma_1 < \gamma_2 < \dots < \gamma_{j-1}$ eşitsizliğinin sağlanması gerekir.⁵³

Sıralı probit model de yine EYO yöntemi ile kolayca tahmin edilebilmektedir. Ancak yoğun hesaplamalar yapılmaksızın bu modeldeki katsayıların nasıl yorumlanması gerektiği pek açık

⁵³ Akın, F., Kalitatif Tercih Modelleri Analizi, s.71

değildir. Dolayısıyla, katsayılar yorumlanırken çok dikkatli olunmalıdır.⁵⁴

Sıralı probit modelin benzerlik fonksiyonu ve logaritmik benzerlik fonksiyonu aşağıdaki gibidir.

$$LL = \prod_{i=1}^N \prod_{j=1}^J [\Phi(\gamma_j - \beta'X_i) - \Phi(\gamma_{j-1} - \beta'X_i)]^{Y_{ij}} \quad (4.105)$$

$$\ln LL = \sum_{i=1}^N \sum_{j=1}^J Y_{ij} [\Phi(\gamma_j - \beta'X_i) - \Phi(\gamma_{j-1} - \beta'X_i)] \quad (4.106)$$

Parametre tahminleri için $\ln LL$ maksimize edilir. Maksimizasyon sürecinde notasyonu kısaltmak için gerekli tanımlamalar aşağıda verilmektedir.

$$Y_{i,j} = \gamma_j - \beta'X_i$$

$$\phi_{i,j} = \phi(\gamma_j - \beta'X_i)$$

$$\delta_{j,k} = \begin{cases} 1, j = k \\ 0, j \neq k \end{cases} \text{ ise}$$

Yukarıdaki tanımlamalardan hareketle, $\ln LL$ 'nin birinci ve ikinci türevleri elde edilmektedir.

$$\frac{\partial \ln LL}{\partial \beta} = \sum_{i=1}^N \sum_{j=1}^J Y_{ij} \frac{\phi_{i,j-1} - \phi_{i,j}}{\Phi_{i,j} - \Phi_{i,j-1}} X_i = 0 \quad (4.107)$$

$$\frac{\partial \ln LL}{\partial \gamma_k} = \sum_{i=1}^N \sum_{j=1}^J Y_{ij} \frac{\delta_{j,k} \phi_{i,j} - \delta_{j-1,k} \phi_{i,j-1}}{\Phi_{i,j} - \Phi_{i,j-1}} = 0 \quad (4.108)$$

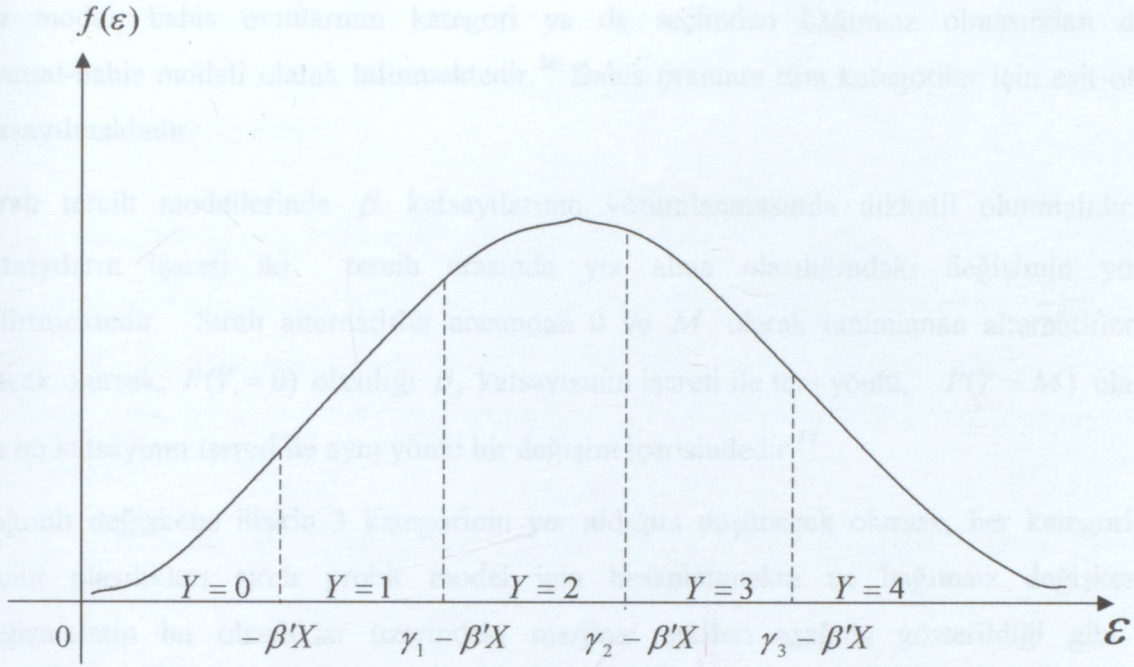
Parametre tahminlerinin elde edilmesi için 2. türevler hesaplanmakta ve Newton-Raphson iterasyon yöntemi kullanılmaktadır.

$$\frac{\partial^2 \ln LL}{\partial \beta \partial \beta'} = \sum_{i=1}^N \sum_{j=1}^J \frac{Y_{ij}}{(\Phi_{i,j} - \Phi_{i,j-1})^2} [(\Phi_{i,j} - \Phi_{i,j-1})(Y_{i,j-1} \phi_{i,j-1} - Y_{i,j} \phi_{i,j}) - (\phi_{i,j-1} - \phi_{i,j})^2] X_i X_i' \quad (4.109)$$

$$\frac{\partial^2 \ln LL}{\partial \beta \partial \gamma_k} = \sum_{i=1}^N \sum_{j=1}^J \frac{Y_{ij}}{(\Phi_{i,j} - \Phi_{i,j-1})^2} \left[\frac{((\Phi_{i,j} - \Phi_{i,j-1})(Y_{i,j} \phi_{i,j} \delta_{j,k} - Y_{i,j-1} \phi_{i,j-1} \delta_{j-1,k})) - ((\phi_{i,j-1} - \phi_{i,j})(\phi_{i,j} \delta_{j,k} - \phi_{i,j-1} \delta_{j-1,k}))}{(\Phi_{i,j} - \Phi_{i,j-1})^2} \right] X_i \quad (4.110)$$

$$\frac{\partial^2 \ln LL}{\partial \gamma_k \partial \gamma_l} = \sum_{i=1}^N \sum_{j=1}^J \frac{Y_{ij}}{(\Phi_{i,j} - \Phi_{i,j-1})^2} \left[\frac{((\Phi_{i,j} - \Phi_{i,j-1})(Y_{i,j-1} \phi_{i,j-1} \delta_{j-1,k} \delta_{j-1,l} - Y_{i,j} \phi_{i,j} \delta_{j,k} \delta_{j,l})) - ((\phi_{i,j} \delta_{j,k} - \phi_{i,j-1} \delta_{j-1,k})(\phi_{i,j} \delta_{j,l} - \phi_{i,j-1} \delta_{j-1,l}))}{(\Phi_{i,j} - \Phi_{i,j-1})^2} \right] \quad (4.111)$$

⁵⁴ Greene, W. H., Econometric Analysis (4th edition), ss. 875-88



Şekil 4.9 Sıralı probit modelde olasılıklar (Ruud, s.760)

Sıralı probit modelin benzerlik fonksiyonunda birikimli normal dağılım ve standart normal olasılık yoğunluk fonksiyonu yerine birikimli lojistik dağılım ve lojistik fonksiyon yerleştirilirse sıralı logit model oluşur.⁵⁵ Sıralı logit modelde alternatiflerin seçilme olasılıkları aşağıdaki gibi hesaplanmaktadır.

$$\begin{aligned}
 P_j &= \Lambda(\beta'X) \\
 P_{j-1} &= \Lambda(\beta'X + \gamma_1) - \Lambda(\beta'X) \\
 P_{j-2} &= \Lambda(\beta'X + \gamma_1 + \gamma_2) - \Lambda(\beta'X + \gamma_1)
 \end{aligned} \tag{4.112}$$

...

...

Sıralı logit model aşağıdaki biçimde de gösterilebilmektedir. Bu gösterimde birey $J+1$ adet alternatifle karşı karşıyadır ve bireyin alternatifleri seçme olasılıkları belirtilmektedir.

$$\begin{aligned}
 P(Y_i = 0) &= \Lambda(\gamma_1 - X_i'\beta) \\
 P(Y_i = 1) &= \Lambda(\gamma_2 - X_i'\beta) - \Lambda(\gamma_1 - X_i'\beta) \\
 &\dots
 \end{aligned} \tag{4.113}$$

...

$$P(Y_i = J) = 1 - \Lambda(\gamma_J - X_i'\beta)$$

⁵⁵ Akın, F., a.g.e, s.76

Bu model, bahis oranlarının kategori ya da seçimden bağımsız olmasından dolayı oransal-bahis modeli olarak bilinmektedir.⁵⁶ Bahis oranının tüm kategoriler için eşit olduğu varsayılmaktadır.

Sıralı tercih modellerinde β katsayılarının yorumlanmasında dikkatli olunmalıdır. Bu katsayıların işareti iki tercih arasında yer alma olasılığındaki değişimin yönünü belirtmektedir. Sıralı alternatifler arasından 0 ve M olarak tanımlanan alternatifleri ele alacak olursak; $P(Y = 0)$ olasılığı β_j katsayısının işareti ile ters yönlü, $P(Y = M)$ olasılığı ise bu katsayının işareti ile aynı yönlü bir değişim içerisindedir.⁵⁷

Bağımlı değişkene ilişkin 3 kategorinin yer aldığını düşünecek olursak, her kategori için seçim olasılıkları sıralı probit model için hesaplanmakta ve bağımsız değişkendeki değişmelerin bu olasılıklar üzerindeki marjinal etkileri aşağıda gösterildiği gibi elde edilmektedir.⁵⁸

$$\begin{aligned}\frac{\partial P(Y = 0)}{\partial X} &= -\phi(\beta'X)\beta, \\ \frac{\partial P(Y = 1)}{\partial X} &= [\phi(-\beta'X) - \phi(\gamma_1 - \beta'X)]\beta, \\ \frac{\partial P(Y = 2)}{\partial X} &= \phi(\gamma_2 - \beta'X)\beta\end{aligned}\tag{4.114}$$

Ekonomik olayları gözlenemeyen tek bir indeks değişkeni ile açıklamanın güç ve karmaşık olması nedeniyle olacak ki, sıralı modeller ekonometrik uygulamalardan ziyade biyometrik uygulamalarda daha sık kullanılmaktadır. Bu tür modellerin kullanımında dikkatli olunmalıdır. Öyle ki, sıralı olmayan gerçek model yerine sıralı bir modelin kullanılması olasılıkların tahmininde ciddi sapmalara yol açabilecek, öte yandan sıralı olan gerçek model yerine sıralı olmayan bir modelin kullanılması da tutarlılıktan ziyade etkinlik kaybına neden olacaktır.

Ayrıca, incelenen modeller arasında söz konusu olan temel bazı farklılıklar da burada verilebilir. Sıralı logit modeli sıralı probit modelden ve çok durumlu logit modeli yuvalanmış logit modelden ayıran temel fark hata terimlerinin dağılımıdır. Öyle ki, hata terimleri sıralı logit modelde lojistik dağılım sergilerken, sıralı probit modelde normal dağılmaktadır. Diğer yandan, çok durumlu logit model tesadüfi fayda fonksiyonundaki hata terimlerinin bağımsız

⁵⁶ Peracchi, F., *Econometrics*, s.552

⁵⁷ E-Views 4.1 User's Guide, s.425

⁵⁸ Greene, W. H., *Econometric Analysis* (3rd Edition), s. 928

uç-değer dağılımına sahip olduğu varsayımına dayalı olarak oluşturulurken, yuvalanmış logit model hata terimlerinin genelleştirilmiş uç-değer dağılımına sahip olduğu varsayımına dayalı olarak oluşturulabilmektedir.

4.2.3 Ardışık (Sequential) Tercih Modelleri

Çok alternatifin söz konusu olduğu ve alternatifler arasında tercih yapılırken kararların ardışık olarak alındığı durumlarda çok durumlu modellerin tahmini daha az sayıda alternatifli modellerin peş peşe tahminine dönüşmekte ve bu durum yapılması gereken hesaplamaları azaltmaktadır.⁵⁹

Çoklu tercih modellerinin giriş kısmında tanımlanan ardışık değişkene ilişkin olasılıklar hesaplanmaktadır.

$$\begin{aligned} P_1 &= F(\beta'_1 X) \\ P_2 &= [1 - F(\beta'_1 X)]F(\beta'_2 X) \\ P_3 &= [1 - F(\beta'_1 X)][1 - F(\beta'_2 X)]F(\beta'_3 X) \end{aligned} \quad (4.115)$$

β_1 katsayısı, örneklemdaki bireylerin ilkökul mezunu olup olmamalarına göre ikiye ayrılarak tahmin edilebilir. β_2 katsayısı ise ilkökul mezunu olan bireylerin oluşturduğu alt grupta bireylerin ortaokul mezunu olup olmamalarına göre ikiye ayrılması yoluyla tahmin edilebilir. Her bir durumda oluşturulan ikili modeller, probit veya logit yöntemi ile tahmin edilebilir.⁶⁰

Üç alternatifin olduğu ve bireyin öncelikle $Y = 2$ veya $Y \neq 2$ olup olmadığını belirlediğini, eğer $Y \neq 2$ ise bu defa da $Y = 1$ veya $Y = 0$ olup olmadığını belirlediğini ve her bir tercihin seçilme olasılıklarının belirlenmesinde probit modelinin kullanıldığını düşünürsek ardışık bir probit modele ilişkin olasılıklar aşağıdaki gibi gösterilmektedir. Bu modelin olabilirlik fonksiyonunun maksimizasyonu bu iki adet iki durumlu probit modelinin olabilirlik fonksiyonunun maksimum kılınmasıyla gerçekleştirilebilir.⁶¹

$$\begin{aligned} P_2 &= \Phi(\beta'_2 X_2) \\ P_1 &= [1 - \Phi(\beta'_2 X_2)]\Phi(\beta'_1 X_1) \end{aligned} \quad (4.116)$$

⁵⁹ Amemiya, T., Advanced Econometrics, s.310

⁶⁰ Maddala, G. S., a.g.e, s.50

⁶¹ Amemiya, T., a.g.e, s.310

4.2.4 Tercih Modellerine İlişkin Genel İfadeler

Aşağıda yer alan tanımlamalar doğrultusunda bağımlı gölge değişken olan Y_i 'nin genel gösterimi ve beklenen değeri belirtilmektedir.

$$I_i = \beta_1 + \beta_2 X_{ik} + \dots + \beta_k X_{ik} \quad (4.117)$$

$$\beta = (\beta_1 \quad \beta_2 \quad \dots \quad \beta_k)' \quad (4.118)$$

$$X_i' = (1 \quad X_{i2} \quad \dots \quad X_{ik}) \quad (4.119)$$



$$I_i = X_i' \beta \quad (4.120)$$

$$Y_i = \beta_1 + \beta_2 X_{ik} + \dots + \beta_k X_{ik} + \varepsilon_i = X_i' \beta + \varepsilon_i \quad (4.121)$$

$$E(Y_i / X_i) = 1.P(Y_i = 1 / X_i) + 0.P(Y_i = 0 / X_i) = P(Y_i = 1 / X_i) = X_i' \beta \quad (4.122)$$

Bağımlı değişkenin aldığı değere bağlı olarak, ε_i 'nin beklenen değeri ile varyansı hesaplanmaktadır.

$$P(Y_i = 0 / X_i) = P(\varepsilon_i = -X_i' \beta / X_i) = 1 - X_i' \beta \quad (4.123)$$

$$P(Y_i = 1 / X_i) = P(\varepsilon_i = 1 - X_i' \beta / X_i) = X_i' \beta \quad (4.124)$$

$$E(\varepsilon_i / X_i) = 0 \quad (4.125)$$

$$V(\varepsilon_i / X_i) = X_i' \beta (1 - X_i' \beta) \quad (4.126)$$

Alternatiflere ilişkin seçim olasılıklarının genel gösterimi ile doğrusal olasılık, probit ve logit model için olasılık hesaplanması sırasıyla ifade edilmektedir.

$$P_i = F(I_i) = F(\beta_1 + \beta_2 X_{ik} + \dots + \beta_k X_{ik}) = F(X_i' \beta) \quad (4.127)$$

$$F(X_i' \beta) = X_i' \beta \quad (4.128)$$

$$F(X_i' \beta) = \Phi(X_i' \beta) = \int_{-\infty}^{X_i' \beta} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z^2\right) dz \quad (4.129)$$

$$F(X_i' \beta) = \frac{\exp(X_i' \beta)}{1 + \exp(X_i' \beta)} \quad (4.130)$$

Bağımsız değişkendeki değişimin olasılık üzerindeki marjinal etkisinin genel gösterimi ile doğrusal olasılık, probit ve logit model için hesaplanması sırasıyla gösterilmektedir.⁶²

$$\frac{\partial P_i}{\partial X_{ik}} = \frac{\partial F(X'_i \beta)}{\partial X_{ik}} = F'(X'_i \beta) \cdot \frac{\partial (X'_i \beta)}{\partial X_{ik}} = f(X'_i \beta) \cdot \beta_k \quad (4.131)$$

$$\frac{\partial P_i}{\partial X_{ik}} = \beta_k \quad (4.132)$$

$$\frac{\partial P_i}{\partial X_{ik}} = \phi(X'_i \beta) \cdot \beta_k \quad (4.133)$$

$$\frac{\partial P_i}{\partial X_{ik}} = \frac{\exp(X'_i \beta)}{(1 + \exp(X'_i \beta))^2} \cdot \beta_k \quad (4.134)$$

N adet gözlemden n_1 adetinin $Y_i = 1$, n_2 adetinin ise $Y_i = 0$ olduğu ve $N = n_1 + n_2$ olduğu varsayılarak benzerlik fonksiyonu aşağıdaki gibi yazılmaktadır.

$$\begin{aligned} LL &= P(Y_1, Y_2, \dots, Y_N) = P(Y_1) \cdot P(Y_2) \cdot \dots \cdot P(Y_N) \\ &= P_1 \cdot P_2 \cdot \dots \cdot P_{n_1} \cdot (1 - P_{n_1+1}) \cdot (1 - P_{n_1+2}) \cdot \dots \cdot (1 - P_N) \\ &= \prod_{i=1}^{n_1} P_i \prod_{i=n_1+1}^N (1 - P_i) = \prod_{i=1}^N P_i^{Y_i} \cdot (1 - P_i)^{1-Y_i} \\ &= \prod_{i=1}^N F(X'_i \beta)^{Y_i} \cdot (1 - F(X'_i \beta))^{1-Y_i} \end{aligned} \quad (4.135)$$

Benzerlik fonksiyonunun logaritması alınır ve parametrelere göre kısmî türevleri 0' a eşitlenirse en çok benzerlik tahmin edicileri elde edilmiş olur.

$$\frac{\partial \log LL}{\partial \beta} = \sum \frac{Y_i - F(X'_i \beta)}{F(X'_i \beta)(1 - F(X'_i \beta))} f(X'_i \beta) \cdot X_i = 0 \quad (4.136)$$

Logit model için işlemler yapıldığında tahmin edilen olasılık değerleri toplamının, $Y_i = 1$ olan gözlemlerin toplam sayısını veren $\sum Y_i$ değerine eşit olduğu görülmektedir.⁶³

$$\frac{\partial \log LL}{\partial \beta} = \sum_{i=1}^N \left[Y_i - \frac{\exp(X'_i \beta)}{1 + \exp(X'_i \beta)} \right] X_i = 0 \quad (4.137)$$

⁶² Verbeek, M., a.g.e, s.179

⁶³ Verbeek, M., a.g.e, s.181

$$\sum_{i=1}^N Y_i X_i = \sum_{i=1}^N P_i X_i \quad (4.138)$$

Çoklu tercih modellerinde bireyler 2' den fazla tercih arasından seçim yapmaktadır ve bu seçime ilişkin olasılık değeri hesaplanmaktadır. Örneğin piyasa araştırmalarında tüketicilere bir ürün hakkında beğeni derecesinin sorulduğu, tüketicinin çok kötü, kötü, ne kötü ne iyi, iyi, çok iyi tercihlerinden birini seçtiği ve bu seçimin bağımlı değişken olarak değerlendirildiği bir model çoklu tercih modeli olarak ele alınabilir.⁶⁴

$$Y_i^* = X_i' \beta + \varepsilon_i$$

$$Y_i = j, \text{ eğer } \gamma_{j-1} < Y_i^* \leq \gamma_j \quad (4.139)$$

$$\begin{aligned} P(Y_i = j | \beta, \gamma) &= P(\gamma_{j-1} < Y_i^* \leq \gamma_j) \\ &= P(\gamma_{j-1} < X_i' \beta + \varepsilon_i \leq \gamma_j) \\ &= P(\gamma_{j-1} - X_i' \beta < \varepsilon_i \leq \gamma_j - X_i' \beta) \\ &= \Phi(\gamma_j - X_i' \beta) - \Phi(\gamma_{j-1} - X_i' \beta) \end{aligned} \quad (4.140)$$

Yukarıda yer alan (4.140) numaralı formülde bireyin j . alternatifi seçme olasılığının probit model yaklaşımıyla hesaplanması gösterilmektedir.

$$Y_i = \begin{cases} 0 & Y_i^* \leq \gamma_1 \\ 1 & \gamma_1 < Y_i^* \leq \gamma_2 \\ 2 & \gamma_2 < Y_i^* \leq \gamma_3 \\ \vdots & \vdots \\ J & \gamma_J < Y_i^* \end{cases} \quad \text{eğer}$$

$$\begin{aligned} P(Y_i = 0) &= F(\gamma_1 - X_i' \beta) \\ P(Y_i = 1) &= F(\gamma_2 - X_i' \beta) - F(\gamma_1 - X_i' \beta) \\ &\dots \\ P(Y_i = J) &= 1 - F(\gamma_J - X_i' \beta) \end{aligned} \quad (4.141)$$

(4.141)' de, $J+1$ adet alternatifle karşılaşan bireyin alternatifleri seçme olasılıklarının genel ifadesi görülmektedir. Sıralı tercih modellerinde bağımsız değişken sayısı kadar parametre değeri ve alternatif sayısının bir eksiği kadar da limit değer hesaplanmaktadır. Hesaplanan bu değerler alternatiflere ilişkin seçim olasılıklarının bulunmasında kullanılmaktadır.

⁶⁴ Koop, G., Bayesian Econometrics, s.218

5. EKONOMİK TÜKETİCİ TEORİSİYLE MODEL KURMA

5.1 Ekonomik Tüketici Teorisi

Bireylerin tercihlerinin açıklanmasında yararlanılan yaklaşımlardan olan ekonomik tüketici teorisi açıklanacaktır. Bu bölümün teorik açıklamalarının çoğunda Lerman, R.S. ve Akiva B.M.'nin kitabından yararlanılmıştır. Bireylerin tercih edecekleri mal ve hizmetler $1'$ den T' ye kadar belirlenmiş olsun ve bu mal ve hizmetlere ilişkin miktarlar aşağıdaki gibi belirtilsin.

$$Q = \{q_1, q_2, \dots, q_T\}$$

Mal ve hizmetlerin miktarlarının pozitif ve sürekli değerler olduğu farzedilmektedir. Belirli bir gelir düzeyi olan Inc ve belirli fiyatlar olan $prc_1, prc_2, \dots, prc_T$ dikkate alınırsa bireyin bütçe kısıtlaması şu şekilde yazılabilmektedir.

$$\sum_{t=1}^T prc_t q_t \leq Inc \quad (5.1)$$

Bireylerin mümkün olan tüm alternatifleri karşılaştırabildikleri farzedilmektedir. Karşılaştırma esnasında bireyler bütçe kısıtlamasına uymakla beraber kendileri için en çok fayda sağlayacak olan alternatifini seçmektedirler. Mümkün olan tüm alternatifler, seçenekler diğer bir ifadeyle alternatiflere ilişkin evrensel küme C ile simgelensin. Bireyler tercihlerini bu kümeden yapacaklarından kişilerin seçim sepetleri için $C_n \subseteq C$ ifadesi yazılabilmektedir. Fayda fonksiyonu ve bireyin i ve j alternatifleri arasından i' yi seçme koşulu şu şekilde yazılmaktadır.

$$U = U(q_1, q_2, \dots, q_L) \longrightarrow \text{fayda fonksiyonu} \quad (5.2)$$

$$U_{in} > U_{jn}, \quad j \neq i, j \in C_n, i \in C_n$$

Fayda fonksiyonu tercihin kendisi ile ilgili olan ve birey tarafından değerlendirilen özelliklerle ilgili vektör (kişinin yolculuk tercihinde seçtiği alternatif ile ilişkin yolculuk süresi, yolculuk giderleri, yolculuğun rahatlık ve güvenlik düzeyi gibi değişkenler) ile bireyin karakteristik özellikleriyle (gelir, yaş, cinsiyet, eğitim durumu gibi değişkenler) ilgili vektör olmak üzere 2 kısımdan oluşmaktadır.

$$U_{in} = U(z_{in}, S_n) \quad (5.3)$$

5.2 Rasgele Fayda (Random Utility)

Bireyin en yüksek fayda sağlayacak olan alternatifini seçeceği farzedilmektedir. Faydalar kesin olarak bilinmemekte ve bu yüzden rasgele değişkenler olarak ele alınmaktadır. i . alternatifin seçilme olasılığı, bu alternatifin seçilmesiyle elde edilecek fayda olan U_{in} ' in seçim sepetindeki diğer tüm alternatiflerin seçilmeleriyle elde edilecek olan faydaya eşit veya bu faydadan büyük olma olasılığına eşittir.⁶⁵

$$P(i/C_n) = P(U_{in} \geq U_{jn}, j \in C_n) \quad (5.4)$$

Bir alternatif ile ilişkin rasgele fayda gözlenen veya sistematik bileşen ile gözlenemeyen bileşen olmak üzere 2 kısımda incelenmektedir. V_{in} ve V_{jn} , i . ve j . alternatiflere ilişkin faydanın sistematik kısmını, ε_{in} ve ε_{jn} ise gözlenemeyen yani rasgele kısımları belirtmektedir. ε_{in} ve ε_{jn} terimleri disturbans olarak da adlandırılmaktadır.

$$U_{in} = V(z_{in}, S_n) + \varepsilon(z_{in}, S_n) = V_{in} + \varepsilon_{in} \quad (5.5)$$

$$P(i/C_n) = P(V_{in} + \varepsilon_{in} \geq V_{jn} + \varepsilon_{jn}, j \in C_n) \quad (5.6)$$

Açıklamalar doğrultusunda n . bireyin i . alternatifi seçme olasılığı aşağıdaki gibi yazılmaktadır.

$$P_n(i) = P(U_{in} \geq U_{jn}) = P(V_{in} + \varepsilon_{in} \geq V_{jn} + \varepsilon_{jn}) = P(\varepsilon_{jn} - \varepsilon_{in} \leq V_{in} - V_{jn}) \quad (5.7)$$

Alternatiflerin sistematik kısımlarının genel ifadeleri aşağıdaki gibi olmaktadır. h vektörel fonksiyonu belirtmektedir.

$$X_{in} = h(z_{in}, S_n) \rightarrow V_{in} = V(X_{in}) \quad (5.8)$$

$$X_{jn} = h(z_{jn}, S_n) \rightarrow V_{jn} = V(X_{jn}) \quad (5.9)$$

$$V_{in} = \beta_1 X_{in1} + \beta_2 X_{in2} + \beta_3 X_{in3} + \dots + \beta_K X_{inK} \quad (5.10)$$

$$V_{jn} = \beta_1 X_{jn1} + \beta_2 X_{jn2} + \beta_3 X_{jn3} + \dots + \beta_K X_{jnK} \quad (5.11)$$

Alternatiflerin gözlenemeyen kısımları olan ε_{in} ve ε_{jn} ' lerin dağılımları ya da bu kısımların farklarının dağılımları üzerine yapılan varsayımlar bizi farklı seçim modellerine yöneltmektedir.

⁶⁵ Kennedy, P., A Guide to Econometrics, s.231

5.3 İkili Seçim Modelleri

Disturbans terimlerinin ya da bu terimler arasındaki farkların dağılımı hakkında bazı varsayımlar yaparak farklı modeller kurulmakta ve bu modeller yardımıyla da i . alternatifin seçilme olasılığı hesaplanmaktadır.

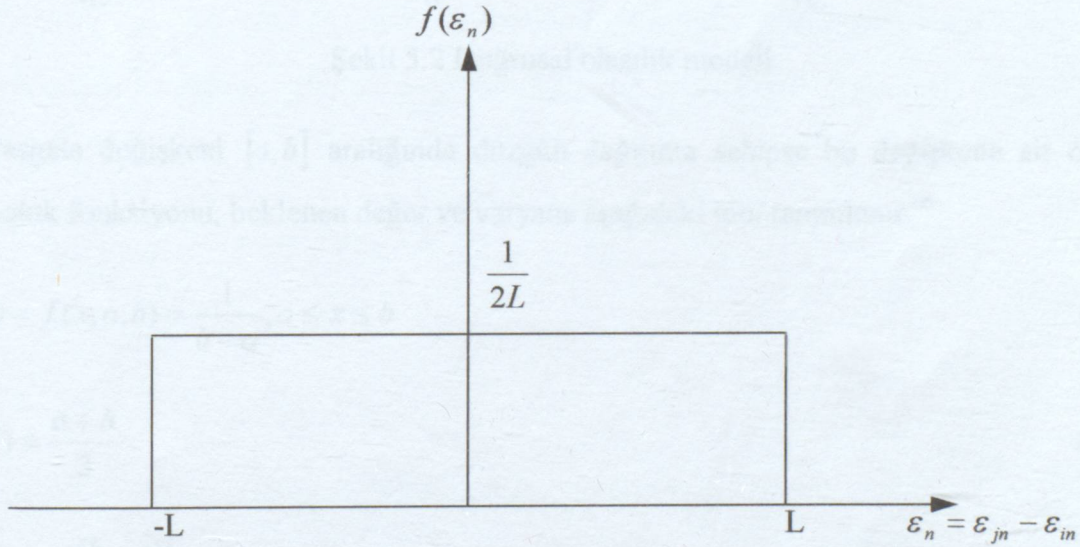
5.3.1 İkili Seçim Model Çeşitleri

4 başlık altında incelemeler yapılacaktır.

5.3.1.1 İkili Doğrusal Olasılık Modeli

i . ve j . alternatiflere ilişkin disturbansların farklarının, $\varepsilon_{jn} - \varepsilon_{in}$, $-L$ ile L arasında düzgün (uniform) dağılıma sahip olduğunu varsayalım. $\varepsilon_{jn} - \varepsilon_{in} = \varepsilon_n$ ve bu terime ilişkin yoğunluk fonksiyonunu $f(\varepsilon_n)$ olarak tanımlayalım. i . alternatifin seçilme olasılığı aşağıdaki gibi hesaplanmaktadır.

$$P_n(i) = P(V_{in} + \varepsilon_{in} \geq V_{jn} + \varepsilon_{jn}) = P(\varepsilon_{jn} - \varepsilon_{in} \leq V_{in} - V_{jn}) = P(\varepsilon_n \leq V_{in} - V_{jn})$$

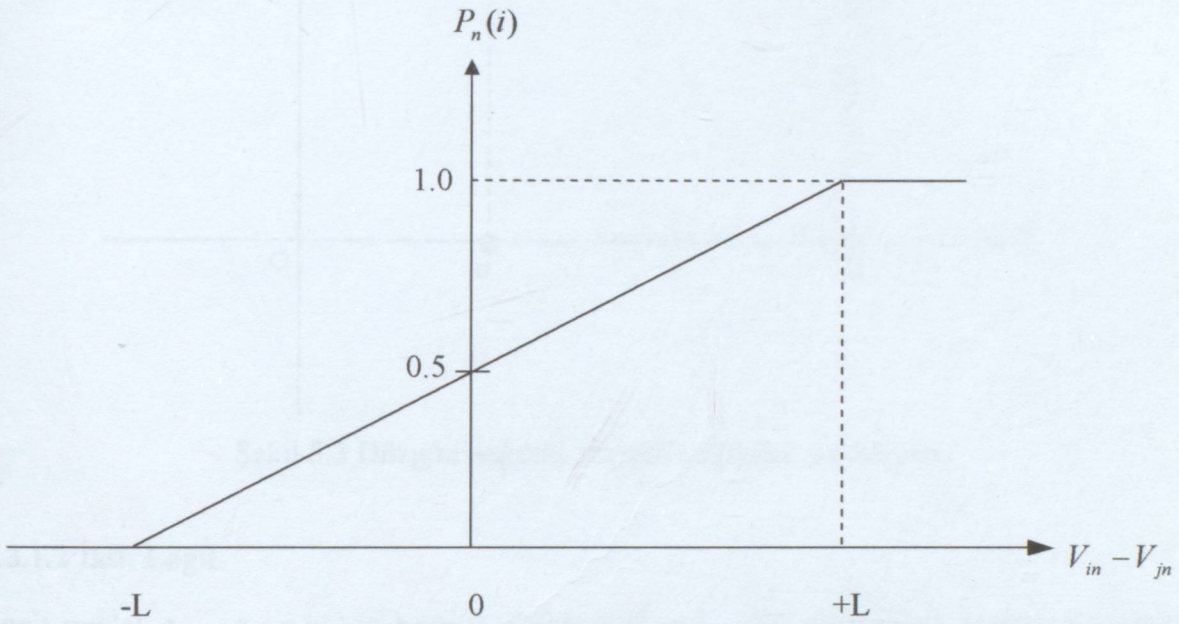


Şekil 5.1 Düzgün dağılım

$$P_n(i) = \begin{cases} 0 & V_{in} - V_{jn} < -L \\ \int_{-L}^{V_{in} - V_{jn}} f(\varepsilon_n) d\varepsilon_n = \frac{V_{in} - V_{jn} + L}{2L} & -L \leq V_{in} - V_{jn} \leq L \\ 1 & V_{in} - V_{jn} > L \end{cases} \quad (5.12)$$

Sistematik kısımların parametrelerinde doğrusal fonksiyonlar olduğunu varsayarsak $V_{in} - V_{jn}$ farklarının $P_n(i)$ ile olan doğrusal ilişkisini aşağıdaki şekilde görebiliriz.

$$V_{in} - V_{jn} = \beta'X_{in} - \beta'X_{jn} = \beta'(X_{in} - X_{jn}) \quad (5.13)$$



Şekil 5.2 Doğrusal olasılık modeli

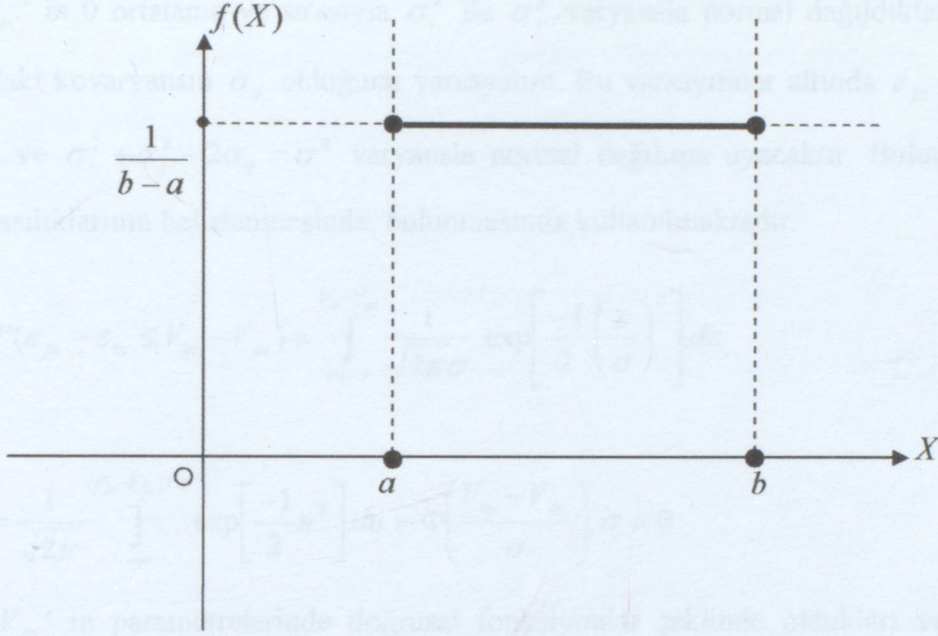
X rasgele değişkeni $[a, b]$ aralığında düzgün dağılıma sahipse bu değişkene ait olasılık yoğunluk fonksiyonu, beklenen değer ve varyans aşağıdaki gibi tanımlanır.⁶⁶

$$f(x) = f(x; a; b) = \frac{1}{b-a}, a \leq x \leq b \quad (5.14)$$

$$E(X) = \frac{a+b}{2} \quad (5.15)$$

$$Var(X) = \frac{(b-a)^2}{12} \quad (5.16)$$

⁶⁶ Akdeniz, F., Olasılık ve İstatistik, s.249



Şekil 5.3 Düzgün dağılım olasılık yoğunluk fonksiyonu

5.3.1.2 İkili Logit

Logit model $\varepsilon_{jn} - \varepsilon_{in} = \varepsilon_n$ ' in logistik dağılıma uygun olduğu varsayımı üzerine kuruludur. Lojistik dağılıma ilişkin birikimli dağılım ve olasılık yoğunluk fonksiyonları aşağıdaki gibi belirtilmektedir.

$$F(\varepsilon_n) = \frac{1}{1 + e^{-\mu\varepsilon_n}}, \mu > 0, -\infty < \varepsilon_n < +\infty \quad (5.17)$$

$$f(\varepsilon_n) = \frac{\mu e^{-\mu\varepsilon_n}}{(1 + e^{-\mu\varepsilon_n})^2} \quad (5.18)$$

i. alternatifin seçilme olasılığı $P_n(i)$ aşağıdaki gibi belirtilmektedir.

$$P_n(i) = P(U_{in} \geq U_{jn}) = \frac{1}{1 + e^{-\mu(V_{in} - V_{jn})}} = \frac{e^{\mu V_{in}}}{e^{\mu V_{in}} + e^{\mu V_{jn}}} \quad (5.19)$$

V_{in} ve V_{jn} ' in parametrelerinde doğrusal fonksiyonlar şeklinde oldukları varsayılırsa;

$$P_n(i) = \frac{e^{\mu\beta'X_{in}}}{e^{\mu\beta'X_{in}} + e^{\mu\beta'X_{jn}}} = \frac{1}{1 + e^{-\mu\beta'(X_{in} - X_{jn})}} \quad (5.20)$$

olarak elde edilir.

5.3.1.3 İkili Probit

ε_{in} ve ε_{jn} 'in 0 ortalama ve sırasıyla σ_i^2 ile σ_j^2 varyansla normal dağıldıklarını ve ayrıca aralarındaki kovaryansın σ_{ij} olduğunu varsayalım. Bu varsayımlar altında $\varepsilon_{jn} - \varepsilon_{in}$ terimi 0 ortalama ve $\sigma_i^2 + \sigma_j^2 - 2\sigma_{ij} = \sigma^2$ varyansla normal dağılıma uyacaktır. Bulunan bu sonuç seçim olasılıklarının belirlenmesinde, bulunmasında kullanılmaktadır.

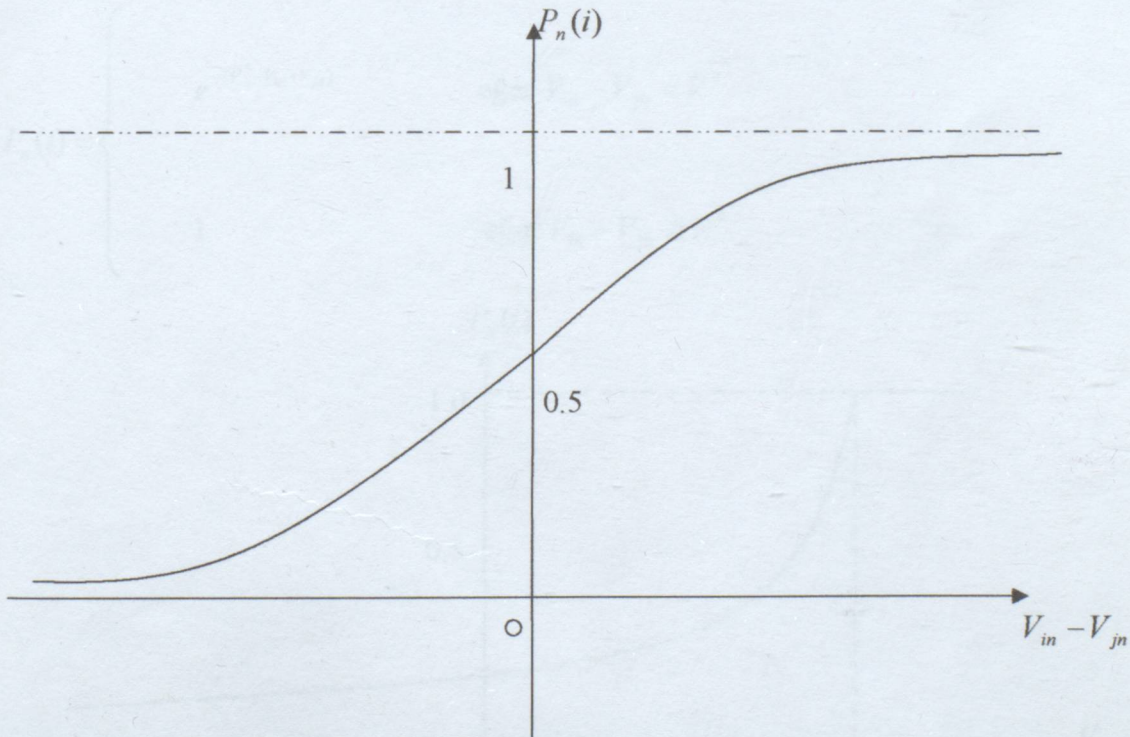
$$P_n(i) = P(\varepsilon_{jn} - \varepsilon_{in} \leq V_{in} - V_{jn}) = \int_{\varepsilon=-\infty}^{V_{in}-V_{jn}} \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left[\frac{-1}{2} \cdot \left(\frac{\varepsilon}{\sigma}\right)^2\right] \cdot d\varepsilon$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(V_{in}-V_{jn})/\sigma} \exp\left[\frac{-1}{2} u^2\right] \cdot du = \Phi\left(\frac{V_{in}-V_{jn}}{\sigma}\right), \sigma > 0 \quad (5.21)$$

V_{in} ve V_{jn} 'in parametrelerinde doğrusal fonksiyonlar şeklinde oldukları varsayılarak i . alternatifte ait seçim olasılığı şu şekilde bulunmaktadır.

$$P_n(i) = \Phi\left(\frac{\beta'(X_{in} - X_{jn})}{\sigma}\right) \quad (5.22)$$

$\sigma = 1$ için ikili probit modeline ilişkin grafik aşağıda çizilmiştir.



Şekil 5.4 İkili probit model

İkili probit modeli çeşitli alanlarda kullanılmaktadır. Finney (1971) tarafından bildirildiğine göre, en yaygın kullanım alanı zehirli maddeleri ilgilendiren toksikolojidir. Toksikolojide $V_{in} - V_{jn}$ ' in deneğe verilen zehirli maddenin dozajını ifade ettiği düşünülürse, $P_n(i)$ deneğin hastalığa yakalanma veya ölme olasılığı olarak yorumlanır.

5.3.1.4 Diğer İkili Seçim Modelleri

$\varepsilon_{jn} - \varepsilon_{in} = \varepsilon_n$ ' in dağılımına ilişkin yapılacak varsayımlar değişik bir seçim modeli oluşturmamıza yol açmaktadır. Aşağıdaki 3 modelde i . alternatifine ilişkin seçim olasılıkları verilmektedir.

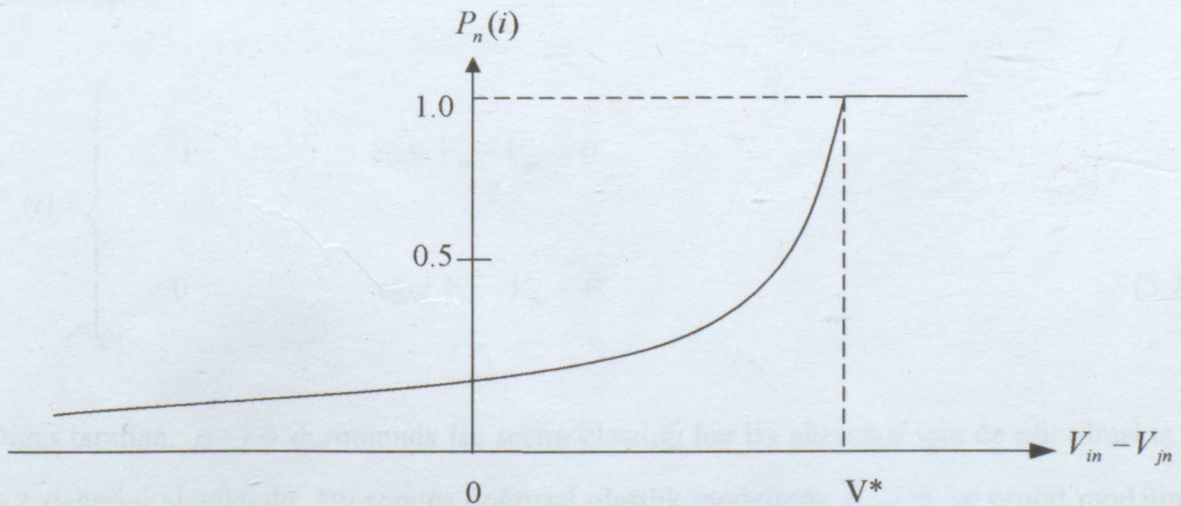
5.3.1.4.1 Arctan Olasılık Modeli

Probit ve logit modelin çizimlerine benzemektedir. Modelin varsayımına göre ε_{in} ve ε_{jn} Cauchy dağılımına uymaktadır.

$$P_n(i) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1}(V_{in} - V_{jn}) \quad (5.23)$$

5.3.1.4.2 Sağdan Kesikli Üstel Model

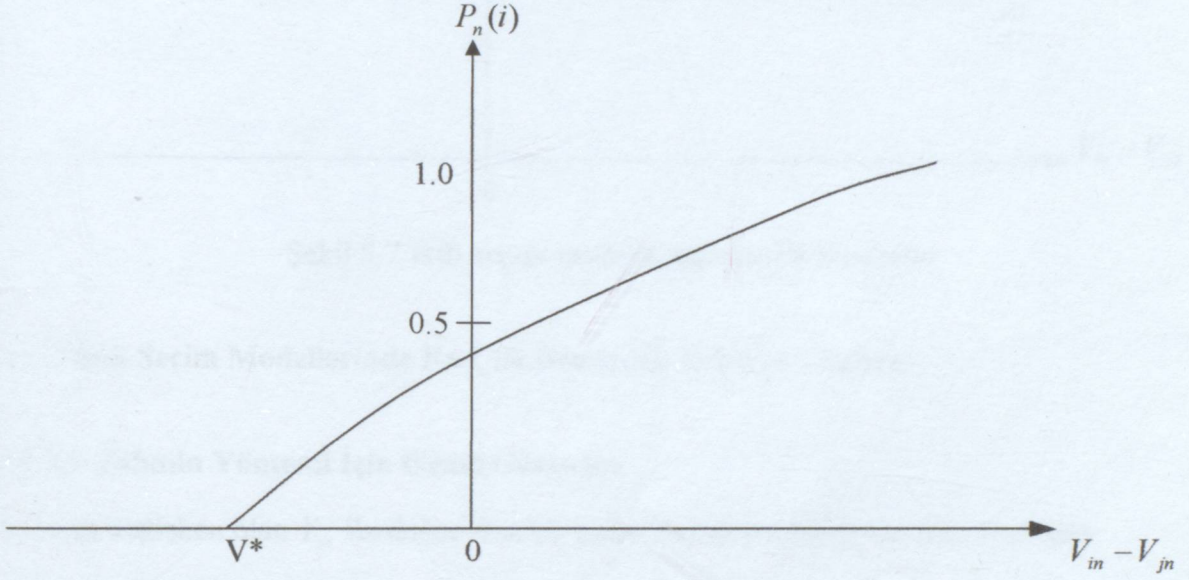
$$P_n(i) = \begin{cases} e^{-(V^* - V_{in} + V_{jn})} & \text{eğer } V_{in} - V_{jn} < V^* \\ 1 & \text{eğer } V_{in} - V_{jn} \geq V^* \end{cases} \quad (5.24)$$



Şekil 5.5 Sağdan kesikli üstel model

5.3.1.4.3 Soldan Kesikli Üstel Model

$$P_n(i) = \begin{cases} 1 - e^{-(V_{in} - V_{jn} - V^*)} & \text{eğer } V_{in} - V_{jn} > V^* \\ 0 & \text{eğer } V_{in} - V_{jn} \leq V^* \end{cases} \quad (5.25)$$



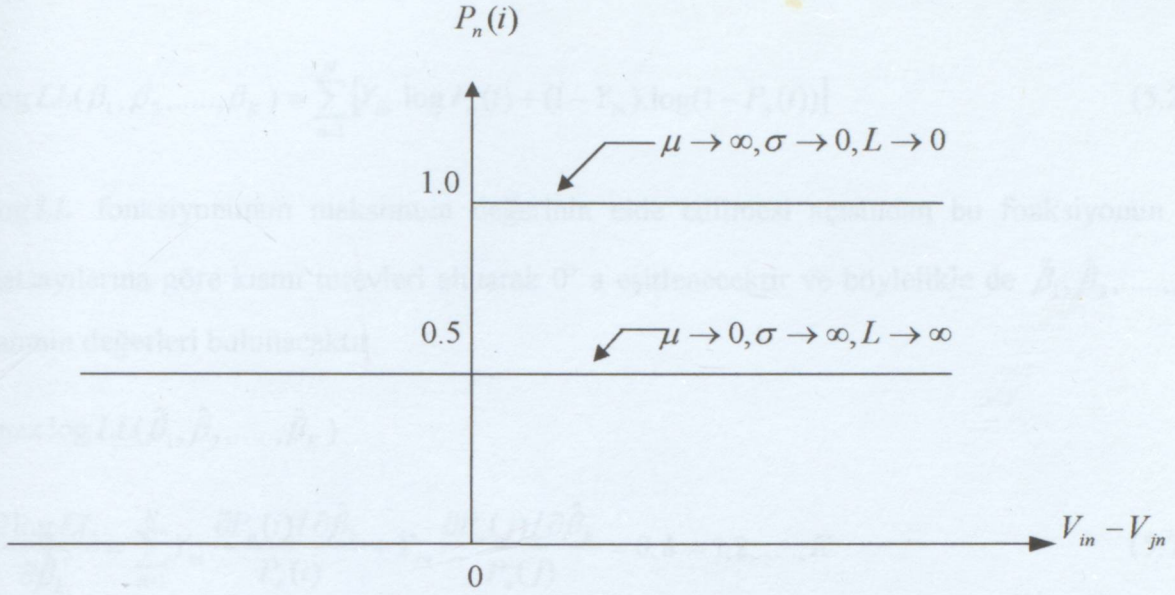
Şekil 5.6 Soldan kesikli üstel model

5.3.2 İkili Seçim Modellerinde Sınırlı Durumlar (Limiting Cases)

Logit modelinde $\mu \rightarrow \infty$ olması durumunda i . alternatife ilişkin seçim olasılığı aşağıda belirtilmiştir.

$$P_n(i) = \begin{cases} 1 & \text{eğer } V_{in} - V_{jn} > 0 \\ 0 & \text{eğer } V_{in} - V_{jn} < 0 \end{cases} \quad (5.26)$$

Diğer taraftan, $\mu \rightarrow 0$ durumunda ise seçim olasılığı her iki alternatif için de eşit olmakta ve $1/2$ değerini almaktadır. Bu sonuca doğrusal olasılık modelinde $L \rightarrow \infty$ ve probit modelinde ise $\sigma \rightarrow \infty$ olması durumlarında ulaşılmaktadır.



Şekil 5.7 İkili seçim modellerinde sınırlı durumlar

5.3.3 İkili Seçim Modellerinde En Çok Benzerlik Tahmin Yöntemi

5.3.3.1 Tahmin Yöntemi İçin Genel Gösterim

Bağımlı değişken olan Y_{in} iki değer alan bir gölge değişken olarak tanımlanmaktadır.

$$Y_{in} = \begin{cases} 1 & \text{eğer } n. \text{ birey } i. \text{ alternatifi seçerse} \\ 0 & \text{eğer } n. \text{ birey } j. \text{ alternatifi seçerse} \end{cases}$$

N birimlik bir örneğin anakütleden rasgele seçildiğini varsayarsak, örneğe ilişkin benzerlik fonksiyonu LL ve benzerlik fonksiyonunun logaritması olan $\log LL$ aşağıdaki gibi ifade edilir.

$$LL(\beta_1, \beta_2, \dots, \beta_K) = \prod_{n=1}^N P_n(i)^{Y_{in}} \cdot P_n(j)^{Y_{jn}} \quad (5.27)$$

$$\log LL(\beta_1, \beta_2, \dots, \beta_K) = \sum_{n=1}^N Y_{in} \cdot \log P_n(i) + Y_{jn} \cdot \log P_n(j) \quad (5.28)$$

$Y_{jn} = 1 - Y_{in}$ ve $P_n(j) = 1 - P_n(i)$ olduğu dikkate alınır;

$$\log LL(\beta_1, \beta_2, \dots, \beta_K) = \sum_{n=1}^N [Y_{in} \cdot \log P_n(i) + (1 - Y_{in}) \cdot \log(1 - P_n(i))] \quad (5.29)$$

$\log LL$ fonksiyonunun maksimum değerinin elde edilmesi açısından bu fonksiyonun β katsayılarına göre kısmî türevleri alınarak 0' a eşitlenecektir ve böylelikle de $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_K$ tahmin değerleri bulunacaktır.

$$\max \log LL(\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_K)$$

$$\frac{\partial \log LL}{\partial \hat{\beta}_k} = \sum_{n=1}^N Y_{in} \frac{\partial P_n(i) / \partial \hat{\beta}_k}{P_n(i)} + Y_{jn} \frac{\partial P_n(j) / \partial \hat{\beta}_k}{P_n(j)} = 0, k = 1, 2, \dots, K \quad (5.30)$$

Benzerlik fonksiyonu global konkav biçimindedir ve eğer 1.mertebeden türevlerin çözümü varsa, bu tek çözüm olarak ele alınacaktır. En çok benzerlik tahmin edicileri tutarlı, asimptotik etkin ve asimptotik normal tahmin edicilerdir. Asimptotik varyans-kovaryans matrisi $-\xi(\nabla^2 \log LL)^{-1}$ olarak simgelendiğinde, $\nabla^2 \log LL$ fonksiyonun 2. türevlerini göstermektedir. Matris şeklinde gösterilen $\nabla^2 \log LL$ ' nin k . sıra ve l . sütun elemanı aşağıdaki gibi bulunmaktadır.

$$[\nabla^2 \log LL]_{kl} = \frac{\partial^2 \log LL}{\partial \beta_k \partial \beta_l} \quad (5.31)$$

Tahmin edilen parametre değerleriyle hesaplanan varyans - kovaryans matrisi gerçek değerlerin tutarlı bir tahmincisi olarak değerlendirilmektedir.

$$\xi \left[\frac{\partial^2 \log LL}{\partial \beta_k \partial \beta_l} \right] \cong \sum_{n=1}^N \left[\frac{\partial^2 [Y_{in} \log P_n(i) + Y_{jn} \log P_n(j)]}{\partial \beta_k \partial \beta_l} \right]_{\beta=\hat{\beta}} \quad (5.32)$$

Asimptotik varyans-kovaryans matrisi (5.33)' de gösterilmektedir.

$$Var(\hat{\beta}) = \sum \left[\frac{f^2(\beta' X_i)}{F(\beta' X_i)[1 - F(\beta' X_i)]} X_i X_i' \right]^{-1} \quad (5.33)$$

2. türevlerin hesaplanabildiği ve benzerlik fonksiyonunun global konkav olduğu durumlarda Newton-Raphson iterasyon yöntemi kullanılarak parametre tahminleri elde edilebilmektedir. Bu yöntemin adımları aşağıda açıklanmaktadır.

1.adım: Başlangıç değeri olarak $\hat{\beta}_0 = [\hat{\beta}_{01}, \hat{\beta}_{02}, \dots, \hat{\beta}_{0K}]$ belirlenir. İterasyon sayacı (iterasyon sayısını belirten değer) $m = 0$ ve $e_1 = 10^{-4}, e_2 = 10^{-2}$ olarak ele alınacaktır.

2.adım: $\nabla \log LL(\beta)$ fonksiyonu $\hat{\beta}_w$ etrafında doğrusallaştırılır.

$$\nabla \log LL(\hat{\beta}_m) + \nabla^2 \log LL(\hat{\beta}_m)(\hat{\beta} - \hat{\beta}_m) = 0$$

3.adım: Bir sonraki iterasyonda çözümleme yapılır.

$$\hat{\beta}_{m+1} = \hat{\beta}_m - [\nabla^2 \log LL(\hat{\beta}_m)^{-1} \nabla \log LL(\hat{\beta}_m)]$$

4.adım: Aşağıda belirtilen kriterlerin sağlanıp sağlanmadığına bakılır.

$$\left[\frac{1}{K} \sum_{k=1}^K (\hat{\beta}_{m+1,k} - \hat{\beta}_{m,k})^2 \right]^{1/2} < e_1 \quad \text{ve/veya}$$

$$\left| \frac{\hat{\beta}_{m+1,k} - \hat{\beta}_{m,k}}{\hat{\beta}_{m,k}} \right| < e_2, k = 1, 2, \dots, K$$

Bu koşullar sağlanırsa $\hat{\beta}_{m+1}$ çözüm olarak belirlenir, aksi taktirde ise iterasyon sayacı $m = m + 1$ olarak alınır ve işlemler yeniden yapılarak çözüm bulunmaya çalışılır.

EYO tahmincilerini elde etmenin diğer bir yolu scoring yönteminin kullanılmasıdır. Newton-Raphson ve scoring yöntemlerinin karşılaştırılması açısından her iki yöntem için $\hat{\beta}_m$ veri iken, $\hat{\beta}_{m+1}$ tahmin edicisi hesaplanması aşağıda verilmektedir. İki yöntem, scoring yönteminde ikinci türevlerin beklenen değerlerin alınmasıyla farklılaşmaktadır.⁶⁷

$$\hat{\beta}_{m+1} = \hat{\beta}_m - \left[\frac{\partial^2 \ln LL}{\partial \beta \partial \beta'} \Big|_{\hat{\beta}_m} \right]^{-1} \cdot \frac{\partial \ln LL}{\partial \beta} \Big|_{\hat{\beta}_m} \quad (5.34)$$

$$\hat{\beta}_{m+1} = \hat{\beta}_m - \left[E \frac{\partial^2 \ln LL}{\partial \beta \partial \beta'} \Big|_{\hat{\beta}_m} \right]^{-1} \cdot \frac{\partial \ln LL}{\partial \beta} \Big|_{\hat{\beta}_m} \quad (5.35)$$

⁶⁷ Akın, F., a.g.e, s.31

5.3.3.2 İkili Logit Modelde En Çok Benzerlik Tahmini

Logit modelde benzerlik fonksiyonunun logaritması aşağıdaki gibi belirtilmektedir.

$$\log LL = \sum_{n=1}^N \left[Y_{in} \log \left(\frac{e^{\beta'X_{in}}}{e^{\beta'X_{in}} + e^{\beta'X_{jn}}} \right) + Y_{jn} \log \left(\frac{e^{\beta'X_{jn}}}{e^{\beta'X_{jn}} + e^{\beta'X_{in}}} \right) \right] \quad (5.36)$$

$X_n = X_{in} - X_{jn}$ olarak tanımlanır ve seçim olasılıkları bulunur.

$$P_n(i) = \frac{1}{1 + e^{-\beta'X_n}}, P_n(j) = \frac{e^{-\beta'X_n}}{1 + e^{-\beta'X_n}}$$

$$\frac{\partial \log P_n(i)}{\partial (\beta'X_n)} = \frac{e^{-\beta'X_n}}{(1 + e^{-\beta'X_n})^2} \cdot (1 + e^{-\beta'X_n}) = P_n(j)$$

$$\frac{\partial \log P_n(j)}{\partial (\beta'X_n)} = -P_n(i)$$

Bu sonuçlardan hareketle;

$$\begin{aligned} \frac{\partial \log LL(\beta)}{\partial \beta_K} &= \sum_{n=1}^N [Y_{in} P_n(j) X_{nk} - Y_{jn} P_n(i) X_{nk}] = \sum_{n=1}^N [Y_{in} (1 - P_n(i)) - (1 - Y_{in}) P_n(i)] X_{nk} \\ &= \sum_{n=1}^N [Y_{in} - P_n(i)] X_{nk}; k = 1, 2, \dots, K; X_{nk} = X_{in} - X_{jn} \end{aligned} \quad (5.37)$$

olarak elde edilir ve en çok benzerlik tahmin edicilerine ulaşmak için de K adet doğrusal olmayan denklem sistemi çözülür.

$$\sum_{n=1}^N [Y_{in} - P_n(i)] X_{nk} = 0; k = 1, 2, \dots, K$$

2.mertebeden kısmî türevler hesaplanabilmektedir.

$$\frac{\partial^2 \log LL}{\partial \beta_k \partial \beta_l} = - \sum_{n=1}^N P_n(i)(1 - P_n(i)) \cdot X_{nk} \cdot X_{nl} \quad (5.38)$$

İlk kısmî türevler alındığında bulunan sonuçların çözüm olarak değerlendirilebilmesi için fonksiyonunun global konkav olmasına gereksinim duyulmaktadır. Bunun tespiti için de 2.kısmî türevlerin yer aldığı matrisin negatif yarı-tanımli olup olmadığına bakılır. A matrisi $N * K$ boyutlu bir matris olsun ve matrisin elemanları şu şekilde bulunsun:

$$a_{nk} = [P_n(i)(1 - P_n(i))]^{1/2} \cdot X_{nk} \quad (5.39)$$

Matrisin tüm elemanları gerçek sayılar olduğundan $A'A$ pozitif yarı-tanımlı ve $-A'A$ ise negatif yarı-tanımlı olmaktadır.

$$A'A = \nabla^2 \log LL$$

5.3.3.3 İkili Probit Modelde En Çok Benzerlik Tahmini

Probit modelinde en çok benzerlik tahminlerinin elde edilmesi logit modeli için yapılan çözümlemelere benzemektedir. Benzerlik fonksiyonunun logaritması yazılır ve kısmî türevler alınır.

$$\log LL = \sum_{n=1}^N [Y_{in} \log \Phi(\beta'X_n) + (1 - Y_{in}) \log(1 - \Phi(\beta'X_n))] \quad (5.40)$$

$$\frac{\partial \log LL}{\partial \beta_k} = \sum_{n=1}^N \left[\frac{Y_{in} \phi(\beta'X_n)}{\Phi(\beta'X_n)} - \frac{(1 - Y_{in}) \phi(\beta'X_n)}{1 - \Phi(\beta'X_n)} \right] X_{nk}; k = 1, 2, \dots, K \quad (5.41)$$

5.4 Çoklu Seçim Modelleri

İkili seçim modellerinde C_n iki alternatif içerirken çoklu seçim modellerinde alternatif ya da seçenek sayısı 2' den fazla olmaktadır ve bu yüzden dolayı da çoklu seçim modellerine ilişkin uygun modelin elde edilmesi, tahmin yöntemlerinin kullanılması daha karışık bir hale gelmektedir.

5.4.1 Çoklu Seçim Teorisi

Seçilmesi mümkün olan tüm alternatifleri içeren küme evrensel, temel ya da seçim kümesi olarak adlandırılmakta ve C simgesiyle gösterilmektedir. Evrensel kümedeki eleman sayısı da J olarak belirtilmektedir. Bireyler evrensel kümedeki alternatifler arasından seçim yapmakta ve oluşturdukları seçim kümeleri evrensel kümenin bir alt kümesi olarak ortaya çıkmaktadır. Seçim kümesindeki i . alternatifin n . birey tarafından tercih edilmesi olasılığı aşağıda ifade edilmektedir.

$$\begin{aligned} P_n(i) &= P(U_{in} \geq U_{jn}) = P(V_{in} + \varepsilon_{in} \geq V_{jn} + \varepsilon_{jn}) \\ &= P(\varepsilon_{jn} \leq V_{in} - V_{jn} + \varepsilon_{in}); \forall j \in C_n, j \neq i \end{aligned}$$

Ana kütle hata paylarının birleşik yoğunluk fonksiyonu $f(\varepsilon_{1n}, \varepsilon_{2n}, \dots, \varepsilon_{J_n})$ olarak simgelenir ve alternatifler arasından da i birinci alternatif olarak belirlensin.

$$P_n(1) = \int_{\varepsilon_{1n}=-\infty}^{\infty} \int_{\varepsilon_{2n}=-\infty}^{V_{1n}-V_{2n}+\varepsilon_{1n}} \dots \int_{\varepsilon_{J_n n}=-\infty}^{V_{1n}-V_{J_n n}+\varepsilon_{1n}} f(\varepsilon_{1n}, \varepsilon_{2n}, \dots, \varepsilon_{J_n n}) d\varepsilon_{J_n n} \dots \varepsilon_{1n} \quad (5.42)$$

Yukarıdaki eşitlik (5.42) seçim olasılıklarının belirlenmesinde kullanışlı bir özelliğe sahip olmadığından dolayı daha uygun yollar kullanılarak seçim olasılıklarının belirlenmesi sağlanacaktır. F , anakütle hata paylarının birikimli dağılım fonksiyonunu göstermekte ve $F_i(\varepsilon_{1n}, \varepsilon_{2n}, \dots, \varepsilon_{J_n n})$ ise ε_{in} 'e bağlı olarak kısmî türevleri simgelemektedir. Buradan hareket edecek olursak $P_n(1)$ aşağıdaki gibi hesaplanabilir.

$$P_n(1) = \int_{\varepsilon_{1n}=-\infty}^{\infty} F_1(\varepsilon_{1n}, V_{1n} - V_{2n} + \varepsilon_{1n}, V_{1n} - V_{3n} + \varepsilon_{1n}, \dots, V_{1n} - V_{J_n n} + \varepsilon_{1n}) d\varepsilon_{1n} \quad (5.43)$$

ε_{1n} belirli bir değer olarak ayarlandığında integral işlemi ε_{1n} 'in bu değere eşit olma olasılığını vermektedir ve ayrıca tüm diğer hata payları da $V_{1n} + \varepsilon_{1n} \geq V_{jn} + \varepsilon_{jn}, \forall j \in C_n$ koşulunu gerçekleştirmelidir. ε_{1n} 'in mümkün olan tüm değerlerine göre integral işlemleri yapılırsa, 1. alternatifin seçilme olasılığı bulunmuş olur.

Diğer bir yol da çoklu seçim modelini ikili seçim modeline indirgeyerek seçim olasılıklarının ifade edilmesine yöneliktir.

$$U_{in} \geq U_{jn}, \forall j \in C_n; j \neq i$$

$$U_{in} \geq \max U_{jn} \Rightarrow P_n(i) = P[V_{in} + \varepsilon_{in} \geq \max(V_{jn} + \varepsilon_{jn})] \quad (5.44)$$

5.4.2 Çoklu Logit Model

5.4.2.1 Çoklu Logit Modelin Tanımlanması

Çoklu logit modelde i . alternatifin seçilme olasılığı genel bir ifadeyle şu şekilde belirtilir. Modelde $J_n = 2$ yani seçenek sayısının 2 olması durumunda çoklu model ikili modele indirgenmektedir.

$$P_n(i) = \frac{e^{V_{in}}}{\sum_{j \in C_n} e^{V_{jn}}}$$

Olasılıklara ilişkin aşağıdaki ifadeler yazılmaktadır.

$$0 \leq P_n(i) \leq 1, \forall i \in C_n$$

$$\sum_{i \in C_n} P_n(i) = 1$$

Modelde ε_{in} ' lere ilişkin birbirlerinden bağımsız olarak dağılma, aynı dağılıma sahip olma, η ve $\mu > 0$ parametreleriyle Gumbel dağılımına uyma varsayımları yapılmaktadır. Seçim olasılıkları yeniden yazılacak olursa (5.45) elde edilir.

$$P_n(i) = \frac{e^{\mu V_{in}}}{\sum_{j \in C_n} e^{\mu V_{jn}}} \quad (5.45)$$

5.4.2.2 Gumbel Dağılımına İlişkin Genel Özellikler

ε ' in gumbel dağılımına uygun olarak dağıldığı düşünülürse birikimli dağılım ve olasılık yoğunluk fonksiyonları aşağıdaki gibi yazılmaktadır.

$$F(\varepsilon) = \exp(-e^{-\mu(\varepsilon-\eta)}), \mu > 0 \quad (5.46)$$

$$f(\varepsilon) = \mu e^{-\mu(\varepsilon-\eta)} \exp(-e^{-\mu(\varepsilon-\eta)}) \quad (5.47)$$

Dağılıma ilişkin genel özellikler maddeler halinde verilmektedir.

- 1) Dağılımın modu η değerine eşittir.
- 2) Dağılımın ortalaması $(\eta + \gamma / \mu)$ değerine eşittir. ($\gamma \approx 0,577$)
- 3) Dağılımın varyansı $(\pi^2 / 6\mu^2)$ değerine eşittir.
- 4) Eğer ε değişkeni (η, μ) parametreleriyle gumbel dağılımına sahipse, $\alpha\varepsilon + V$ ($\alpha\eta + V, \mu/\alpha$) parametreleriyle gumbel dağılımına sahip olur. (V ve $\alpha > 0$ herhangi 2 sabit sayıyı belirtmektedir.)
- 5) Eğer ε_1 ve ε_2 sırasıyla (η_1, μ) ve (η_2, μ) parametrelerine sahip gumbel dağılımına uyan bağımsız değişkenlerse $\varepsilon^* = \varepsilon_1 - \varepsilon_2$ lojistik dağılıma uyacaktır.

$$F(\varepsilon^*) = \frac{1}{1 + e^{\mu(\eta_2 - \eta_1 - \varepsilon^*)}}$$

- 6) Eğer ε_1 ve ε_2 sırasıyla (η_1, μ) ve (η_2, μ) parametrelerine sahip gumbel dağılımına uyan bağımsız değişkenlerse $\max(\varepsilon_1, \varepsilon_2)$ gumbel dağılımına uyacaktır ve parametreleri aşağıdaki gibi olacaktır.

$$\left(\frac{1}{\mu} \ln(e^{\mu\eta_1} + e^{\mu\eta_2}), \mu \right)$$

- 7) Eğer $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_j)$ sırasıyla $(\eta_1, \mu), (\eta_2, \mu), \dots, (\eta_j, \mu)$ parametrelerine sahip gumbel dağılımına uyan bağımsız rasgele değişkenlerse $\max(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_j)$ aşağıdaki parametrelerle gumbel dağılımına uyacaktır.

$$\left(\frac{1}{\mu} \ln \sum_{j=1}^J e^{\mu\eta_j}, \mu \right)$$

5.4.2.3 Çoklu Logit Modelin Elde Edilmesi

Kolaylık olması açısından anakütle hata paylarının tümü için $\eta = 0$ alınarak işlemler yapılmaktadır.

$$P_n(1) = P[V_{1n} + \varepsilon_{1n} \geq \max(V_{jn} + \varepsilon_{jn}); j = 2, 3, \dots, J_n]$$

$U_n^* = \max(V_{jn} + \varepsilon_{jn})$ olarak tanımlanır ve gumbel dağılımına ilişkin 7 numaralı özellik kullanılırsa, U_n^* ' in aşağıda belirtilen parametrelerle gumbel dağılımına uyduğu söylenir.

$$\left(\frac{1}{\mu} \ln \sum_{j=2}^J e^{\mu V_{jn}}, \mu \right)$$

$U_n^* = V_n^* + \varepsilon_n^*$ olarak tanımlanır ve 4 numaralı özellik kullanılırsa ε_n^* ' in $(0, \mu)$ parametreleriyle gumbel dağılımına uyduğu söylenir.

$$V_n^* = \frac{1}{\mu} \ln \sum_{j=2}^J e^{\mu V_{jn}}$$

$$P_n(1) = P(V_{1n} + \varepsilon_{1n} \geq V_n^* + \varepsilon_n^*) = P[(V_n^* + \varepsilon_n^*) - (V_{1n} + \varepsilon_{1n}) \leq 0]$$

5 numaralı özellik kullanılırsa çoklu logit modeline ilişkin genel ifade elde edilmiş olur.

$$P_n(1) = \frac{1}{1 + e^{\mu(V_n^* - V_{1n})}} = \frac{e^{\mu V_{1n}}}{e^{\mu V_{1n}} + e^{\mu V_n^*}} = \frac{e^{\mu V_{1n}}}{e^{\mu V_{1n}} + \exp\left(\ln \sum_{j=2}^J e^{\mu V_{jn}}\right)} = \frac{e^{\mu V_{1n}}}{\sum_{j=1}^J e^{\mu V_{jn}}}$$

5.4.2.4 Çoklu Logit Modelde Sınırlı Durumlar (Limiting Cases)

İkili logit modelde bahsedildiği gibi çoklu modelde de μ parametresine göre 2 durum söz konusu olmaktadır. $\mu \rightarrow 0$ olması durumunda hata paylarının varyansı sonsuza yaklaşmakta ve alternatiflerin seçilme olasılığı aynı olmaktadır.

$$\lim_{\mu \rightarrow 0} P_n(i) = \frac{1}{J_n}, \forall i \in C_n \quad (5.48)$$

2.durum olarak $\mu \rightarrow \infty$ olması hali incelenecektir.

$$\lim_{\mu \rightarrow \infty} P_n(i) = \lim_{\mu \rightarrow \infty} \frac{1}{1 + \sum_{j \in C_n} e^{\mu(V_{jn} - V_{in})}} = \begin{cases} 1 & \text{eğer } V_{in} > \max_{i \neq j} V_{jn} \\ 0 & \text{eğer } V_{in} < \max_{i \neq j} V_{jn} \end{cases} \quad (5.49)$$

Bazı alternatifler arasında (J_n^* sayıdaki alternatif arasında) $V_{in} = \max_{i \neq j} V_{jn}$ durumu ortaya çıkabilmektedir. Bu durumda limit değeri J_n^* alternatif için $1/J_n^*$ değerine, geri kalan $(J_n - J_n^*)$ alternatif için de 0 değerine eşit olmaktadır. $\mu \rightarrow \infty$ olması durumunda hata paylarının varyansı 0'a yaklaşmakta ve deterministik seçim modeliyle karşılaşmaktadır.

5.4.2.5 Parametrelerinde Doğrusal Çoklu Logit Model

Fayda fonksiyonunun sistematik kısmı olan V_{in} parametrelerinde doğrusal bir fonksiyon şeklinde olduğunda n . bireyin i . alternatifi seçme olasılığı aşağıdaki gibi bulunmaktadır.

$$P_n(i) = \frac{e^{\beta'X_{in}}}{\sum_{j \in C_n} e^{\beta'X_{jn}}} \quad (5.50)$$

5.4.2.6 Logit Modelin Özellikleri

5.4.2.6.1 Alternatifler Arası Bağımsızlık (Independent from Irrelevant Alternatives)

Bu özelliğe göre herhangi 2 alternatifin seçim olasılıklarına ilişkin oran diğer tüm alternatiflerin sistematik kısımlarından etkilenmemektedir, diğer ifadeyle bu oran diğer alternatiflerden tamamıyla bağımsızdır.

$$\frac{P_n(i)}{P_n(l)} = \frac{e^{V_{in}} / \sum e^{V_{jn}}}{e^{V_{ln}} / \sum e^{V_{jn}}} = \frac{e^{V_{in}}}{e^{V_{ln}}} = e^{V_{in} - V_{ln}} \quad (5.51)$$

Seçim kümesi C_n ve bu seçim kümesinin bir alt kümesi olan $\tilde{C}_n$ için alternatiflere ilişkin seçim olasılıkları belirlenecek olursa;

$$P_n(i) = \frac{e^{\beta'X_{in}}}{\sum_{j \in C_n} e^{\beta'X_{jn}}}$$

$$P_n(i) = \frac{e^{\beta'X_{in}}}{\sum_{j \in \tilde{C}_n} e^{\beta'X_{jn}}}, \quad \tilde{C}_n \subseteq C_n$$

Alternatifler arasındaki bağımsızlık özelliğinin test edilmesinde Hausman ve McFadden tarafından geliştirilen test istatistiği kullanılmaktadır.

$$H_0 : \hat{\beta}_{\tilde{C}} = \hat{\beta}_C$$

$$(\hat{\beta}_{\tilde{C}} - \hat{\beta}_C)' (\Sigma_{\hat{\beta}_{\tilde{C}}} - \Sigma_{\hat{\beta}_C})^{-1} (\hat{\beta}_{\tilde{C}} - \hat{\beta}_C) \quad (5.52)$$

Bu test istatistiği χ^2 dağılımı yardımıyla test edilmektedir.⁶⁸

5.4.2.6.2 Çoklu Logit Modelde Elastikiyetler (Esneklikler)

Seçim modellerinde bütünleşik olmayan (disaggregate) ve bütünleşik (aggregate) esneklik çeşitlerinden söz edilecektir.

Disaggregate esneklik, herhangi bir özellikteki değer değişimine bağlı olarak bireyin alternatifini seçme olasılığındaki değişimi ifade etmektedir. En basit haliyle, bağımsız değişken

⁶⁸ McFadden hipotezin test edilmesinde Lagrange Çarpan testinin de kullanılabildiğini göstermiştir.

olarak belirtilen X_{ink} 'daki değişime bağlı olarak bireyin i . alternatifi seçme olasılığındaki esneklik olarak tanımlanmaktadır.

$$E_{X_{ink}}^{P_n(i)} = \frac{\partial P_n(i)}{\partial X_{ink}} \cdot \frac{X_{ink}}{P_n(i)} = \frac{\partial \ln P_n(i)}{\partial \ln X_{ink}} = [1 - P_n(i)] X_{ink} \cdot \beta_k \quad (5.53)$$

Disaggregate çapraz esneklik, j . alternatifin özelliğindeki değişime bağlı olarak i . alternatifi seçme olasılığındaki esneklik olarak tanımlanmaktadır.

$$E_{X_{jnk}}^{P_n(i)} = \frac{\partial \ln P_n(i)}{\partial \ln X_{jnk}} = -P_n(j) \cdot X_{jnk} \cdot \beta_k; i \neq j \quad (5.54)$$

Aggregate esneklik çeşidinde ise bireysel seçimler yerine grup bazında seçim olasılığının değişimi söz konusu olmaktadır. Gruptaki aynı özelliğe yani aynı bağımsız değişken değerine sahip birey sayısını N olarak ele aldığımızda, grubun i . alternatifi seçme olasılığı $\bar{P}(i)$ olarak tanımlanır.

$$\bar{P}(i) = \frac{\sum_{n=1}^N P_n(i)}{N} \quad (5.55)$$

Bağımsız değişken X_{jnk} 'nın değeri her bir birey için artışa tabi tutulacak olursa;

$$\frac{\partial X_{jnk}}{X_{jnk}} = \frac{\partial X_{jn'k}}{X_{jn'k}} = \frac{\partial X_{jk}}{X_{jk}}, n = 1, 2, \dots, N$$

$$X_{jk} = \frac{1}{N} \sum_{n=1}^N X_{jnk}$$

Esneklik ise aşağıdaki gibi tanımlanmaktadır:

$$E_{X_{jk}}^{\bar{P}(i)} = \frac{\sum_{n=1}^N P_n(i) \cdot E_{X_{jnk}}^{P_n(i)}}{\sum_{n=1}^N P_n(i)} \quad (5.56)$$

Her 2 esneklik çeşidi de δ_{ij} olarak simgelenen kronecker delta fonksiyonu yardımıyla gösterilebilmektedir. Bu fonksiyon $i = j$ olduğu durumlarda 1 değerine, diğer durumlarda ise 0 değerine eşittir.

$$E_{X_{jnk}}^{P_n(i)} = [\delta_{ij} - P_n(j)] X_{jnk} \beta_k \quad (5.57)$$

$$E_{X_{jnk}}^{\bar{P}(i)} = \frac{\beta_k}{N \cdot \bar{P}(i)} \cdot \sum_{n=1}^N P_n(i) [\delta_{ij} - P_n(j)] X_{jnk} \quad (5.58)$$

5.4.2.7 Arttırımlı (Incremental) Çoklu Logit Model

Parametrelerinde doğrusal logit modelde seçim olasılığının bulunmasında kullanılacak olan formül 5.4.2.5 numaralı alt bölümde verilmişti. Sistematik kısımda meydana gelecek bir değişmeye bağlı olarak elde edilecek olan seçim olasılıkları şu şekilde hesaplanmaktadır:

$$P'_n(i) = \frac{e^{V_{in} + \Delta V_{in}}}{\sum_{j \in C_n} e^{V_{jn} + \Delta V_{jn}}} \quad (5.59)$$

Formülde yer alan ΔV_{in} , i . alternatif için fayda fonksiyonunda meydana gelecek değişmeyi ve bu değişimin hesaplanmasında kullanılan formüldeki ΔX_{ink} ise n . birey ve i . alternatif için k . bağımsız değişkende meydana gelen değişmeyi göstermektedir.

$$\Delta V_{in} = \sum_{k=1}^K \beta_k \Delta X_{ink} \quad (5.60)$$

Pay ve paydanın $\sum_{j \in C_n} e^{V_{jn}}$ terimine bölünmesiyle incremental logit model elde edilmektedir.

$$P'_n(i) = \frac{P_n(i) \cdot e^{\Delta V_{in}}}{\sum_{j \in C_n} P_n(j) \cdot e^{\Delta V_{jn}}} \quad (5.61)$$

5.4.2.8 Çoklu Logit Modelin Tahmin Edilmesi

5.4.2.8.1 Çoklu Logit Modelde En Çok Benzerlik Yöntemi

Örnek birim mevcudu N olarak belirlensin. Bağımlı değişken tanımlaması yapılmaktadır.

$$Y_{in} = \begin{cases} 1 & \text{eğer } n. \text{ birey } i. \text{ alternatifini seçerse} \\ 0 & \text{diğer durumlarda} \end{cases} \quad (5.62)$$

Çoklu seçim modellerinde kullanılan benzerlik fonksiyonunun genel gösterimi aşağıda verilmektedir. Bu gösterimde $P_n(i)$ yerine parametrelerinde doğrusal logit modeldeki formül yazılır ve bu ifadenin logaritması alınarak işlemler gerçekleştirilir.

$$LL = \prod_{n=1}^N \prod_{i \in C_n} P_n(i)^{Y_{in}}$$

$$\log LL = \sum_{n=1}^N \sum_{i \in C_n} Y_{in} (\beta' X_{in} - \ln \sum_{j \in C_n} e^{\beta' X_{jn}}) \quad (5.63)$$

$\log LL$ fonksiyonunun katsayılar göre kısmî türevleri alınarak 0' a eşitlenir.

$$\frac{\partial \log LL}{\partial \beta_k} = \sum_{n=1}^N \sum_{i \in C_n} Y_{in} \left(X_{ink} - \frac{\sum_{j \in C_n} e^{\beta' X_{jn}} \cdot X_{jnk}}{\sum_{j \in C_n} e^{\beta' X_{jn}}} \right) = 0; k = 1, 2, \dots, K \quad (5.64)$$

$$\sum_{n=1}^N \sum_{i \in C_n} [Y_{in} - P_n(i)] X_{ink} = 0$$

5.4.2.8.2 Tekrarlı Gözlemler İçin En Çok Benzerlik Yöntemi

Tekrarlı gözlem kavramının anlaşılabilmesi açısından kişilerin seyahat tercihleriyle ilgili bir çalışma yapıldığını düşünelim. Alternatiflerin özellikleri sabit kalmakla birlikte kişilerin 1 ay boyunca yaptığı tercihler gün bazında elde edilmiş olsun. D_n , n . bireyin i . alternatifi seçme sayısını göstermek üzere;

$$D_n = \sum_{i \in C_n} D_{in}$$

Benzerlik fonksiyonunu ve bu fonksiyonun logaritmasını yazacak olursak;

$$LL = \prod_{n=1}^N \frac{D_n!}{\prod_{j \in C_n} D_{jn}!} \cdot \prod_{i \in C_n} P_n(i)^{D_{in}}$$

$$\log LL = \sum_{n=1}^N \left(\ln D_n! - \sum_{j \in C_n} \ln D_{jn}! + \sum_{i \in C_n} D_{in} \ln P_n(i) \right) \quad (5.65)$$

Daha önceden de değinildiği üzere kısmî türevlerin alınıp 0'a eşitlenmesi ve denklem sistemlerin çözülmesiyle en çok benzerlik tahmin edicileri elde edilebilmektedir.

6. ÇALIŞMA GRUPLARINI ETKİLEYEN ETKENLERİN ARACILIĞI İLE GRUPLARA AİT OLMA OLASILIKLARININ BELİRLENMESİNDE SIRALI LOGİT ve SIRALI PROBİT MODELLERİNİN KULLANIMI

İş yaşamında bireyler yaptıkları işlere, bulundukları görevlere veya yer aldıkları pozisyonlara göre çeşitli kategorilere ayrılmaktadır. Kategorilere ayırmada bireyin eğitim düzeyi, yaşı, elde ettiği başarılar, iş deneyim yılı gibi etkenler rol oynamaktadır. Çalışmamızda bireyler menial, mavi yaka, esnaf-sanatkar, beyaz yaka ve profesyonel olmak üzere 5 grupta incelenmektedir. Grup değerleri bağımlı değişken olarak kullanılmakta ve bireylerin gruplar arasında geçiş yapabilmesi mümkün olduğundan, örneğin mavi yakalı olan bir bireyin başarılı bulunan bir proje sonrasında beyaz yakalı olarak terfi ettirilmesi, bağımlı değişkenin sıralı özellikte olduğu görülmektedir. Bağımlı değişkeni etkileyen bağımsız değişkenler piyasa deneyim yılı, resmi kurumlarca onaylı eğitim yılı ve ırk olmak üzere 3 adettir. Bağımsız değişkenlerden 2' si nicel, 1' i de nitel özelliktedir. Tüm değişkenlere ilişkin tanımlamalar aşağıda yer almaktadır.

Bağımlı Değişken

- $Y_i = 0$ (çırak-stajyer, menial)
 $Y_i = 1$ (mavi yaka, blue collar)
 $Y_i = 2$ (el sanatları, esnaf - sanatkar, craft)
 $Y_i = 3$ (beyaz yaka, white collar)
 $Y_i = 4$ (profesyonel, professional)

Bağımsız Değişkenler

X_{1i} = piyasa deneyim yılı, çalışma deneyimi (yıl bazında)

X_{2i} = resmi kurumlarca onaylı eğitim yılı

X_{3i} = ırk

eğer kişi beyaz ise $X_{3i} = 1$

eğer kişi siyah ise $X_{3i} = 0$

6.1 İkili Model Analizleri

İlk aşamada ikili model uygulamaları yapılacak ve bu uygulama için bağımlı değişken aşağıda gösterildiği gibi 2 gruba ayrılacaktır. 0,1 ve 2 olarak tanımlanan gruplarda beden gücünün daha çok, diğer geri kalan iki grupta ise beyin gücünün daha çok kullanıldığı göz önüne alınarak ikili model uygulamaları için grup tanımlamaları yapılmaktadır. Aşağıda yer alan tanımlamaya göre 2. grupta yer alan bireylerin daha üstün kademelerde olduğunu söylemek mümkündür.

$$\left. \begin{array}{l} Y_i = 0 \\ Y_i = 1 \\ Y_i = 2 \end{array} \right\} \text{olanlar 1.grup} \quad \left. \begin{array}{l} Y_i = 3 \\ Y_i = 4 \end{array} \right\} \text{olanlar 2.grup}$$

1. grup için 1, 2. grup için ise 0 değeri kullanılarak bağımlı gölge değişken tanımlaması yapılmaktadır. Bu tanımlamanın ardından ikili doğrusal olasılık, logit ve probit model uygulamaları için E-views 3.0 paket programı kullanılmaktadır.

Çizelge 6.1 Modeldeki değişkenlerin tanımlayıcı istatistik değerleri

	Y	X_1	X_2	X_3
Mean	0.545994	20.50148	13.09496	0.916914
Median	1.000000	17.00000	12.00000	1.000000
Maximum	1.000000	66.00000	20.00000	1.000000
Minimum	0.000000	2.000000	3.000000	0.000000
Std. Dev.	0.498620	13.95936	2.946427	0.276423
Skewness	-0.184760	0.647828	-0.064465	-3.020981
Kurtosis	1.034136	2.433952	3.394780	10.12633
Jarque-Bera	56.18303	28.07116	2.421831	1225.695
Probability	0.000000	0.000001	0.297924	0.000000
Observations	337	337	337	337

Kurduğumuz modeldeki değişkenlere ilişkin tanımlayıcı istatistikler çizelge 6.1' de yer almaktadır. Bağımlı ve bağımsız değişkenlere ilişkin ortalama, medyan, maksimum, minimum, standart sapma, çarpıklık, basıklık, normal dağılıma uygunluğun belirlenmesinde kullanılan Jarque-Bera test istatistiği ve bu test istatistiğine ilişkin olasılık değeri, gözlem sayısı değerleri çizelgeden incelenmektedir.

Modelde yer alan bağımlı değişken 2 değer aldığından dolayı ikili tercih modelleri kullanılarak parametre tahminleri yapılacak ve bu modellere ilişkin açıklamalara yer verilecektir.

6.1.1. İkili Doğrusal Olasılık Modeli Sonuçları

İkili doğrusal olasılık modeline ilişkin parametre tahminleri, değişkenlerin katsayılarının yorumları, model sonucunda bağımlı değişkenin tahmin edilen değerleri, sorunlu gözlemler, kalıntılar ve kalıntıların normal dağılıma uygunluğunun test edilmesi incelenmektedir.

Çizelge 6.2 İkili doğrusal olasılık modelinde parametre tahminleri

Dependent Variable: Y
Method: Least Squares

Sample: 1 337

Included observations: 337

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	-0.004291	0.001699	-2.525114	0.0120
X2	-0.096268	0.008027	-11.99243	0.0000
X3	-0.136830	0.082550	-1.657528	0.0984
C	2.020046	0.144446	13.98479	0.0000
R-squared	0.306990	Mean dependent var		0.545994
Adjusted R-squared	0.300747	S.D. dependent var		0.498620
S.E. of regression	0.416953	Akaike info criterion		1.100113
Sum squared resid	57.89205	Schwarz criterion		1.145455
Log likelihood	-181.3690	F-statistic		49.17082
Durbin-Watson stat	0.564415	Prob(F-statistic)		0.000000

En küçük kareler yöntemi kullanılarak parametre tahminleri yapılmakta ve bunun sonucunda da bağımlı ve bağımsız değişkenler arasındaki bağıntı aşağıdaki gibi ifade edilmektedir.

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i}$$

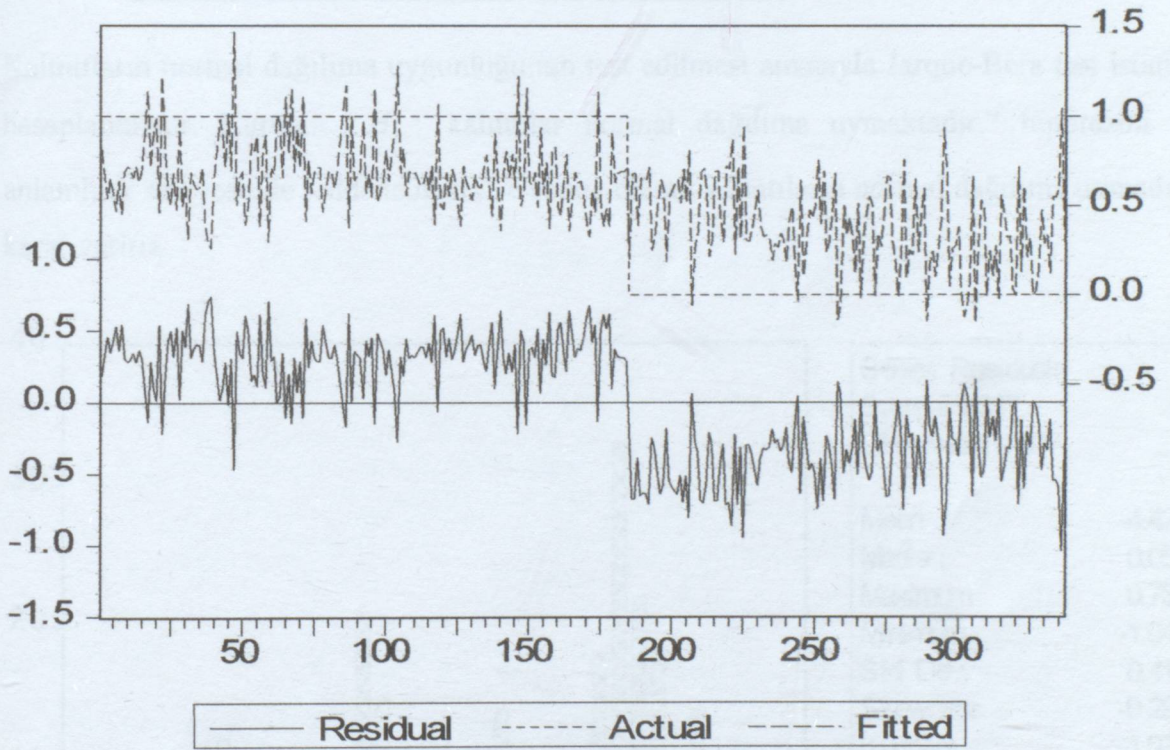
$$\hat{Y}_i = 2,020046 - 0,004291X_{1i} - 0,096268X_{2i} - 0,136830X_{3i}$$

Çizelgede katsayıların yanı sıra bu katsayılara ilişkin standart hatalar, t istatistik değerleri ve katsayıların anlamlılığının belirlenmesinde yararlanılan olasılık (probability) değerleri yer almaktadır. Katsayı değerinin standart hata değerine bölünmesiyle t istatistik değerleri elde edilmektedir. Prob. değeri ise $H_0 : \hat{\beta}_i = 0$ hipotezinin reddedilmesiyle yapılacak olan hatayı göstermektedir. Belirlenen anlamlılık seviyesinde katsayıların anlamlılıkları test edilebilmektedir. Kurulan modelde 0,05 anlamlılık seviyesine göre inceleme yaptığımızda sadece X_3 değişkenine ilişkin katsayının anlamsız olduğu sonucuna varırız. Katsayıların işareti bağımlı ve bağımsız değişkenler arasındaki ilişkinin yönünü göstermektedir. Kurulan modelin sonuçları incelendiğinde bağımsız değişkenler ile bağımlı değişken arasında ters yönlü bir ilişki olduğunu görürüz. Diğer bir ifadeyle bağımsız değişkendeki bir artış (azalış) bağımlı değişkende azalışa (artışa) yol açmaktadır. Katsayıların büyüklüğü ise bağımsız değişkenlerde meydana gelecek 1 birimlik değişimin, $P(Y = 1)$ üzerindeki etkisini

göstermektedir. Bu açıklamalara göre, eğitim durumundaki 1 birimlik yani 1 yıllık bir artış $P(Y = 1)$ olasılığını 0,096268 azaltmaktadır. Eğitim yılındaki artış, bireyin 2. gruba ait olma olasılığını arttırmaktadır.

Bağımlı değişkene ait ortalama ve standart sapma değerleri ile hata kareleri toplamı, R^2 , düzeltilmiş R^2 , katsayıların anlamlılıklarının birlikte test edilmesinde yararlanılan F istatistik değeri ve bu istatistiğe ilişkin olasılık değeri çizelgede görülmektedir. Belirtildiği üzere R^2 değeri nitel bağımlı değişkenli modellerde uyum iyiliğinin iyi bir ölçüsü olarak kullanılmamaktadır ve bu değer düşük bir seviyede olduğu çizelge 6.2' de gösterilmektedir.

Aşağıda yer alan grafikte residual kalıntıları, actual bağımlı değişkene ilişkin gözlenen değerleri ve fitted ise bağımlı değişkenin tahmin edilen değerlerini göstermektedir.



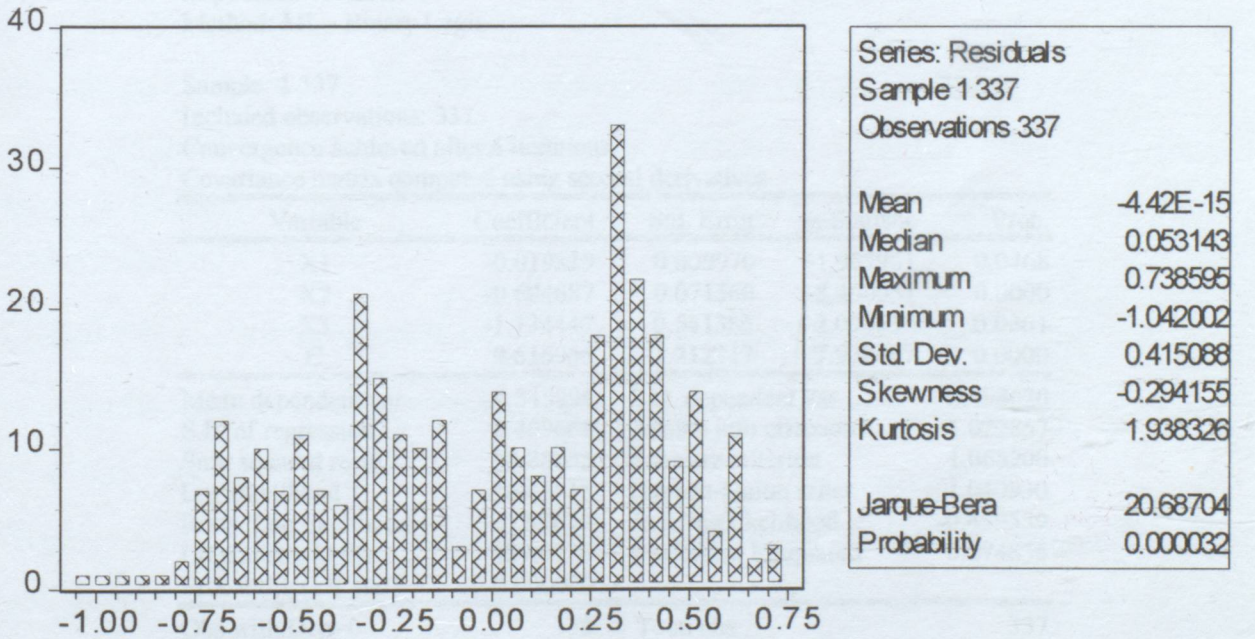
Şekil 6.1 İkili DOM' da kalıntı, gözlenen ve tahmin edilen değerler

Tahmin edilen değerler incelendiğinde 0-1 aralığı dışında yer alan gözlemler olduğunu görmekteyiz ve bu durum doğrusal olasılık modelinin zayıf noktalarından biri olarak ortaya çıkmaktadır. 17 adet tahmin değerinin 1' den büyük ve 15 adet değer de 0' dan küçük olduğunu gösteren çizelge 6.3' de gözlemler gözlem numaraları ile birlikte verilmektedir.

Çizelge 6.3 İkili DOM' da sorunlu gözlemler ve tahmin değerleri

Gözlem No	Tahmin	Gözlem No	Tahmin	Gözlem No	Tahmin
17	1,138270	104	1,279391	279	-0,008364
22	1,219320	118	1,065328	289	-0,149411
45	1,014386	146	1,221739	302	-0,123667
47	1,469981	149	1,157378	303	-0,059306
62	1,001514	174	1,133979	306	-0,157992
65	1,108235	207	-0,059852	310	-0,005945
67	1,146852	244	-0,040271	314	-0,023108
71	1,108235	258	-0,140829	320	-0,004073
85	1,031549	259	-0,053143	326	-0,005945
86	1,170250	266	-0,018817	336	1,042002
96	1,151216	271	-0,005945		

Kalıntıların normal dağılıma uygunluğunun test edilmesi amacıyla Jarque-Bera test istatistiği hesaplanmıştır. Kurulan " H_0 : kalıntılar normal dağılıma uymaktadır." hipotezini 0,05 anlamlılık seviyesinde reddedebiliriz ve sonuç olarak kalıntıların normal dağılıma uymadığına karar veririz.



Şekil 6.2 Jarque – Bera testi için frekans tablosu

6.1.2 İkili Logit Model Sonuçları

İkili logit modele ilişkin parametre tahminleri, değişkenlerin katsayılarının yorumları, bahis oranları, bağımlı değişkenin açıklanan değişkene göre gruplandırılmasıyla elde edilen tanımlayıcı istatistikler, belirlenen bir eşik değere göre çapraz çizelge ve uygunluk testleri incelenmektedir.

En çok benzerlik yöntemi kullanılarak parametre tahminleri yapılmış ve bu tahmin sürecinde Newton-Raphson nümerik iterasyon yöntemi kullanılarak 4 iterasyon sonucunda parametrelerin en çok benzerlik tahmin edicileri bulunmuştur. Elde edilen modele göre bağımlı değişken ile bağımsız değişkenler arasındaki bağlantı aşağıdaki gibi ifade edilmektedir.

$$I_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i}$$

$$\hat{Y}_i = \frac{1}{1 + \exp(-I_i)}$$

$$I_i = 9.616966 - 0.019819 X_{1i} - 0.604687 X_{2i} - 1.134447 X_{3i}$$

Çizelge 6.4 İkili logit modelde parametre tahminleri

Dependent Variable: Y

Method: ML - Binary Logit

Sample: 1 337

Included observations: 337

Convergence achieved after 4 iterations

Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
X1	-0.019819	0.009970	-1.987961	0.0468
X2	-0.604687	0.071560	-8.450051	0.0000
X3	-1.134447	0.541365	-2.095530	0.0361
C	9.616966	1.212717	7.930097	0.0000
Mean dependent var	0.545994	S.D. dependent var		0.498620
S.E. of regression	0.409668	Akaike info criterion		1.022857
Sum squared resid	55.88655	Schwarz criterion		1.068200
Log likelihood	-168.3515	Hannan-Quinn criter.		1.040930
Restr. Log likelihood	-232.1628	Avg. log likelihood		-0.499559
LR statistic (3 df)	127.6226	McFadden R-squared		0.274856
Probability(LR stat)	0.000000			
Obs with Dep=0	153	Total obs		337
Obs with Dep=1	184			

Logit ve probit modelde katsayılar doğrudan, bağımsız değişkenlerdeki bir değişimin bağımlı değişkenin beklenen değeri üzerindeki etkisi olarak yorumlanamamaktadır. Katsayının işareti bağımsız değişken ile olayın gerçekleşme olasılığı ($P(Y = 1)$) arasındaki ilişkinin yönünü gösterir. Katsayının işaretinin pozitif olması bu ilişkinin doğru yönlü ve negatif olması ters yönlü olduğunu belirtmektedir. Modelimizde bağımsız değişkenlerin tümü olayın gerçekleşme olasılığı ile ters yönlü bir ilişki içerisinde.

Teorik bilgilerden hatırlanacağı üzere lojistik regresyonda, regresyon katsayıları olan β_i 'lerin yorumlanmasında bahis oranları olarak bilinen odds ratio değerleri sıklıkla kullanılmaktaydı. Bu değerler bağımlı değişken olan Y 'nin X_p değişkeninin etkisi ile kaç kat fazla gözlenme olasılığına sahip olduğunu belirtmektedir. Modelimize ilişkin bahis oranı değerleri ve bu oranlara ilişkin yorumlar aşağıda verilmektedir.

Çizelge 6.5 İkili logit modelde değişkenlere ilişkin bahis oranları

	Bahis Oranı
Deneyim Yılı	0.980376
Eğitim yılı	0.546245
İrk	0.321599

Bulunan bu değerlere göre deneyim yılındaki bir birimlik artış (modelimizde 1 yıl) $P(Y = 1)$ olma olasılığını 0.98 kat arttırmaktadır. Diğer bir ifadeyle bu oranın 1 değerine yakın olması deneyim yılındaki bir birimlik artışın olasılık üzerinde önemli bir etkiye sahip olmadığını göstermektedir.

Çizelge 6.4' de yer alan değerlere ilişkin tanımlamalar aşağıda yer almaktadır.

Log likelihood: Log likelihood fonksiyonunun maksimum değerini belirtmekte ve $\log LL_1$ simgesiyle gösterilmektedir.

Avg.log likelihood: Log likelihood değerinin gözlem sayısına bölünmesiyle elde edilmektedir.

Restr.log likelihood: Modelde sadece sabit terim yer aldığında log likelihood fonksiyonunun maksimum değerini belirtmekte ve $\log LL_0$ simgesiyle gösterilmektedir.

LR Statistic: Sabit terim dışındaki katsayıların anlamlılıkların test edilmesinde kullanılan istatistik değeridir ve m (kısıt sayısı) serbestlik derecesi ile χ^2 dağılır.

$$LR \text{ Statistic} = -2(\text{Log}LL_0 - \text{Log}LL_1)$$

Probability(LR stat): Katsayıların birlikte anlamlılıklarının testi için kurulan hipotezin reddedilmesiyle yapılacak hatayı göstermektedir. Belirlenen anlamlılık düzeyinden küçük olması durumunda $H_0 : \beta_1 = \beta_2 = \dots = \beta_j = 0$ hipotezinin reddine karar verilir.

Mc Fadden R squared: Nitel bağımlı değişkenli modellerde uyum iyiliğinin bir ölçüsü olarak kullanılmaktadır.

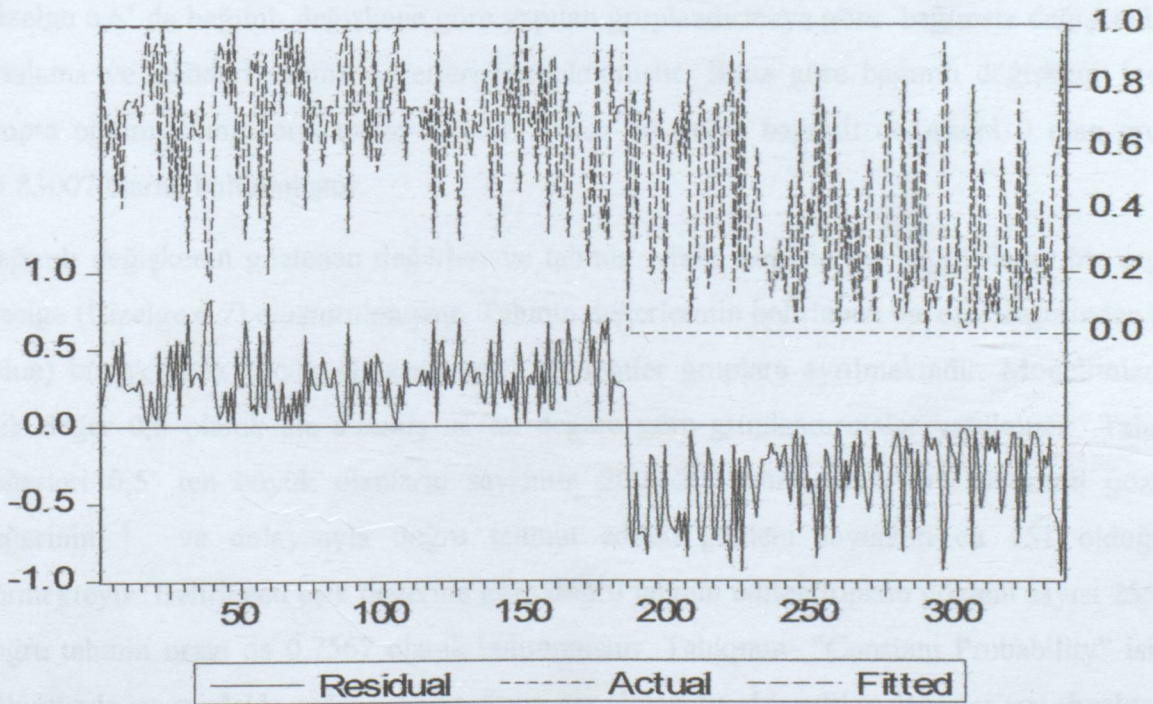
$$R_{MF}^2 = 1 - \frac{\log LL_1}{\log LL_0}$$

Tabloda ayrıca Akaike bilgi, Schwarz ve Hannan-Quinn olmak üzere 3 kritere de yer verilmektedir. Kriterlerin hesaplanmasında log-likelihood değeri ile gözlem ve parametre sayıları kullanılmaktadır.

$$AIC = -2(\log(LL_1)/n) + 2(k/n)$$

$$SC = -2(\log(LL_1)/n) + (k \log(n)/n)$$

$$HQ = -2(\log(LL_1)/n) + 2(k \log(\log(n))/n)$$



Şekil 6.3 İkili logit modelde kalıntı, gözlenen ve tahmin edilen değerler

Şekil 6.3’ de kalıntılar, bağımlı değişkenin gözlenen değerleri ve tahmin edilen değerleri birlikte gösterilmektedir. Tahmin edilen değerlerin tümünün 0-1 aralığı içinde yer aldığı ve doğrusal olasılık modelindeki “tahmin edilen değerlerin 0-1 aralığı dışına çıkması” sorununun logit model sayesinde ortadan kalktığı görülmektedir.

Çizelge 6.6 İkili logit modelde açıklanan değişkene göre tanımlayıcı istatistikler

Dependent Variable: Y

Method: ML – Binary Logit

Sample: 1 337

Included observations: 337

Descriptive statistics for explanatory variables

Variable	Dep=0	Mean Dep=1	All
X1	19.86275	21.03261	20.50148
X2	14.83007	11.65217	13.09496
X3	0.941176	0.896739	0.916914
Variable	Dep=0	Standard Deviation Dep=1	All
X1	13.37813	14.43959	13.95936
X2	2.672449	2.323248	2.946427
X3	0.236067	0.305130	0.276423
Observations	153	184	337

Çizelge 6.6’ da bağımlı değişkene göre yapılan gruplandırmaya göre bağımsız değişkenlerin ortalama ve standart sapma değerleri hesaplanmıştır. Buna göre bağımlı değişkeni 1 olan grupta eğitim yılının ortalaması 11,65217 iken bu değer bağımlı değişkeni 0 olan grupta 14,83007 olarak bulunmuştur.

Bağımlı değişkenin gözlenen değerleri ve tahmin edilen değerlerine bağlı olarak bir çapraz çizelge (Çizelge 6.7) oluşturulmuştur. Tahmin değerlerinin belirlenen bir eşik değerinden (cut value) büyük yada küçük olmasına göre gözlemler gruplara ayrılmaktadır. Modelimizdeki eşik değer 0,5 olarak ele alınmış ve bu değere göre gruplandırmalar yapılmıştır. Tahmin değerleri 0,5’ ten büyük olanların sayısının 200, 200 gözlemden 151 tanesinin gözlem değerinin 1 ve dolayısıyla doğru tahmin edilen gözlem sayısının da 151 olduğunu görmekteyiz. Belirlenen eşik değerine göre doğru tahmin edilen toplam gözlem sayısı 255 ve doğru tahmin oranı da 0,7567 olarak bulunmuştur. Tablonun “Constant Probability” isimli bölümünde ise modelde sadece sabit terimin yer almasıyla elde edilen değerler yer almaktadır. Modelde sadece sabit terim yer aldığında tahmin edilen tüm değerler eşik değerinden büyük ve doğru olarak tahmin edilen gözlem sayısı 184 olmaktadır.

Çizelge 6.7 İkili logit modelde eşik değere göre gruplandırma ve modelin doğruluğu

Dependent Variable: Y
Method: ML - Binary Logit

Sample: 1 337

Included observations: 337

Prediction Evaluation (success cutoff C = 0.5)

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)≤C	104	33	137	0	0	0
P(Dep=1)>C	49	151	200	153	184	337
Total	153	184	337	153	184	337
Correct	104	151	255	0	184	184
% Correct	67.97	82.07	75.67	0.00	100.00	54.60
% Incorrect	32.03	17.93	24.33	100.00	0.00	45.40
Total Gain*	67.97	-17.93	21.07			
Percent Gain**	67.97	NA	46.41			

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
E(# of Dep=0)	97.05	55.95	153.00	69.46	83.54	153.00
E(# of Dep=1)	55.95	128.05	184.00	83.54	100.46	184.00
Total	153.00	184.00	337.00	153.00	184.00	337.00
Correct	97.05	128.05	225.11	69.46	100.46	169.93
% Correct	63.43	69.59	66.80	45.40	54.60	50.42
% Incorrect	36.57	30.41	33.20	54.60	45.40	49.58
Total Gain*	18.03	14.99	16.37			
Percent Gain**	33.03	33.03	33.03			

Çizelgenin ikinci bölümünde ise gözlemlere ilişkin beklenen değerler yer almaktadır. Buna göre gözlem değeri 1 olanların beklenen değerlerinin, diğer bir ifadeyle tahmin edilen değerlerinin toplamı 128,05 olarak hesaplanmıştır. Tahmin edilen değerler, bağımlı değişkenin 1 olma olasılığını göstermekte ve bu olasılıkların toplam değeri, gözlem değeri 1 olanların toplam sayısına eşit olmaktadır. Tablodan görüldüğü gibi bağımlı değişkenin 1'e eşit olma olasılıklarının toplamı ve gözlem değeri 1 olanların sayısı 184 değerine eşittir.

Çizelge 6.8 İkili logit modelde Andrews ve Hosmer – Lemeshow uygunluk testi sonuçları

Dependent Variable: Y
Method: ML - Binary Logit

Sample: 1 337

Included observations: 337

Andrews and Hosmer-Lemeshow Goodness-of-Fit Tests

Grouping based upon predicted risk (randomize ties)

	Quantile of Risk		Dep=0		Dep=1		Total Obs	H-L Value
	Low	High	Actual	Expect	Actual	Expect		
1	0.0156	0.1218	33	31.1081	0	1.89189	33	2.00695
2	0.1250	0.2074	31	28.0105	3	5.98953	34	1.81122
3	0.2123	0.3549	21	24.5994	13	9.40062	34	1.90481
4	0.3743	0.4795	17	18.6404	16	14.3596	33	0.33175
5	0.4796	0.6208	13	14.9811	21	19.0189	34	0.46833
6	0.6231	0.7005	10	11.3949	24	22.6051	34	0.25683
7	0.7026	0.7385	8	9.09380	25	23.9062	33	0.18161
8	0.7404	0.7644	9	8.46991	25	25.5301	34	0.04418
9	0.7661	0.9103	8	5.34417	26	28.6558	34	1.56598
10	0.9142	0.9977	3	1.35781	31	32.6422	34	2.06874
Total			153	153.000	184	184.000	337	10.6404
H-L Statistic:			10.6404		Prob[Chi-Sq(8 df)]:		0.2229	
Andrews Statistic:			36.7336		Prob[Chi-Sq(10 df)]:		0.0001	

Bağımlı değişkenin tahmin edilen değerleri kantil şeklinde 10 gruba ayrılmıştır ve her kantile ilişkin en düşük ve en yüksek olasılık değerleri ile gerçekleşen ve beklenen gözlem sayıları tabloda görülmektedir. Andrews ve Hosmer-Lemeshow tarafından ortaya konulan uyum iyiliği ölçülerindeki ana fikir gerçekleşen ve beklenen gözlem frekanslarının karşılaştırılmasıdır. “ H_0 : gözlenen ve beklenen frekanslar arasında fark yoktur.” şeklinde kurulan hipotezin belirlenen anlamlılık seviyelerine göre kabul ya da reddine karar verilmektedir. Veriler J adet gruba ayrıldığında her bir gruptaki gözlem sayısı n_j , bu gruptaki bağımlı değişkeni 1 olan gözlemlerin sayısı n_{yj} ve tahmin değerleri toplamının gözlem sayısına bölünmesiyle elde edilen değer de $\bar{p}_{yj}$ ile gösterildiğinde Hosmer-Lemeshow istatistiği aşağıdaki gibi hesaplanmaktadır.

$$HL = \sum_{j=1}^J \frac{(n_{yj} - n_j \cdot \bar{p}_{yj})^2}{n_j \bar{p}_{yj} (1 - \bar{p}_{yj})}$$

İlk 2 kantile ilişkin hesaplamalar yapılmıştır:

$$HL_1 = \frac{(0 - 33 \cdot 0,057330)^2}{33 \cdot 0,057330 \cdot 0,942670} = \frac{3,57925}{1,783428} = 2,006949$$

$$HL_2 = \frac{(3 - 34 \cdot 0,176163)^2}{34 \cdot 0,176163 \cdot 0,823837} = \frac{8,937284}{4,934398} = 1,811221$$

HL test istatistiğine göre 0,05 anlamlılık düzeyinde H_0 hipotezini reddetmek için yeterli bir neden bulunmamaktadır şeklinde bir sonuca varılır.

6.1.3 İkili Probit Model Sonuçları

En çok benzerlik yöntemi kullanılarak parametre tahminleri yapılmış ve bu tahmin sürecinde Newton-Raphson nümerik iterasyon yöntemi kullanılarak 4 iterasyon sonucunda parametrelerin en çok benzerlik tahmin edicileri bulunmuştur. Elde edilen modele göre bağımlı değişken ile bağımsız değişkenler arasındaki bağlantı aşağıdaki gibi ifade edilmektedir.

$$I_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i}$$

$$P_i = P(Y = 1) = F(I_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{I_i} e^{-z^2/2} dz$$

Çizelge 6.9 İkili probit modelde parametre tahminleri

Dependent Variable: Y

Method: ML - Binary Probit

Sample: 1 337

Included observations: 337

Convergence achieved after 4 iterations

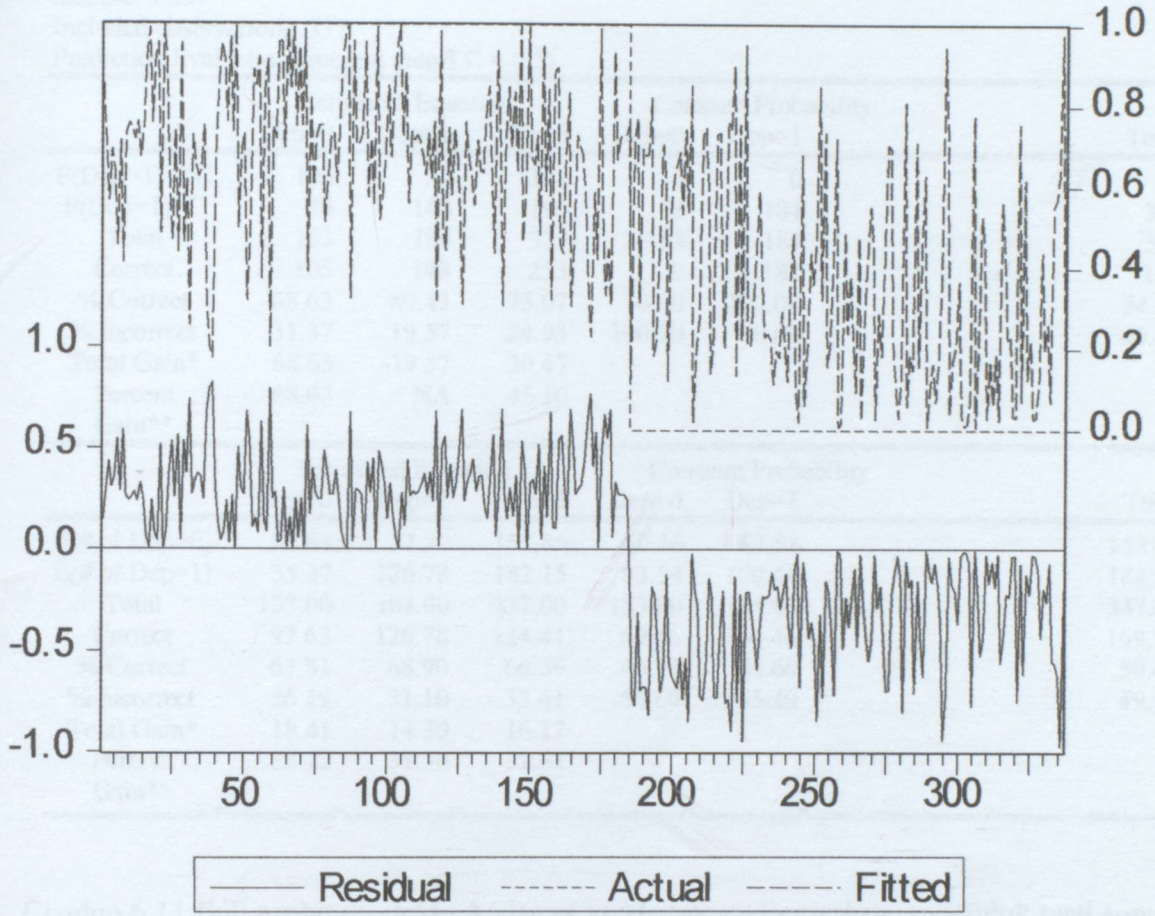
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
X1	-0.012895	0.005879	-2.193513	0.0283
X2	-0.353522	0.038615	-9.155129	0.0000
X3	-0.640050	0.305845	-2.092728	0.0364
C	5.613197	0.659304	8.513827	0.0000
Mean dependent var	0.545994	S.D. dependent var		0.498620
S.E. of regression	0.410148	Akaike info criterion		1.024092
Sum squared resid	56.01783	Schwarz criterion		1.069434
Log likelihood	-168.5594	Hannan-Quinn criter.		1.042164
Restr. log likelihood	-232.1628	Avg. Log likelihood		-0.500176
LR statistic (3 df)	127.2067	McFadden R-squared		0.273960
Probability(LR stat)	0.000000			
Obs with Dep=0	153	Total obs		337
Obs with Dep=1	184			

Öncelikle her gözlem için endeks değeri bulunmakta ve bu endeks değeri gözleme ilişkin olasılığın hesaplanmasında kullanılmaktadır. Normal dağılım tablosundan $-\infty$ dan I_i değerine kadar olan alan bulunduğunda, bu değer i . gözlemin 1' e eşit olma olasılığı olarak ifade edilmektedir.

Şekil 6.4' de kalıntılar, bağımlı değişkenin gözlenen değerleri ve tahmin edilen değerleri birlikte gösterilmektedir. Tahmin edilen değerlerin tümünün 0-1 aralığı içinde yer aldığı ve

doğrusal olasılık modelindeki “tahmin edilen değerlerin 0-1 aralığı dışına çıkması” sorununun probit model sayesinde ortadan kalktığı görülmektedir.



Şekil 6.4 İkili probit modelde kalıntı, gözlenen ve tahmin edilen değerler

Bağımlı değişkenin gözlenen değerleri ve tahmin edilen değerlerine bağlı olarak logit modeldekine benzer bir çapraz çizelge (Çizelge 6.10) oluşturulmuştur. Tahmin değerlerinin belirlenen bir eşik değerinden (cut value) büyük yada küçük olmasına göre gözlemler gruplara ayrılmaktadır. Modelimizdeki eşik değer 0,5 olarak ele alınmış ve bu değere göre gruplandırmalar yapılmıştır. Tahmin değerleri 0,5’ ten büyük olanların sayısının 196, 196 gözlemden 148 tanesinin gözlem değerinin 1 ve dolayısıyla doğru tahmin edilen gözlem sayısının da 148 olduğunu görmekteyiz. Belirlenen eşik değerine göre doğru tahmin edilen toplam gözlem sayısı 253 ve doğru tahmin oranı da 0,7507 olarak bulunmuştur. Tablonun “Constant Probability” isimli bölümünde ise modelde sadece sabit terimin yer almasıyla elde edilen değerler yer almaktadır. Modelde sadece sabit terim yer aldığında tahmin edilen tüm değerler eşik değerinden büyük ve doğru olarak tahmin edilen gözlem sayısı 184 olmaktadır.

Çizelge 6.10 İkili probit modelde eşik değere göre gruplandırma ve modelin doğruluğu

Dependent Variable: Y

Method: ML – Binary Probit

Sample: 1 337

Included observations: 337

Prediction Evaluation (success cutoff C = 0.5)

	Estimated Equation			Constant Probability		Total
	Dep=0	Dep=1	Total	Dep=0	Dep=1	
P(Dep=1)≤C	105	36	141	0	0	0
P(Dep=1)>C	48	148	196	153	184	337
Total	153	184	337	153	184	337
Correct	105	148	253	0	184	184
% Correct	68.63	80.43	75.07	0.00	100.00	54.60
% Incorrect	31.37	19.57	24.93	100.00	0.00	45.40
Total Gain*	68.63	-19.57	20.47			
Percent Gain**	68.63	NA	45.10			

	Estimated Equation			Constant Probability		Total
	Dep=0	Dep=1	Total	Dep=0	Dep=1	
E(# of Dep=0)	97.63	57.22	154.85	69.46	83.54	153.00
E(# of Dep=1)	55.37	126.78	182.15	83.54	100.46	184.00
Total	153.00	184.00	337.00	153.00	184.00	337.00
Correct	97.63	126.78	224.41	69.46	100.46	169.93
% Correct	63.81	68.90	66.59	45.40	54.60	50.42
% Incorrect	36.19	31.10	33.41	54.60	45.40	49.58
Total Gain*	18.41	14.30	16.17			
Percent Gain**	33.72	31.50	32.61			

Çizelge 6.11 İkili probit modelde Andrews ve Hosmer – Lemeshow uygunluk testi sonuçları

Dependent Variable: Y

Method: ML - Binary Probit

Sample: 1 337

Included observations: 337

Andrews and Hosmer-Lemeshow Goodness-of-Fit Tests

Grouping based upon predicted risk (randomize ties)

	Quantile of Risk		Dep=0		Dep=1		Total	H-L Value
	Low	High	Actual	Expect	Actual	Expect	Obs	
1	0.0072	0.1204	33	31.3353	0	1.66465	33	1.75309
2	0.1230	0.2158	31	27.8773	3	6.12270	34	1.94244
3	0.2235	0.3611	21	24.3947	13	9.60529	34	1.67216
4	0.3670	0.4787	17	18.7042	16	14.2958	33	0.35841
5	0.4808	0.6080	13	15.4325	21	18.5675	34	0.70209
6	0.6101	0.6819	9	11.9522	25	22.0478	34	1.12449
7	0.6865	0.7264	9	9.56177	24	23.4382	33	0.04647
8	0.7307	0.7556	9	8.78817	25	25.2118	34	0.00689
9	0.7556	0.9132	8	5.61463	26	28.3854	34	1.21387
10	0.9132	0.9998	3	1.19351	31	32.8065	34	2.83375
Total			153	154.854	184	182.146	337	11.6537
H-L Statistic:			11.6537			Prob[Chi-Sq(8 df)]:		0.1673
Andrews Statistic:			46.5047			Prob[Chi-Sq(10 df)]:		0.0000

Bağımlı değişkenin tahmin edilen değerleri logit modelde olduğu gibi kantil şeklinde 10 gruba ayrılmıştır ve her kantile ilişkin en düşük ve en yüksek olasılık değerleri ile gerçekleşen ve beklenen gözlem sayıları tabloda görülmektedir.

İlk 2 kantile ilişkin hesaplamalar yapılmıştır:

$$HL_1 = \frac{(0 - 33.0,050444)^2}{33.0,050444.0,949556} = \frac{2,77106}{1,580678} = 1,75309$$

$$HL_2 = \frac{(3 - 34.0,180079)^2}{34.0,180079.0,819921} = \frac{9,751255}{5,020128} = 1,94244$$

HL test istatistiğine göre 0,05 anlamlılık düzeyinde H_0 hipotezini reddetmek için yeterli bir neden bulunmamaktadır şeklinde bir sonuca varılır.

Doğrusal olasılık modelinin parametre tahminlerinin elde edilmesinde en küçük kareler yöntemi kullanılırken, logit ve probit modelde en çok benzerlik tahmin yöntemi kullanıldı. Doğrusal olasılık modelinde sorunlu, diğer bir ifadeyle tahmin edilen olasılık değeri 0-1 aralığı dışında olan gözlem sayısı 32 olarak bulundu. Logit ve probit model sayesinde olasılık değerlerinin 0-1 aralığı dışında olması sorununun ortadan kalktığı şekil 6.3 ve şekil 6.4' de gösterilmektedir. İkili logit modelde doğru olarak tahmin edilen gözlem sayısı 255 ve doğru gözlem oranı 0,7567 iken ikili probit modelde gözlem sayısı 253 ve gözlem oranı ise 0,7507 olarak elde edildi. Bu değerlerden hareketle modellerin performanslarının birbirine yakın olduğu söylenebilir. Diğer yandan, her iki modeldeki *McFaddenR²* değerleri birbirine çok yakındır. Hem ikili logit hem de ikili probit modelde beklenen ve gözlenen frekanslar arasında fark yoktur şeklinde kurulan hipotezleri reddetmek için yeterli neden bulunamamaktadır sonucuna varıldı.

6.2 Çoklu Model Analizleri

İkili modellerden daha karmaşık yapıda olan çoklu tercih modellerinde bireyler ikiden fazla tercihle karşı karşıyadır ve dolayısıyla bağımlı değişken ikiden fazla değer almaktadır. Çalışmamızda bireyler çalışma statülerine göre 5 gruba ayrılmakta ve grup değerleri bağımlı değişken olarak modelde yer almaktadır.

Bağımsız değişkenlere ilişkin tanımlayıcı istatistikler ikili modellerdeki değerlerle aynı olmakta fakat bağımlı değişkene ait değerler, bağımlı değişkenin tanımlanmasından dolayı değişiklik göstermektedir. Bağımlı değişkene ilişkin ortalama, medyan, maksimum, minimum, standart sapma, çarpıklık, basıklık, normal dağılıma uygunluğun belirlenmesinde kullanılan Jarque-Bera test istatistiği ve bu test istatistiğine ilişkin olasılık değeri, gözlem sayısı değerleri çizelge 6.12’ de gösterilmektedir.

Çizelge 6.12 Çoklu modelde bağımlı değişkenin tanımlayıcı istatistik değerleri

	Y
Mean	2,397626
Median	2,000000
Maximum	4,000000
Minimum	0,000000
Std. Dev.	1,367913
Skewness	-0,176035
Kurtosis	1,733727
Jarque-Bera	24,255590
Probability	0,000005
Observations	337

Bağımlı değişkenin değerlerine bakıldığında sıralı ölçekli olduğu görülmektedir ve bu yüzden dolayı sıralı logit ve probit modellere ilişkin tahmin yöntemleri uygulanarak sonuçlar elde edilmektedir.

6.2.1 Çoklu Logit Model Sonuçları

Bağımlı değişkenin her grubuna ait gözlem sayısı çizelge 6.13 gösterilmektedir. Gözlem sayısı en çok olan çalışan grubu profesyonel, en az olan grup ise menial grubudur. Çizelge 6.13’ de ayrıca, her grubun yüzdesel olarak dağılımı, birikimli gözlem sayısı değerleri ve birikimli yüzde dağılımları gösterilmektedir.

Çizelge 6.14’ de gözlem sayısı, sıralı bağımlı değişkene ilişkin grup sayısı ve tahmin sürecindeki iterasyon sayısı gösterilmektedir. En çok benzerlik yöntemi kullanılarak parametre tahminleri yapılmış ve bu tahmin sürecinde Newton-Raphson nümerik iterasyon

yöntemi kullanılarak 5 iterasyon sonucunda parametrelerin en çok benzerlik tahmin edicileri bulunmuştur.

Çizelge 6.13 Sıralı modelde bağımlı değişkenin gruplara göre dağılımı

Dependent Variable: Y
Method: ML – Ordered Logit

Sample: 1 337
Included observations: 337
Frequencies for dependent variable

Value	Count	Percent	Cumulative Count	Percent
0	31	9.20	31	9.20
1	69	20.47	100	29.67
2	84	24.93	184	54.60
3	41	12.17	225	66.77
4	112	33.23	337	100.00

Çizelge 6.14 Sıralı logit modelde parametre ve limit değerleri tahminleri

Dependent Variable: Y
Method: ML - Ordered Logit

Sample: 1 337
Included observations: 337
Number of ordered indicator values: 5
Convergence achieved after 5 iterations
Covariance matrix computed using second derivatives

	Coefficient	Std. Error	z-Statistic	Prob.
X1	0.024909	0.007731	3.221757	0.0013
X2	0.455863	0.044564	10.22946	0.0000
X3	0.383338	0.391735	0.978565	0.3278
Limit Points				
LIMIT_1:C(4)	3.931627	0.712980	5.514358	0.0000
LIMIT_2:C(5)	5.591815	0.725059	7.712217	0.0000
LIMIT_3:C(6)	6.974040	0.757057	9.212040	0.0000
LIMIT_4:C(7)	7.703901	0.779408	9.884294	0.0000
Akaike info criterion	2.675616	Schwarz criterion		2.754965
Log likelihood	-443.8413	Hannan-Quinn criter.		2.707243
Restr. log likelihood	-509.8441	Avg. log likelihood		-1.317037
LR statistic (3 df)	132.0055	LR index (Pseudo-R2)		0.129457
Probability(LR stat)	0.000000			

Çizelgede bağımsız değişkenlere ilişkin katsayı değerleri ve bağımlı değişkenin herhangi bir gruba ait olma olasılığının hesaplanmasında kullanılan limit değerler verilmektedir. Katsayıların anlamlılıkları 0,05 anlamlılık seviyesinde test edildiğinde ırk değişkenine ilişkin katsayının anlamsız olduğu sonucuna varılır. Aşağıda yer alan log olabilirlik fonksiyonunun maksimizasyonu ile katsayı ve limit değerleri tespit edilmektedir. Log olabilirlik fonksiyonunun maksimum değeri -443.8413 olarak hesaplanmıştır.

Olabilirlik oranı testine göre $H_0 : \beta_1 = \beta_2 = \dots = \beta_j = 0$ şeklinde kurulan hipotezi reddetmekle yapılacak olan hata 0 olarak bulunmuştur ve bu sonuç doğrultusunda 0,05 anlamlılık seviyesinde en az bir katsayının 0'dan farklı olduğu kararına varılmaktadır.

$$\log LL = \sum_{i=1}^N \sum_{j=0}^M \log(P(Y_i = j / X_i, \beta, \gamma)) \cdot l(Y_i = j)$$

$l(Y_i = j)$ ifadesi, $Y_i = j$ olduğunda 1 değerini, diğer durumlarda ise 0 değerini almaktadır.

Bağımlı değişkenin 5 grubu olduğundan her bir gözlem için 5 olasılık değeri hesaplanmaktadır ve bu hesaplamalarda lojistik ile normal dağılıma ilişkin birikimli dağılım fonksiyonları kullanılmaktadır. Modelimizde her gözlem için gruplara ait olma olasılıkları aşağıda ifade edildiği gibi bulunmaktadır.

$$I_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i}$$

$$I_i = 0.024909 X_{1i} + 0.455863 X_{2i} + 0.038338 X_{3i}$$

$$P(Y_i = 0) = \Lambda(3.931627 - I_i)$$

$$P(Y_i = 1) = \Lambda(5.591815 - I_i) - \Lambda(3.931627 - I_i)$$

$$P(Y_i = 2) = \Lambda(6.974040 - I_i) - \Lambda(5.591815 - I_i)$$

$$P(Y_i = 3) = \Lambda(7.703901 - I_i) - \Lambda(6.974040 - I_i)$$

$$P(Y_i = 4) = 1 - \Lambda(7.703901 - I_i)$$

Her gözlem için 5 olasılık değeri hesaplanmakta ve gözlemin, en yüksek olasılık değerine sahip gruba ait olduğu şeklinde bir karara varılabilmektedir.

Çizelge 6.15' de bağımlı değişkenin her bir grubuna ilişkin gerçekleşen ve beklenen gözlem sayıları ile bu sayılar arasındaki farkı gösteren gözlem hata değeri gösterilmektedir. Çizelgeye göre gözlenen değeri 0 olan bağımlı değişken sayısı 31 ve modelin tahmin edilmesiyle grubu 0 olarak belirlenen gözlem sayısı da 12 olmaktadır. Bağımlı değişkenin 0 olma olasılıklarının toplamı 30.336 olarak hesaplanmıştır ve gözlem sayısından bu olasılık değerinin çıkarılmasıyla olasılık hata değeri elde edilmiştir.

Çizelge 6.15 Sıralı logit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları

Dependent Variable: Y
Method: ML - Ordered Logit

Sample: 1 337

Included observations: 337

Prediction table for ordered dependent variable

Value	Count	Count of obs With Max Prob	Error	Sum of all Probabilities	Error
0	31	12	19	30.336	0.664
1	69	42	27	63.922	5.078
2	84	131	-47	84.466	-0.466
3	41	0	41	44.550	-3.550
4	112	152	-40	113.726	-1.726

Katsayıların anlamlılıkları 0,05 anlamlılık seviyesinde test edildiğinde ırk değişkenine ilişkin katsayının anlamsız olduğu sonucuna varılmıştı. Irk değişkenini modele dahil etmeden 2 bağımsız değişkenle modeli yeniden kuracak olursak 5 iterasyon sonucunda elde edilen EYO benzerlik tahmin edicilerinin 0,05 anlamlılık seviyesinde anlamlı oldukları sonucuna ulaşılır.

Çizelge 6.16 Sorunlu değişkenin çıkartılmasıyla elde edilen sıralı logit modelde parametre ve limit değerleri

Dependent Variable: Y
Method: ML - Ordered Logit

Sample: 1 337

Included observations: 337

Number of ordered indicator values: 5

Convergence achieved after 5 iterations

Covariance matrix computed using second derivatives

	Coefficient	Std. Error	z-Statistic	Prob.
X1	0.024567	0.007719	3.182533	0.0015
X2	0.456567	0.044580	10.24158	0.0000
Limit Points				
LIMIT_1:C(3)	3.580458	0.610686	5.863015	0.0000
LIMIT_2:C(4)	5.238953	0.622871	8.410980	0.0000
LIMIT_3:C(5)	6.620904	0.659246	10.04314	0.0000
LIMIT_4:C(6)	7.349283	0.683712	10.74909	0.0000
Akaike info criterion	2.672520	Schwarz criterion	2.740533	
Log likelihood	-444.3195	Hannan-Quinn criter.	2.699629	
Restr. log likelihood	-509.8441	Avg. log likelihood	-1.318456	
LR statistic (2 df)	131.0490	LR index (Pseudo-R2)	0.128519	
Probability(LR stat)	0.000000			

Katsayılar ve limit değerlerden hareketle bağımlı değişkene ilişkin olasılıklı ifadeler yazılmaktadır.

$$I_i = \beta_1 X_{1i} + \beta_2 X_{2i}$$

$$I_i = 0.024567 X_{1i} + 0.456567 X_{2i}$$

$$P(Y_i = 0) = \Lambda(3.580458 - I_i)$$

$$P(Y_i = 1) = \Lambda(5.238953 - I_i) - \Lambda(3.580458 - I_i)$$

$$P(Y_i = 2) = \Lambda(6.620904 - I_i) - \Lambda(5.238953 - I_i)$$

$$P(Y_i = 3) = \Lambda(7.349283 - I_i) - \Lambda(6.620904 - I_i)$$

$$P(Y_i = 4) = 1 - \Lambda(7.349283 - I_i)$$

Çizelge 6.17 Sorunlu değişkenin çıkartılmasıyla elde edilen sıralı logit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları

Dependent Variable: Y
Method: ML - Ordered Logit

Sample: 1 337

Included observations: 337

Prediction table for ordered dependent variable

Value	Count	Count of obs with Max Prob	Error	Sum of all Probabilities	Error
0	31	9	22	30.153	0.847
1	69	43	26	64.087	4.913
2	84	131	-47	84.430	-0.430
3	41	0	41	44.398	-3.398
4	112	154	-42	113.932	-1.932

Kurulan bu modele göre 2 ve 3 olarak tanımlanmış gruplara ilişkin tahmin edilen gözlem sayılarının 3 bağımsız değişkenin yer aldığı ilk modelle aynı, diğer gruplarda da sayıların birbirine çok yakın olduğunu görmekteyiz.

6.2.2 Çoklu Probit Model Sonuçları

Bireyler çalışma statülerine göre 5 gruba ayrıldığından her birey için gruplara ait olma olasılıklarını gösteren 5 değer hesaplanmakta ve probit modelde bu hesaplamalar normal dağılıma ilişkin birikimli dağılım fonksiyonu kullanılarak yapılmaktadır.

En çok benzerlik yöntemi kullanılarak parametre tahminleri yapılmış ve bu tahmin sürecinde Newton-Raphson nümerik iterasyon yöntemi kullanılarak 5 iterasyon sonucunda parametrelerin en çok benzerlik tahmin edicileri ve limit değerler bulunmuştur.

Katsayıların anlamlılıkları 0,05 anlamlılık seviyesinde test edildiğinde ırk değişkenine ilişkin katsayının anlamsız olduğu sonucuna varılır. Log olabilirlik fonksiyonunun maksimizasyonu ile katsayı ve limit değerleri tespit edilmektedir. Log olabilirlik fonksiyonunun maksimum değeri -446.0793 olarak hesaplanmıştır. Olabilirlik oranı testine göre $H_0 : \beta_1 = \beta_2 = \dots = \beta_j = 0$ şeklinde kurulan hipotezi reddetmekle yapılacak olan hata

0 olarak bulunmuştur ve bu sonuç doğrultusunda 0,05 anlamlılık seviyesinde en az bir katsayının 0' dan farklı olduğu kararına varılmaktadır.

Çizelge 6.18 Sıralı probit modelde parametre ve limit değerleri tahminleri

Dependent Variable: Y

Method: ML - Ordered Probit

Sample: 1 337

Included observations: 337

Number of ordered indicator values: 5

Convergence achieved after 5 iterations

Covariance matrix computed using second derivatives

	Coefficient	Std. Error	z-Statistic	Prob.
X1	0.014290	0.004470	3.196757	0.0014
X2	0.265385	0.024753	10.72145	0.0000
X3	0.310604	0.217068	1.430908	0.1525
Limit Points				
LIMIT_1:C(4)	2.426885	0.404375	6.001568	0.0000
LIMIT_2:C(5)	3.331937	0.409303	8.140520	0.0000
LIMIT_3:C(6)	4.150456	0.423407	9.802529	0.0000
LIMIT_4:C(7)	4.581432	0.433686	10.56395	0.0000
Akaike info criterion	2.688898	Schwarz criterion	2.768247	
Log likelihood	-446.0793	Hannan-Quinn criter.	2.720525	
Restr. Log likelihood	-509.8441	Avg. Log likelihood	-1.323677	
LR statistic (3 df)	127.5296	LR index (Pseudo-R2)	0.125067	
Probability(LR stat)	0.000000			

Bağımlı değişkenin 5 grubu olduğundan her bir gözlem için 5 olasılık değeri hesaplanmaktadır ve probit modelde bu hesaplamalar normal dağılıma ilişkin birikimli dağılım fonksiyonu kullanılarak yapılmaktadır. Modelimizde her gözlem için gruplara ait olma olasılıkları aşağıda ifade edildiği gibi bulunmaktadır.

$$I_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i}$$

$$I_i = 0.014290 X_{1i} + 0.265385 X_{2i} + 0.310604 X_{3i}$$

$$P(Y_i = 0) = \Phi(2.426885 - I_i)$$

$$P(Y_i = 1) = \Phi(3.331937 - I_i) - \Phi(2.426885 - I_i)$$

$$P(Y_i = 2) = \Phi(4.150456 - I_i) - \Phi(3.331937 - I_i)$$

$$P(Y_i = 3) = \Phi(4.581432 - I_i) - \Phi(4.150456 - I_i)$$

$$P(Y_i = 4) = 1 - \Phi(4.581432 - I_i)$$

Olasılıklar hesaplandıktan sonra, gözlemin, en yüksek olasılık değerine sahip gruba ait olduğu şeklinde bir karara varılabilmektedir. Ek 3 incelendiğinde her birey için gruplara ait olma olasılık değerleri görülmektedir. Gözlem numarası 1 olan bireye ilişkin olasılık değerleri incelendiğinde, en yüksek olasılık değerinin, 1 olarak tanımlanan Alt.2' ye ait olduğu ve bireyin 0,32478 olasılıkla mavi yaka olduğu söylenir.

Çizelge 6.19 Sıralı probit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları

Dependent Variable: Y
Method: ML - Ordered Probit

Sample: 1 337

Included observations: 337

Prediction table for ordered dependent variable

Value	Count	Count of obs with Max Prob	Error	Sum of all Probabilities	Error
0	31	20	11	32.361	-1.361
1	69	34	35	62.893	6.107
2	84	131	-47	84.498	-0.498
3	41	0	41	44.552	-3.552
4	112	152	-40	112.695	-0.695

Yukarıda yer alan çizelgede bağımlı değişkenin her bir grubuna ilişkin gerçekleşen ve beklenen gözlem sayıları ile bu sayılar arasındaki farkı gösteren gözlem hata değeri gösterilmektedir. Çizelgeye göre gözlenen değeri 0 olan bağımlı değişken sayısı 31 ve modelin tahmin edilmesiyle grubu 0 olarak belirlenen gözlem sayısı da 20 olmaktadır. Bağımlı değişkenin 0 olma olasılıkların toplamı 32.361 olarak hesaplanmıştır ve gözlem sayısından bu olasılık değerinin çıkarılmasıyla olasılık hata değeri bulunmuştur. Logit model ile karşılaştırıldığında 2, 3 ve 4 olarak belirlenen grupların tahmin edilen gözlem sayılarının eşit olduğu görülmektedir. Her iki model sonucuna göre tahmin edilen değeri 3 yani beyaz yakalı olan gözlem sayısı 0 olarak bulunmuştur. Diğer bir ifadeyle her bir gözlem için gruplara ait olma olasılıkları hesaplandığında, hiçbir gözlemde bağımlı değişkenin 3 olma olasılığı maksimum değer olamamıştır ve beyaz yakalı olarak tahmin edilen gözlem sayısı da 0 olarak bulunmuştur.

Katsayıların anlamlılıkları 0,05 anlamlılık seviyesinde test edildiğinde ırk değişkenine ilişkin katsayının anlamsız olduğu sonucuna varıldığından bu değişken modelde yer almadan 2 bağımsız değişkenle model yeniden kurulmak istenirse çizelge 6.20' de yer alan sonuçlar elde edilecektir. Elde edilen katsayı ve limit değerleri birikimli normal dağılım fonksiyonunda kullanılarak her bir gözlem için gruplara ait olma olasılıkları hesaplanmaktadır.

Çizelge 6.20 Sorunlu değişkenin çıkartılmasıyla elde edilen sıralı probit modelde parametre ve limit değerleri

Dependent Variable: Y

Method: ML - Ordered Probit

Sample: 1 337

Included observations: 337

Number of ordered indicator values: 5

Convergence achieved after 4 iterations

Covariance matrix computed using second derivatives

	Coefficient	Std. Error	z-Statistic	Prob.
X1	0.013816	0.004454	3.102117	0.0019
X2	0.264994	0.024747	10.70795	0.0000
Limit Points				
LIMIT_1:C(3)	2.132203	0.346956	6.145452	0.0000
LIMIT_2:C(4)	3.034255	0.351509	8.632094	0.0000
LIMIT_3:C(5)	3.849942	0.366464	10.50564	0.0000
LIMIT_4:C(6)	4.278717	0.377372	11.33819	0.0000
Akaike info criterion	2.689016	Schwarz criterion	2.757030	
Log likelihood	-447.0992	Hannan-Quinn criter.	2.716125	
Restr. Log likelihood	-509.8441	Avg. log likelihood	-1.326704	
LR statistic (2 df)	125.4897	LR index (Pseudo-R2)	0.123067	
Probability(LR stat)	0.000000			

Kurulan bu modele göre 3 olarak tanımlanmış gruba ilişkin tahmin edilen gözlem sayısının 3 bağımsız değişkenin yer aldığı ilk modelle aynı olduğunu, diğer gruplarda da sayıların birbirine çok yakın olarak bulunduğunu aşağıda yer alan çizelgeden görmekteyiz.

Çizelge 6.21 Sorunlu değişkenin çıkartılmasıyla elde edilen sıralı probit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları

Dependent Variable: Y

Method: ML - Ordered Probit

Sample: 1 337

Included observations: 337

Prediction table for ordered dependent variable

Value	Count	Count of obs with Max Prob	Error	Sum of all Probabilities	Error
0	31	21	10	32.340	-1.340
1	69	30	39	63.166	5.834
2	84	130	-46	84.410	-0.410
3	41	0	41	44.332	-3.332
4	112	156	-44	112.753	-0.753

Bağımlı değişkenin gözlem sayıları her grupta farklılık göstermektedir ve dolayısıyla gözlem sayılarına ilişkin yüzdelik dilimler de birbirinden oldukça farklıdır. Bu farklılığı gidermek açısından gözlem sayıları eşit olan, rasgele örnekleme yöntemiyle (random sampling) 3 örneklem ele alınarak analizler yapılacaktır. Örneklem elde edilmesinde SPSS 11.5 paket programından yararlanılmaktadır.

Her bir örneklemdaki gözlem sayısı 155 ve her gruptaki gözlem sayısı da 31 olarak alınmaktadır. Gruplardaki gözlem sayıları eşit olarak alınıp tahmin edilen frekansların ve olasılık hata değerlerinin dengeli diğer anlamda birbirlerine yakın olup olmadıkları sıralı logit ve sıralı probit model uygulamaları esas alınarak incelenmektedir.

Çizelge 6.22' de bağımlı değişkenin her bir grubuna ilişkin gerçekleşen ve beklenen gözlem sayıları ile bu sayılar arasındaki farkı gösteren gözlem hata değeri gösterilmektedir. Gruplara ait olma olasılıkları toplandığında her bir grup için bu toplamaların birbirlerine yakın oldukları görülmektedir. Çizelge 6.15 ve çizelge 6.19 incelendiğinde ise her bir gruba ait olasılık değerleri toplamının da gözlem sayıları gibi dengesiz oldukları yani birbirlerinden farklı oldukları sonucuna varılmaktadır.

Çizelge 6.22 Sıralı logit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları

Dependent Variable: Y
Method: ML - Ordered Logit

Sample: 1 155

Included observations: 155

Prediction table for ordered dependent variable

Value	Count	Count of obs with Max Prob	Error	Sum of all Probabilities	Error
0	31	27	4	29.775	1.225
1	31	29	2	29.038	1.962
2	31	31	0	31.567	-0.567
3	31	33	-2	33.368	-2.368
4	31	35	-4	31.252	-0.252

Aşağıda yer alan çizelgede ise sıralı probit model için çözümlemeler yapılmakta ve gruplar için toplam olasılık değerlerinin birbirlerine yakın oldukları görülmektedir. Ayrıca hem logit hem de probit modelde elde edilen toplam değerlerin de birbirlerine yakın oldukları sonucuna da varılabilmektedir.

Çizelge 6.23 Sıralı probit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları

Dependent Variable: Y
Method: ML - Ordered Probit

Sample: 1 155

Included observations: 155

Prediction table for ordered dependent variable

Value	Count	Count of obs with Max Prob	Error	Sum of all Probabilities	Error
0	31	45	-14	30.710	0.290
1	31	11	20	28.954	2.046
2	31	27	4	31.601	-0.601
3	31	35	-4	33.312	-2.312
4	31	37	-6	30.423	0.577

Çizelge 6.24' de ise ele alınan 2. örneklem için yapılan sıralı logit model çözümlemesi sonucunda elde edilen frekans ve hata değerleri gösterilmektedir. İlk örneklemdekine benzer olarak gruplara ilişkin toplam olasılık değerlerinin birbirlerine yakın oldukları görülmektedir. Göze çarpan diğer bir nokta ise ilk örneklemde 2 olarak tanımlanan grubun hata değeri mutlak değer olarak en düşük değerde iken 2. örneklemde bu gruba ilişkin hata değeri en yüksek seviyededir.

Çizelge 6.24 Sıralı logit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları

Dependent Variable: Y
Method: ML - Ordered Logit

Sample: 1 155

Included observations: 155

Prediction table for ordered dependent variable

Value	Count	Count of obs with Max Prob	Error	Sum of all Probabilities	Error
0	31	28	3	29.257	1.743
1	31	21	10	29.724	1.276
2	31	44	-13	32.381	-1.381
3	31	30	1	32.705	-1.705
4	31	32	-1	30.933	0.067

Çizelge 6.25 incelendiğinde gruplar için olasılık değerleri toplamının logit modelde olduğu gibi birbirlerine yakın oldukları görülmektedir. 1. ve 2. örneklemelerin sıralı probit model uygulamalarına bakıldığında mavi yakalı gruba ait hata değerlerinin mutlak olarak diğer gruplardan daha büyük olduğu sonucuna varılmaktadır.

Çizelge 6.25 Sıralı probit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları

Dependent Variable: Y
Method: ML - Ordered Probit

Sample: 1 155

Included observations: 155

Prediction table for ordered dependent variable

Value	Count	Count of obs with Max Prob	Error	Sum of all Probabilities	Error
0	31	39	-8	30.214	0.786
1	31	18	13	29.749	1.251
2	31	36	-5	32.281	-1.281
3	31	27	4	32.259	-1.259
4	31	35	-4	30.497	0.503

Çizelge 6.26' da, ele alınan 3. örneklem için yapılan sıralı logit model çözümlemesi sonucunda elde edilen frekans ve hata değerleri gösterilmektedir. İlk 2 örneklemdekine benzer olarak gruplara ilişkin toplam olasılık değerlerinin ve ayrıca hem logit hem de probit modelden elde edilen toplam olasılık değerlerinin de birbirlerine yakın oldukları görülmektedir. 2 olarak tanımlanan esnaf-sanatkar grubuna ilişkin gözlem hata değerlerinin hem 2. hem de 3. örneklemde, mutlak olarak en yüksek seviyede olduğu göze çarpan diğer bir noktadır.

Çizelge 6.26 Sıralı logit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları

Dependent Variable: Y
Method: ML - Ordered Logit

Sample: 1 155

Included observations: 155

Prediction table for ordered dependent variable

Value	Count	Count of obs with Max Prob	Error	Sum of all Probabilities	Error
0	31	30	1	30.873	0.127
1	31	33	-2	28.694	2.306
2	31	24	7	30.452	0.548
3	31	32	-1	33.578	-2.578
4	31	36	-5	31.404	-0.404

Çizelge 6.27 incelendiğinde gruplar için olasılık değerleri toplamının logit modelde olduğu gibi birbirlerine yakın oldukları görülmektedir. Diğer iki örneklemde farklı olarak, 3. örneklemin sıralı probit model sonuçlarına bakıldığında mutlak hata değeri en yüksek olan grubun çırak-stajyer grubu olduğu sonucuna varılmaktadır.

Çizelge 6.27 Sıralı probit modelde bağımlı değişkenin gözlenen ve tahmin edilen frekansları

Dependent Variable: Y
Method: ML - Ordered Probit

Sample: 1 155

Included observations: 155

Prediction table for ordered dependent variable

Value	Count	Count of obs with Max Prob	Error	Sum of all Probabilities	Error
0	31	48	-17	31.152	-0.152
1	31	15	16	28.611	2.389
2	31	20	11	30.727	0.273
3	31	31	0	33.566	-2.566
4	31	41	-10	30.944	0.056

Çizelgelerde elde edilen sonuçların bulunmasında kullanılan parametre ve limit değerleri tahminleri, her örneklem için ayrı ayrı hesaplanmaktadır.

Çizelge 6.28 Sıralı logit modelde parametre ve limit değerleri tahminleri (örneklem 1)

Dependent Variable: Y
Method: ML - Ordered Logit

Sample: 1 155
Included observations: 155
Number of ordered indicator values: 5
Convergence achieved after 5 iterations
Covariance matrix computed using second derivatives

	Coefficient	Std. Error	z-Statistic	Prob.
X1	0.023087	0.010933	2.111583	0.0347
X2	0.402025	0.063173	6.363839	0.0000
X3	0.560031	0.600296	0.932925	0.3509
Limit Points				
LIMIT_1:C(4)	4.366788	1.074998	4.062135	0.0000
LIMIT_2:C(5)	5.488094	1.096998	5.002827	0.0000
LIMIT_3:C(6)	6.498377	1.130138	5.750074	0.0000
LIMIT_4:C(7)	7.764208	1.183063	6.562804	0.0000
Akaike info criterion	3.007538	Schwarz criterion	3.144983	
Log likelihood	-226.0842	Hannan-Quinn criter.	3.063365	
Restr. log likelihood	-249.4629	Avg. log likelihood	-1.458608	
LR statistic (3 df)	46.75741	LR index (Pseudo-R2)	0.093716	
Probability(LR stat)	3.91E-10			

Çizelge 6.29 Sıralı probit modelde parametre ve limit değerleri tahminleri (örneklem 1)

Dependent Variable: Y
Method: ML - Ordered Probit

Sample: 1 155
Included observations: 155
Number of ordered indicator values: 5
Convergence achieved after 4 iterations
Covariance matrix computed using second derivatives

	Coefficient	Std. Error	z-Statistic	Prob.
X1	0.012964	0.006362	2.037686	0.0416
X2	0.242899	0.036540	6.647428	0.0000
X3	0.341889	0.327699	1.043300	0.2968
Limit Points				
LIMIT_1:C(4)	2.662650	0.611419	4.354870	0.0000
LIMIT_2:C(5)	3.317732	0.619697	5.353800	0.0000
LIMIT_3:C(6)	3.929415	0.634694	6.191034	0.0000
LIMIT_4:C(7)	4.682392	0.660230	7.092066	0.0000
Akaike info criterion	3.004807	Schwarz criterion	3.142252	
Log likelihood	-225.8725	Hannan-Quinn criter.	3.060634	
Restr. log likelihood	-249.4629	Avg. log likelihood	-1.457242	
LR statistic (3 df)	47.18065	LR index (Pseudo-R2)	0.094564	
Probability(LR stat)	3.18E-10			

Çizelge 6.30 Sıralı logit modelde parametre ve limit değerleri tahminleri (örneklem 2)

Dependent Variable: Y
Method: ML - Ordered Logit

Sample: 1 155
Included observations: 155
Number of ordered indicator values: 5
Convergence achieved after 5 iterations
Covariance matrix computed using second derivatives

	Coefficient	Std. Error	z-Statistic	Prob.
X1	0.022746	0.011047	2.059068	0.0395
X2	0.389685	0.065295	5.968079	0.0000
X3	0.525361	0.669675	0.784501	0.4327
Limit Points				
LIMIT_1:C(4)	4.085599	1.042112	3.920497	0.0001
LIMIT_2:C(5)	5.219116	1.067168	4.890622	0.0000
LIMIT_3:C(6)	6.236326	1.099868	5.670068	0.0000
LIMIT_4:C(7)	7.471035	1.146527	6.516230	0.0000
Akaike info criterion	3.030454	Schwarz criterion	3.167899	
Log likelihood	-227.8601	Hannan-Quinn criter.	3.086281	
Restr. log likelihood	-249.4629	Avg. log likelihood	-1.470065	
LR statistic (3 df)	43.20546	LR index (Pseudo-R2)	0.086597	
Probability(LR stat)	2.23E-09			

Çizelge 6.31 Sıralı probit modelde parametre ve limit değerleri tahminleri (örneklem 2)

Dependent Variable: Y
Method: ML - Ordered Probit

Sample: 1 155
Included observations: 155
Number of ordered indicator values: 5
Convergence achieved after 5 iterations
Covariance matrix computed using second derivatives

	Coefficient	Std. Error	z-Statistic	Prob.
X1	0.013627	0.006428	2.119910	0.0340
X2	0.223834	0.036649	6.107511	0.0000
X3	0.410377	0.353319	1.161491	0.2454
Limit Points				
LIMIT_1:C(4)	2.457821	0.599393	4.100516	0.0000
LIMIT_2:C(5)	3.115685	0.610003	5.107660	0.0000
LIMIT_3:C(6)	3.724918	0.624959	5.960262	0.0000
LIMIT_4:C(7)	4.439477	0.644382	6.889516	0.0000
Akaike info criterion	3.044848	Schwarz criterion	3.182293	
Log likelihood	-228.9757	Hannan-Quinn criter.	3.100675	
Restr. log likelihood	-249.4629	Avg. log likelihood	-1.477263	
LR statistic (3 df)	40.97435	LR index (Pseudo-R2)	0.082125	
Probability(LR stat)	6.62E-09			

Çizelge 6.32 Sıralı logit modelde parametre ve limit değerleri tahminleri (örneklem 3)

Dependent Variable: Y

Method: ML - Ordered Logit

Sample: 1 155

Included observations: 155

Number of ordered indicator values: 5

Convergence achieved after 5 iterations

Covariance matrix computed using second derivatives

	Coefficient	Std. Error	z-Statistic	Prob.
X1	0.021240	0.010865	1.954857	0.0506
X2	0.450654	0.069863	6.450580	0.0000
X3	1.145907	0.619826	1.848754	0.0645
Limit Points				
LIMIT_1:C(4)	5.528240	1.122689	4.924109	0.0000
LIMIT_2:C(5)	6.654014	1.142793	5.822591	0.0000
LIMIT_3:C(6)	7.662728	1.176959	6.510613	0.0000
LIMIT_4:C(7)	8.987781	1.247399	7.205219	0.0000
Akaike info criterion	2.969495	Schwarz criterion	3.106940	
Log likelihood	-223.1358	Hannan-Quinn criter.	3.025322	
Restr. log likelihood	-249.4629	Avg. log likelihood	-1.439586	
LR statistic (3 df)	52.65408	LR index (Pseudo-R2)	0.105535	
Probability(LR stat)	2.17E-11			

Çizelge 6.33 Sıralı probit modelde parametre ve limit değerleri tahminleri (örneklem 3)

Dependent Variable: Y

Method: ML - Ordered Probit

Sample: 1 155

Included observations: 155

Number of ordered indicator values: 5

Convergence achieved after 5 iterations

Covariance matrix computed using second derivatives

	Coefficient	Std. Error	z-Statistic	Prob.
X1	0.012012	0.006280	1.912571	0.0558
X2	0.258648	0.038835	6.660148	0.0000
X3	0.681864	0.329078	2.072041	0.0383
Limit Points				
LIMIT_1:C(4)	3.176541	0.616697	5.150896	0.0000
LIMIT_2:C(5)	3.832067	0.622919	6.151790	0.0000
LIMIT_3:C(6)	4.437370	0.636754	6.968738	0.0000
LIMIT_4:C(7)	5.207867	0.667784	7.798727	0.0000
Akaike info criterion	2.974605	Schwarz criterion	3.112050	
Log likelihood	-223.5319	Hannan-Quinn criter.	3.030432	
Restr. log likelihood	-249.4629	Avg. log likelihood	-1.442141	
LR statistic (3 df)	51.86205	LR index (Pseudo-R2)	0.103947	
Probability(LR stat)	3.20E-11			

7. SONUÇLAR

Bağımlı bir değişken ile bir veya daha fazla bağımsız değişken arasındaki ilişkiyi açıklama ve değerlendirme ile ilgili olan regresyon analizinde bağımlı değişken sadece kolayca sayısallaştırılabilen bağımsız değişkenlerin değil, bazı durumlarda, özünde nitel olan bağımsız değişkenlerin de etkisi altında bulunabilmektedir. Nitel değişkenlerin regresyon analizinde kullanılması sırasında bu değişkenler 0, 1, 2 gibi değerler olarak nicel hale getirilmekte ve oluşturulan bu yapay değişkenler gölge değişken olarak adlandırılmaktadır. Bağımsız değişkenlerin nitel özellik taşıdığı ve dolayısıyla gölge değişkenler ile temsil edildiği regresyon modellerine uygulamada sıkça rastlamak mümkündür.

İki veya daha fazla değer alan gölge değişkenler bağımlı değişken olarak da regresyon modellerinde yer alabilmektedir. Bağımlı değişkenin evet-hayır, satın alma-almama, borsa ile ilgilenip ilgilenmeme gibi tercih veya kararın belirtildiği bir değişken olduğu ve dolayısıyla iki değer aldığı modeller ikili tercih modelleri olarak adlandırılmaktadır. Bağımlı değişken ikiden fazla değer de alabilmektedir. Bağımlı değişkenin evet-hayır-belki, marka tercihi, ziyaret edilmek istenen ülke tercihi, çok zayıf-zayıf-orta-iyi-çok iyi gibi ikiden fazla değer aldığı modeller de çoklu tercih modelleri olarak tanımlanmaktadır. Her iki tip modelde de temel amaç tercih olasılığının belirlenmesidir. Karar sürecinin olasılıklı yapısı hakkında geliştirilen varsayımlara dayalı olarak çeşitli tercih modelleri oluşturulmaktadır.

İkili tercih modelleri arasından ilk olarak incelenen doğrusal olasılık modelinde bir karar biriminin belli bir tercihte bulunma olasılığının, onun sahip olduğu özelliklerin katsayılar itibarıyla doğrusal bir fonksiyonu olduğu varsayılmaktadır. Bu modelde bağımlı değişken bağımsız değişkenlerin doğrusal bir fonksiyonu olarak gösterilmektedir. DOM, hata teriminin normal dağılmaması ve değişen varyanslı olması, olasılık değerlerinin 0-1 aralığı dışına düşebilmesi, bağımsız değişkendeki değişmelerin olasılık üzerinde sabit bir etkiye sahip olması, belirlilik katsayısının genellikle küçük değerde olması ve uyum iyiliğinin iyi bir ölçüsü olarak kullanılamaması sorunlarını taşıdığından pek tercih edilmeyen bir modeldir.

Doğrusal olasılık modelinde değinilen sorunları gidermek için birikimli dağılım fonksiyonları yardımıyla kurulan modeller geliştirilmiştir. S biçimli BDF' leri esas alan bu modellerden, normal BDF' den elde edilen probit model ve lojistik BDF' den elde edilen ise logit model olarak bilinmektedir.

Logit modelde olasılık değerleri ile bağımsız değişkenler arasında doğrusal olmayan bir ilişki vardır ve bundan hareketle olasılığın bağımsız değişkene göre kısmi türevi alındığında marjinal değişimin yalnız regresyon katsayısına değil, aynı zamanda değişimin ölçüldüğü

olasılığın düzeyine de bağlı olduğu sonucuna varılmaktadır. Bahis oranı, bir olayın gerçekleşme olasılığının gerçekleşmeme olasılığına bölünmesiyle elde edilmekte ve bahis oranının logaritması alındığında logit olarak adlandırılan ifadeye ulaşılmaktadır. Logit hem bağımsız değişken hem de regresyon katsayılarına göre doğrusaldır. Logit modelde parametre tahminlerinin elde edilmesi için logit değerlerinin bulunması gerekmektedir. Gruplandırılmamış veriler söz konusu olduğunda logit değerlerinin anlamsız olmasından dolayı en küçük kareler yöntemi kullanılamamakta ve parametre tahminleri için EYO yöntemi kullanılmaktadır. EYO tahmin yöntemi için benzerlik fonksiyonu yazılmakta ve bu fonksiyonun 0' a eşitlenen kısmî türevleri çözülerek parametre tahminleri elde edilmektedir.

Birikimli normal dağılım fonksiyonu üzerine oluşturulan probit modelin anlaşılması açısından fayda endeksi ve endeksin eşik değeri incelendi. i . bireyin j . alternatiften elde ettiği fayda olarak belirtilen U_{ij} , sistematik (V_{ij}) ve tesadüfi (ε_{ij}) olmak üzere 2 kısımdan oluşmaktadır. Sistematik kısım hem alternatiflerin hem de bireylerin özelliklerinin katsayılar itibariyle doğrusal bir fonksiyonu şeklinde ifade edilmektedir. Tesadüfi kısım gözlenemediğinden dolayı yapılacak tercih ile ilgili olasılıklı ifadelerde bulunulur.

Endeks değeri ile eşik değeri arasındaki ilişkiden yararlanılarak olasılık formülleri ve benzerlik fonksiyonu yazılmaktadır. Benzerlik fonksiyonunu maksimum yapacak parametre değerlerini elde etmek için de nümerik iterasyon yöntemleri kullanılmaktadır.

Probit modelde, bağımsız değişkendeki bir değişimin olasılık üzerine etkisi doğrusal olasılık modelindeki gibi sabit değildir. Bağımsız değişkenin marjinal etkisinin büyüklüğü, standart normal OYF ve ilgili parametre değerine bağlıdır. OYF her zaman pozitif olduğundan dolayı marjinal etki ilgili parametrenin işaretine bağlı olmaktadır. Bu parametrenin değeri negatif olduğunda bağımsız değişkendeki bir artış olasılık değerini azaltmakta, pozitif olduğunda ise bağımsız değişkendeki bir artış olasılık değerini arttırmaktadır.

Logit ve probit modeller benzer sonuçlar vermesine rağmen, iki modelin parametre değerleri doğrudan karşılaştırılamamaktadır. Karşılaştırma için Amemiya' nın önermesinden faydalanılabilir. Normal dağılımın integral hesaplamaları içermesinden dolayı, logit model probit modelden daha sık olarak kullanılmaktadır.

Kategorik değişkenler sıralı, sıralı olmayan ve ardışık değişkenler olarak sınıflandırılmakta ve bu değişkenler kullanılarak çeşitli çoklu tercih modelleri oluşturulmaktadır.

Sıralı olmayan tercih modelleri 5 ana başlık altında incelendi. İlk olarak ele alınan çok durumlu DOM iki alternatif için söz konusu olan doğrusal olasılık modelinin üç veya daha

fazla karşılıklı dışlayıcı alternatif için genelleştirilmesinden oluşmaktadır. Çok durumlu probit model, rasgele fayda modelindeki hata terimlerinin çok değişkenli normal dağılıma uyduğu varsayımı üzerine kurulmaktadır. Çok durumlu logit modelde ise hata terimleri lojistik dağılıma uymaktadır. Hata terimlerinin birbirinden bağımsız ve Weibull dağılımına uymaları durumunda her bir alternatife ilişkin seçim olasılıkları koşullu logit model yardımıyla elde edilmektedir. Yuvalanmış logit modelde alternatifler alt gruplara ayrılmakta ve tercih yapılırken önce grup seçimi yapılmakta ve daha sonra bu gruptaki belli bir alternatif tercih edilmektedir. Yuvalanmış modelin katsayı tahminlerinde sınırlı bilgiye dayalı iki aşamalı en çok benzerlik ve tam bilgiye dayalı en çok benzerlik yöntemleri kullanılmaktadır.

Nitel bağımlı değişkenin sıralı olması durumunda sıralı tercih modelleri kullanılmaktadır. Bağımsız değişkenlere ilişkin katsayı değerleri ve limit değerler kullanılarak bağımlı değişkenin herhangi bir gruba ait olma olasılığı hem probit hem de logit model için hesaplanmaktadır. Sıralı tercih modellerinde regresyon katsayılarının yorumlanmasında dikkatli olunmalıdır. Bu katsayıların işareti iki tercih arasında yer alma olasılığındaki değişimin yönünü belirtmektedir.

Çok alternatifin söz konusu olduğu ve alternatifler arasında tercih yapılırken kararların ardışık olarak alındığı durumlarda çok durumlu modellerin tahmini daha az sayıda alternatifli modellerin peş peşe tahminine dönüşmekte ve bu durum yapılması gereken hesaplamaları azaltmaktadır.

Uygulama aşamasında bireyler menial, mavi yaka, esnaf-sanatkar, beyaz yaka ve profesyonel olmak üzere 5 grupta incelenmektedir. Grup değerleri bağımlı değişken olarak kullanılmakta ve bireylerin gruplar arasında geçiş yapabilmesi mümkün olduğundan, örneğin mavi yakalı olan bir bireyin başarılı bulunan bir proje sonrasında beyaz yakalı olarak terfi ettirilmesi, bağımlı değişken sıralı kategorik değişken özelliğindedir. Bağımlı değişkeni etkileyen bağımsız değişkenler piyasa deneyim yılı, resmi kurumlarca onaylı eğitim yılı ve ırk olmak üzere 3 adettir.

İlk olarak, bireyin iş statüsünü belirten bağımlı değişken kendi arasında 2 alt gruba ayrıldı ve ikili DOM, probit ve logit model sonuçları elde edildi. 0,1 ve 2 olarak tanımlanan gruplarda beden gücünün daha çok, diğer geri kalan iki grupta ise beyin gücünün daha çok kullanıldığı göz önüne alınarak ikili model uygulamaları için grup tanımlamaları yapıldı.

En küçük kareler yönteminin uygulandığı ikili doğrusal olasılık modeli incelendiğinde, bireyin eğitim yılı ve piyasa deneyim yılı arttıkça 2. gruba (profesyonel ve beyaz yakalıların toplandığı alt grup) ait olma olasılığı artmaktadır şeklinde bir sonuca ulaşıldı. Bu modeldeki

sorunlu yani olasılık değeri 0-1 aralığı dışına çıkan gözlem sayısı 32 olarak bulundu. İkili logit ve probit modellerde Newton-Raphson nümerik iterasyon yöntemi kullanılarak parametrelerin EYO tahmin edicileri bulunmaktadır. Bağımsız değişkenlerin tümü olayın ($P(Y=1)$, 1.alt gruba ait olma olasılığı) gerçekleşme olasılığı ile ters yönlü bir ilişki içerisindedir. İkili modellerde modelin doğruluğunun belirlenmesi için tespit edilen bir eşik değere göre gruplandırmalar yapıldı ve bu gruplandırmalar neticesinde doğru ve yanlış tahmin edilen gözlem sayıları bulundu. Logit modelde 0,7567 olarak hesaplanan doğru tahmin oranı probit modelde 0,7507 düzeyindedir. Gerçekleşen ve beklenen gözlem frekanslarının karşılaştırıldığı Hosmer-Lemeshow test sonuçlarına göre her iki modelde de " H_0 : gözlenen ve beklenen frekanslar arasında fark yoktur." şeklinde kurulan hipotezi reddetmek için yeterli bir neden bulunamadı.

İkili modellerden daha karmaşık yapıda olan çoklu tercih modellerinde bireyler ikiden fazla tercihle karşı karşıyadır ve dolayısıyla bağımlı değişken ikiden fazla değer almaktadır. Bu çalışmada bireyler çalışma statülerine göre 5 gruba ayrılmakta ve grup değerleri bağımlı değişken olarak modelde yer almaktadır.

Bağımlı değişkenin her grubuna ait gözlem sayısını gösteren çizelge incelendiğinde gözlem sayısı en çok olan çalışan grubunun profesyonel, en az olanın ise çırak-stajyer grubu olduğu görülmektedir. Sıralı logit ve probit modelde en çok benzerlik yöntemi kullanılarak parametre tahminleri yapılmakta ve bu tahmin sürecinde Newton-Raphson nümerik iterasyon yöntemi kullanılarak 5 iterasyon sonucunda parametrelerin en çok benzerlik tahmin edicileri bulunmaktadır. Modellerde bağımsız değişken sayısı kadar parametre değeri (3 adet), grup sayısının (5) bir eksiği kadar (4) limit değer hesaplanmakta ve bu değerler yardımıyla da her bir gözlem için gruplara ait olma olasılıkları elde edilmektedir. Bağımlı değişkenin 5 grubu olduğundan her bir gözlem için 5 olasılık değeri hesaplanmakta ve olasılıklar hesaplandıktan sonra, gözlemin en yüksek olasılık değerine sahip gruba ait olduğu şeklinde bir karara varılabilmektedir. İki modelde de gözlemler için maksimum olasılık değerleri incelendiğinde, beyaz yaka olarak tahmin edilen gözlem sayısının 0 olduğunu görmekteyiz. Yani her iki modelde de bireylerin hiçbiri için beyaz yaka olma olasılık değeri maksimum seviyeye ulaşmamaktadır. Bağımlı değişkenin her bir grubuna ilişkin gerçekleşen ve beklenen gözlem sayıları arasındaki farkı gösteren gözlem hata değerleri incelendiğinde 2, 3, 4 olarak tanımlanan gruplara ilişkin değerlerin iki modelde de aynı olduğu görülmektedir. Diğer bir ifadeyle hem sıralı probit hem de sıralı logit modelde esnaf-sanatkar, beyaz yaka ve profesyonel olarak tahmin edilen birey sayısı eşittir.

Bağımlı değişkenin gözlem sayıları her grupta farklılık göstermektedir ve dolayısıyla gözlem sayılarına ilişkin yüzdelik dilimler de birbirinden oldukça farklıdır. Bu farklılığı gidermek açısından gözlem sayıları eşit olan, rasgele örnekleme yöntemiyle (random sampling) 3 örneklem ele alınarak analizler yapıldı.

Gruplardaki gözlem sayıları eşit olarak alınıp tahmin edilen frekansların ve olasılık hata değerlerinin dengeli diğer anlamda birbirlerine yakın olup olmadıkları sıralı logit ve sıralı probit model uygulamaları esas alınarak incelendi.

Elde edilen 1.örnekleme hem sıralı logit hem de sıralı probit sonuçlarına bakıldığında, gruplar için toplam olasılık değerlerinin birbirlerine yakın oldukları görülmektedir. Diğer bir ifadeyle her bir gruba ait olasılık değerleri toplamının, gözlem sayıları gibi dengeli oldukları sonucuna varıldı. Diğer 2 örnekleme de aynı sonuçlara varıldı.

Sıralı logit model sonuçlarına göre, ilk örnekleme esnaf-sanatkar grubunun hata değeri mutlak değer olarak en düşük değerde iken 2. örnekleme bu gruba ilişkin hata değeri en yüksek seviyededir. Ayrıca, bu gruba ilişkin gözlem hata değerlerinin hem 2. hem de 3. örneklemelerde, mutlak olarak en yüksek değerde olduğu göze çarpmaktadır.

1. ve 2. örneklemelerin sıralı probit model uygulamalarına bakıldığında mavi yakalı gruba ait hata değerlerinin mutlak olarak diğer gruplardan daha büyük olduğu sonucuna varıldı. İlk iki örneklemden farklı olarak, 3. örneklemin sıralı probit model sonuçlarına göre mutlak hata değeri en yüksek olan grubun çırak-stajyer grubu olduğu sonucuna varıldı.

337 adet gözlemin yer aldığı ilk incelemede gruplara ilişkin gözlem sayılarının birbirinden çok farklı oldukları görülmektedir. Bu gözlemlerle sıralı logit ve sıralı probit analizleri yapıldığında, gruplara ilişkin toplam olasılık değerlerinin de birbirinden oldukça farklı olduğu sonucuna ulaşıldı. Rasgele örnekleme yoluyla elde edilen ve gözlem sayıları eşit olan 3 adet örnekleme ilişkin sıralı logit ve sıralı probit sonuçlarına bakıldığında ise, gruplar için toplam olasılık değerlerinin birbirlerine yakın oldukları sonucuna ulaşıldı. Diğer bir ifadeyle her bir gruba ait olasılık değerleri toplamının, gözlem sayıları gibi dengeli oldukları gösterildi.

KAYNAKLAR

- Akdeniz, F., (1998), Olasılık ve İstatistik, Baki Kitabevi, İstanbul.
- Akın, F., (2002), Ekonometri, Ekin Kitabevi, Bursa.
- Akın, F., (2002), Kalitatif Tercih Modelleri Analizi, Ekin Kitabevi, Bursa.
- Akın, F., (2002), "Kadınların İş Gücüne Katılımı ve İşteki Durum Tercihinin Nested Logit Model ile Analizi", Uluslararası Ekonomi Kongresi, 11-14 Eylül 2002, ODTÜ.
- Aldrich, J. ve Nelson, F., (1984), Linear Probability, Logit and Probit Models, Beverly Hills, California; Sage Publications.
- Alpar, R., (2003), Uygulamalı Çok Değişkenli İstatistiksel Yöntemlere Giriş 1, Nobel Yayın Dağıtım, Ankara.
- Amemiya, T., (1985), Advanced Econometrics, Harvard University Press, Cambridge.
- Chow, C. G., (1983), Econometrics, McGraw-Hill International Editing, New York.
- Cox, D. R., (1970), Analysis Of Binary Data, Methuen, London.
- Ertek, T., (1996), Ekonometriye Giriş, Beta Basım Yayın Dağıtım A.Ş., İstanbul.
- E-Views 4.1 User' s Guide, Quantitative Micro Software.
- Frank, R. C., (1971), Statistics and Econometrics, Holt, Rinehart and Winston Inc., Austin, Texas.
- Genceli, M., (2001), Ekonometri ve İstatistik İlkeleri, Filiz Kitabevi, İstanbul.
- Greene, W. H., (1997), Econometric Analysis (3rd Edition), Prentice Hall Inc., New Jersey.
- Greene, W. H., (2000), Econometric Analysis (4th Edition), Prentice Hall Inc., New Jersey.
- Griffiths, W. E., Hill, R. C. ve Judge, G. G., (1993), Learning and Practicing Econometrics, John Wiley & Sons Ltd., New York.
- Griffiths, W. E., Hill, R. C. ve Judge, G. G., (1997), Undergraduate Econometrics, John Wiley & Sons Ltd., New York.
- Gujarati, D. N., (1999), Essentials of Econometrics, McGraw-Hill International Editions, New York.
- Gujarati, D. N., (2001), Temel Ekonometri, (Çev., Ü. Şenesen ve G. G. Şenesen), Literatür Yayıncılık, İstanbul.
- Güriş, S. ve Çağlayan, E., (2000), Ekonometri: Temel Kavramlar, Der Yayınları, İstanbul.
- Hausman, J. A. ve Wise, D. A., (1978), "A Conditional Probit Model for Qualitative Choice Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences", *Econometrica*, 46: 403-426.
- Intriligator, M. D., Bodkin G. R. ve Hsiao, C., (1996), Econometric Models, Techniques and Applications (2nd Edition), Prentice-Hall International Inc., New Jersey.
- Kalkan, E., (2004), Yatay Birleşmelerin İncelenmesinde İktisadi Tekniklerin Kullanılması, Rekabet Kurumu Uzmanlık Tezi, Ankara.
- Kennedy, P., (1996), A Guide to Econometrics (3rd Ed.), Blackwell Publishers Ltd., Oxford.
- Kılıçbay, A., (1965), Ekonometri, İstanbul Üniversitesi Yayınları (No:1109), İstanbul.

- Klein, I. ve Mitnik, S., (2003), *Contributions to Modern Econometrics - From Data Analysis to Economic Policy*, Kluwer Academic Publishers, Dordrecht – Boston – London.
- Kleinbaum, D. G., Kupper, L. L., Muller, E. K. ve Nizam, A., (1998), *Applied Regression Analysis and Other Multivariable Methods*, Duxbury Press, Pacific Grove.
- Koop, G., (2003), *Bayesian Econometrics*, John Wiley & Sons Ltd., England.
- Koutsoyiannis, A., (1989), *Ekonometri Kuramı, Ekonometri Yöntemlerinin Tanıtımına Giriş* (Çev., Ü. Şenesen ve G. G. Şenesen), Verso Yayıncılık, Ankara.
- Lerman, R. S. ve Akiva, B. M., (1985), *Discrete Choice Analysis – Theory and Application to Travel Demand*, The MIT Press, London.
- Maddala, G. S., (1983), *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge.
- Neter, J., Kutner, H. M., Nachtsheim, C. J. ve Wasserman, W., (1996), *Applied Linear Statistical Models*, McGraw Hill International Editions, New York.
- Özdamar, K., (2002), *Paket Programlar ile İstatistiksel Veri Analizi 1*, Kaan Kitabevi, Eskişehir.
- Peracchi, F., (2001), *Econometrics*, John Wiley & Sons Ltd., New York.
- Pindyck, S. R. ve Rubinfeld, D. L., (1991), *Econometric Models and Economic Forecasts*, McGraw-Hill International Editions, New York.
- Ramanathan, R., (2002), *Introductory Econometrics with Applications*, Thomson Learning and South – Western Publishing, California.
- Ruud, A. P., (2000), *An Introduction to Classical Econometrics Theory*, Oxford University Press., England.
- Salvatore, D., (1976), *Schaum's Outline Series – Theory and Problems of Statistics and Econometrics*, McGraw Hill Book Company, New York.
- Thomas, R. L., (1997), *Modern Econometrics – An Introduction*, Addison – Wesley Longman Limited, Essex.
- Verbeek, M., (2000), *A Guide to Modern Econometrics*, John Wiley & Sons Ltd., Chichester.
- Watsham, T. J. ve Parromore, K., (1997), *Quantitative Methods In Finance* (1st Ed.), International Thomson Business Pres.

İNTERNET KAYNAKLARI

- [1] <http://www.marketingpower.com/mg-dictionary-view2121.php>
- [2] <http://rosowwww.epfl.ch/mbi/papers/discretechoice>
- [3] <http://www.rff.org/Documents/RFF-DP-03-38.pdf>
- [4] http://www.nyu.edu/classes/nbeck/q2/disc_choice7.nagler.pdf
- [5] <http://www.utdt.edu/~mrozada/mecon1/discretemodels.pdf>
- [6] <http://www.econ.umn.edu/~kortum/courses/fall03/lecture1k.pdf>
- [7] <http://www.et.bs.ehu.es/docpub/manuales/limdep/chap24.pdf>
- [8] <http://gulliver.trb.org/publications/nchrp/cd-22/v2chapter5.html>
- [9] <http://eco.ieu.edu.tr/courses/econ324/LECTURE/Exercise9.pdf>
- [10] http://econ.la.psu.edu/~hbierens/ML_LOGIT.PDF

VERİ KAYNAĞI

<http://www.stat.psu.edu>

<http://joni.soc.surrey.ac.uk/~scs1ps/SOCM20/>

EKLER

Ek 1 Veri Seti

Gözlem	Y	X1	X2	X3	Gözlem	Y	X1	X2	X3	Sayfa
Ek 1	0				Veri Seti.....					127
Ek 2	0				İkili Logit Modelde Gözlem, Tahmin ve Kalıntı Değerleri					131
Ek 3	0				Sıralı Probit Modelde Alternatiflere İlişkin Olasılıklar.....					135
1	0	10	12	1	45	1	30	12	1	
2	0	24	14	0	46	1	17	23	1	
3	0	54	15	1	47	1	3	48	1	
4	0	0	14	0	48	1	11	23	1	
5	0	15	14	1	49	1	8	14	1	
6	0	2	12	1	50	1	17	34	1	
7	0	3	12	1	51	1	4	11	1	
8	0	6	12	1	52	1	20	9	1	
9	0	50	11	1	53	1	32	15	1	
10	0	2	12	1	54	1	29	9	1	
11	0	22	12	1	55	1	5	13	1	
12	0	4	12	1	56	1	11	20	1	
13	0	22	0	1	57	1	5	13	1	
14	0	40	5	1	58	1	12	12	1	
15	0	8	12	1	59	1	25	9	1	
16	0	3	46	1	60	1	35	8	1	
17	0	7	0	1	61	1	17	12	1	
18	0	27	15	0	62	1	43	6	1	
19	0	52	5	0	63	1	8	17	1	
20	0	3	12	0	64	1	9	5	1	
21	0	30	12	1	65	1	52	14	1	
22	0	8	14	1	66	1	49	8	1	
23	0	34	21	1	67	1	6	12	1	
24	0	14	12	1	68	1	32	6	1	
25	0	0	8	1	69	1	51	13	1	
26	0	2	14	1	70	1	39	10	1	
27	0	5	13	1	71	1	37	12	1	
28	0	32	12	1	72	1	10	12	1	
29	1	10	12	1	73	1	2	12	1	
30	1	45	12	1	74	1	5	12	1	
31	1	32	12	1	75	1	45	12	1	
32	1	25	12	1	76	1	12	12	1	
33	1	3	12	1	77	1	9	12	1	
34	1	12	15	1	78	1	3	12	1	
35	1	11	15	1	79	1	3	13	1	
36	1	13	5	1	80	1	6	12	1	
37	1	15	12	1	81	1	36	3	1	
38	1	25	10	1	82	1	19	8	1	
39	1	1	5	1	83	1	41	7	0	
40	1	6	10	1	84	1	32	14	1	
41	1	2	12	1	85	1	5	9	1	

Ek 1 Veri Seti

Gözlem	Y	X1	X2	X3	Gözlem	Y	X1	X2	X3
1	0	3	11	1	45	1	23	8	1
2	0	14	12	1	46	1	8	12	1
3	0	44	12	1	47	1	29	3	1
4	0	18	12	1	48	1	30	13	1
5	0	24	14	0	49	1	17	12	1
6	0	38	13	1	50	1	9	12	1
7	0	8	14	0	51	1	11	15	1
8	0	19	14	1	52	1	9	14	1
9	0	8	12	1	53	1	17	10	1
10	0	3	12	1	54	1	4	11	1
11	0	6	12	1	55	1	30	9	1
12	0	40	11	1	56	1	22	16	0
13	0	2	12	1	57	1	29	10	1
14	0	22	12	1	58	1	6	10	1
15	0	4	12	1	59	1	11	16	1
16	0	22	9	1	60	1	5	12	1
17	0	39	6	1	61	1	12	12	1
18	0	3	12	1	62	1	26	8	1
19	0	3	10	1	63	1	35	8	1
20	0	7	9	1	64	1	17	12	1
21	0	27	15	0	65	1	46	6	1
22	0	52	6	0	66	1	6	11	1
23	0	3	12	0	67	1	37	6	1
24	0	35	12	1	68	1	32	11	1
25	0	8	14	1	69	1	43	8	1
26	0	34	11	1	70	1	4	12	1
27	0	14	14	1	71	1	46	6	1
28	0	9	9	1	72	1	51	13	1
29	0	9	14	1	73	1	39	10	1
30	0	5	13	1	74	1	37	12	1
31	0	52	14	1	75	1	10	12	1
32	1	10	12	1	76	1	4	12	1
33	1	25	12	1	77	1	4	12	1
34	1	32	12	1	78	1	49	14	0
35	1	29	12	1	79	1	32	12	1
36	1	4	12	1	80	1	9	12	1
37	1	11	15	1	81	1	9	12	1
38	1	11	16	1	82	1	8	13	1
39	1	19	16	1	83	1	5	12	1
40	1	18	12	1	84	1	34	9	1
41	1	25	10	1	85	1	19	8	1
42	1	8	9	1	86	1	41	7	0
43	1	6	10	1	87	1	37	14	1
44	1	2	12	1	88	1	4	9	1

Gözlem	Y	X1	X2	X3	Gözlem	Y	X1	X2	X3
89	1	43	11	1	134	2	50	10	1
90	1	14	12	1	135	2	32	10	1
91	1	9	12	1	136	2	29	12	1
92	1	33	8	1	137	2	9	16	0
93	1	15	13	1	138	2	49	8	1
94	1	12	12	1	139	2	9	14	0
95	1	19	13	1	140	2	41	14	1
96	1	23	8	0	141	2	9	12	1
97	1	26	13	1	142	2	5	11	1
98	1	13	13	1	143	2	17	12	1
99	1	22	12	1	144	2	9	11	1
100	1	4	11	1	145	2	30	12	1
101	2	22	12	1	146	2	29	7	0
102	2	10	11	1	147	2	9	14	1
103	2	21	9	1	148	2	37	12	1
104	2	38	6	0	149	2	44	7	0
105	2	11	12	1	150	2	22	12	1
106	2	47	9	1	151	2	26	12	1
107	2	18	13	1	152	2	10	12	1
108	2	8	12	1	153	2	33	13	1
109	2	13	12	1	154	2	41	8	1
110	2	10	12	1	155	2	39	12	1
111	2	41	11	1	156	2	29	12	0
112	2	49	11	1	157	2	38	14	1
113	2	4	13	1	158	2	12	12	0
114	2	9	12	1	159	2	9	12	0
115	2	33	12	1	160	2	10	14	1
116	2	2	12	1	161	2	9	16	0
117	2	11	15	1	162	2	20	12	1
118	2	56	6	1	163	2	9	11	1
119	2	31	13	1	164	2	41	14	1
120	2	13	14	1	165	2	6	14	1
121	2	33	11	1	166	2	10	12	1
122	2	41	12	1	167	2	11	14	0
123	2	6	12	1	168	2	21	12	1
124	2	21	12	1	169	2	20	13	1
125	2	25	13	1	170	2	31	14	1
126	2	13	15	1	171	2	4	16	1
127	2	2	12	1	172	2	12	13	1
128	2	23	12	1	173	2	17	14	1
129	2	32	12	1	174	2	40	6	1
130	2	46	12	1	175	2	53	12	1
131	2	13	12	1	176	2	35	14	1
132	2	29	12	1	177	2	12	14	1
133	2	30	12	1	178	2	13	15	1

Gözlem	Y	X1	X2	X3	Gözlem	Y	X1	X2	X3
179	2	48	8	1	224	3	6	16	1
180	2	23	12	1	225	3	38	8	1
181	2	11	12	1	226	4	9	16	1
182	2	9	12	1	227	4	10	16	1
183	2	9	12	1	228	4	37	14	1
184	2	4	12	1	229	4	13	14	1
185	3	34	16	1	230	4	11	13	1
186	3	12	12	1	231	4	21	16	0
187	3	21	13	0	232	4	4	16	1
188	3	12	15	1	233	4	2	16	1
189	3	17	12	1	234	4	6	16	1
190	3	21	12	1	235	4	13	17	1
191	3	20	12	1	236	4	13	16	1
192	3	35	12	0	237	4	18	16	1
193	3	44	15	1	238	4	44	14	1
194	3	6	16	1	239	4	9	16	1
195	3	5	14	1	240	4	38	16	1
196	3	42	11	1	241	4	25	14	1
197	3	34	12	1	242	4	32	13	1
198	3	37	16	1	243	4	18	17	1
199	3	19	13	1	244	4	22	19	1
200	3	32	12	1	245	4	20	12	1
201	3	25	12	1	246	4	11	18	1
202	3	19	12	1	247	4	9	16	1
203	3	50	12	1	248	4	14	14	1
204	3	6	12	1	249	4	2	15	1
205	3	49	12	1	250	4	8	14	1
206	3	3	11	1	251	4	26	12	0
207	3	49	18	1	252	4	10	15	1
208	3	39	15	1	253	4	8	12	1
209	3	20	15	1	254	4	25	16	1
210	3	10	12	1	255	4	8	16	1
211	3	5	12	1	256	4	20	14	1
212	3	10	13	1	257	4	11	12	1
213	3	30	16	1	258	4	23	20	1
214	3	31	15	1	259	4	25	19	1
215	3	9	12	1	260	4	31	13	1
216	3	8	12	1	261	4	2	16	1
217	3	49	13	1	262	4	25	14	1
218	3	11	16	1	263	4	3	16	1
219	3	2	12	1	264	4	16	19	0
220	3	6	12	1	265	4	6	14	1
221	3	12	10	1	266	4	17	19	1
222	3	5	17	1	267	4	23	15	1
223	3	3	12	1	268	4	16	14	1

Gözlem	Y	X1	X2	X3	Gözlem	Y	X1	X2	X3
269	4	19	13	1	304	4	29	19	0
270	4	9	14	1	305	4	47	12	0
271	4	14	19	1	306	4	27	20	1
272	4	45	14	1	307	4	11	16	1
273	4	17	16	1	308	4	24	13	1
274	4	42	13	1	309	4	12	14	1
275	4	15	12	1	310	4	14	19	1
276	4	17	12	1	311	4	22	16	1
277	4	8	16	1	312	4	11	18	1
278	4	15	17	1	313	4	16	16	1
279	4	37	18	1	314	4	18	19	1
280	4	33	16	1	315	4	17	13	1
281	4	16	11	1	316	4	41	13	1
282	4	13	18	1	317	4	6	16	1
283	4	44	15	1	318	4	3	16	1
284	4	28	16	1	319	4	6	19	1
285	4	41	11	1	320	4	36	18	1
286	4	25	16	1	321	4	3	12	1
287	4	66	12	1	322	4	6	18	1
288	4	31	12	1	323	4	6	13	1
289	4	25	20	1	324	4	9	17	1
290	4	17	17	1	325	4	6	16	1
291	4	23	16	1	326	4	14	19	1
292	4	34	17	1	327	4	7	19	1
293	4	21	16	1	328	4	11	19	0
294	4	4	15	1	329	4	14	14	1
295	4	41	8	1	330	4	13	16	1
296	4	8	12	1	331	4	24	15	1
297	4	2	19	1	332	4	7	18	1
298	4	20	16	1	333	4	19	13	1
299	4	38	15	1	334	4	43	12	1
300	4	23	16	0	335	4	31	12	1
301	4	31	18	1	336	4	39	7	1
302	4	19	20	1	337	4	12	16	1
303	4	4	20	1					

Ek 2 İkili Logit Modelde Gözlem, Tahmin ve Kalıntı Değerleri

Gözlem	Gözlenen	Tahmin	Kalıntı	Gözlem	Gözlenen	Tahmin	Kalıntı
1	1	0,85464488	0,14535512	46	1	0,74416353	0,25583647
2	1	0,72087497	0,27912503	47	1	0,99774803	0,00225197
3	1	0,58764140	0,41235860	48	1	0,50675173	0,49324827
4	1	0,70464753	0,29535247	49	1	0,70875553	0,29124447
5	1	0,66277508	0,33722492	50	1	0,74037201	0,25962799
6	1	0,46716066	0,53283934	51	1	0,30878705	0,69121295
7	1	0,72963986	0,27036014	52	1	0,45971985	0,54028015
8	1	0,41104075	0,58895925	53	1	0,89077882	0,10922118
9	1	0,74416353	0,25583647	54	1	0,85216542	0,14783458
10	1	0,76256963	0,23743037	55	1	0,92025050	0,07974950
11	1	0,75163683	0,24836317	56	1	0,37894017	0,62105983
12	1	0,73850021	0,26149979	57	1	0,86540044	0,13459956
13	1	0,76613938	0,23386062	58	1	0,91025354	0,08974646
14	1	0,68788514	0,31211486	59	1	0,19615792	0,80384208
15	1	0,75896253	0,24103747	60	1	0,75531821	0,24468179
16	1	0,93113850	0,06886150	61	1	0,72878062	0,27121938
17	1	0,98339719	0,01660281	62	1	0,95810215	0,04189785
18	1	0,76256963	0,23743037	63	1	0,95032703	0,04967297
19	1	0,91499378	0,08500622	64	1	0,70875553	0,29124447
20	1	0,94792557	0,05207443	65	1	0,98097336	0,01902664
21	1	0,50288512	0,49711488	66	1	0,84710172	0,15289828
22	1	0,99302370	0,00697630	67	1	0,98403213	0,01596787
23	1	0,90898189	0,09101811	68	1	0,76794491	0,23205509
24	1	0,63008891	0,36991109	69	1	0,94228496	0,05771504
25	1	0,46464632	0,53535368	70	1	0,75896253	0,24103747
26	1	0,76080617	0,23919383	71	1	0,98097336	0,01902664
27	1	0,43522419	0,56477581	72	1	0,40391012	0,59608988
28	1	0,94593379	0,05406621	73	1	0,84059932	0,15940068
29	1	0,45971985	0,54028015	74	1	0,62080335	0,37919665
30	1	0,62773073	0,37226927	75	1	0,73654418	0,26345582
31	1	0,26625536	0,73374464	76	1	0,75896253	0,24103747
32	1	0,73654418	0,26345582	77	1	0,75896253	0,24103747
33	1	0,67497913	0,32502087	78	1	0,54492855	0,45507145
34	1	0,64383686	0,35616314	79	1	0,64383686	0,35616314
35	1	0,65735172	0,34264828	80	1	0,74037201	0,25962799
36	1	0,75896253	0,24103747	81	1	0,74037201	0,25962799
37	1	0,30878705	0,69121295	82	1	0,61373407	0,38626593
38	1	0,19615792	0,80384208	83	1	0,75531821	0,24468179
39	1	0,17235356	0,82764644	84	1	0,91423509	0,08576491
40	1	0,70464753	0,29535247	85	1	0,96333066	0,03666934
41	1	0,87437036	0,12562964	86	1	0,98976407	0,01023593
42	1	0,94693850	0,05306150	87	1	0,32818312	0,67181688
43	1	0,91025354	0,08974646	88	1	0,95078363	0,04921637
44	1	0,76613938	0,23386062	89	1	0,72685608	0,27314392
45	1	0,96042499	0,03957501	90	1	0,72087497	0,27912503

Gözlem	Gözlenen	Tahmin	Kalıntı	Gözlem	Gözlenen	Tahmin	Kalıntı
91	1	0,74037201	0,25962799	136	1	0,65735172	0,34264828
92	1	0,95216515	0,04783485	137	1	0,44117451	0,55882549
93	1	0,58037178	0,41962822	138	1	0,93546726	0,06453274
94	1	0,72878062	0,27121938	139	1	0,72571243	0,27428757
95	1	0,56095090	0,43904910	140	1	0,31094786	0,68905214
96	1	0,98692156	0,01307844	141	1	0,74037201	0,25962799
97	1	0,52654657	0,47345343	142	1	0,84965112	0,15034888
98	1	0,58999352	0,41000648	143	1	0,70875553	0,29124447
99	1	0,68788514	0,31211486	144	1	0,83924075	0,16075925
100	1	0,85216542	0,14783458	145	1	0,65287377	0,34712623
101	1	0,68788514	0,31211486	146	1	0,99191314	0,00808686
102	1	0,83654879	0,16345121	147	1	0,45971985	0,54028015
103	1	0,93239850	0,06760150	148	1	0,62080335	0,37919665
104	1	0,99470513	0,00529487	149	1	0,98914381	0,01085619
105	1	0,73268030	0,26731970	150	1	0,68788514	0,31211486
106	1	0,89175801	0,10824199	151	1	0,67061611	0,32938389
107	1	0,56582607	0,43417393	152	1	0,73654418	0,26345582
108	1	0,74416353	0,25583647	153	1	0,49188830	0,50811170
109	1	0,72484541	0,27515459	154	1	0,94440326	0,05559674
110	1	0,73654418	0,26345582	155	1	0,61142845	0,38857155
111	1	0,73465468	0,26534532	156	1	0,85643149	0,14356851
112	1	0,70262163	0,29737837	157	1	0,32382833	0,67617167
113	1	0,63235038	0,36764962	158	1	0,89310833	0,10689167
114	1	0,74037201	0,25962799	159	1	0,89865332	0,10134668
115	1	0,63927921	0,36072079	160	1	0,45480125	0,54519875
116	1	0,76613938	0,23386062	161	1	0,44117451	0,55882549
117	1	0,30878705	0,69121295	162	1	0,69633155	0,30366845
118	1	0,97689908	0,02310092	163	1	0,83924075	0,16075925
119	1	0,50179728	0,49820272	164	1	0,31094786	0,68905214
120	1	0,44010199	0,55989801	165	1	0,47451903	0,52548097
121	1	0,76439424	0,23560576	166	1	0,73654418	0,26345582
122	1	0,60197037	0,39802963	167	1	0,71775199	0,28224801
123	1	0,75163683	0,24836317	168	1	0,69212442	0,30787558
124	1	0,69212442	0,30787558	169	1	0,55606393	0,44393607
125	1	0,53148470	0,46851530	170	1	0,35491642	0,64508358
126	1	0,30039137	0,69960863	171	1	0,21895836	0,78104164
127	1	0,76613938	0,23386062	172	1	0,59477916	0,40522084
128	1	0,68361416	0,31638584	173	1	0,42066946	0,57933054
129	1	0,64383686	0,35616314	174	1	0,98307047	0,01692953
130	1	0,57800394	0,42199606	175	1	0,54384927	0,45615073
131	1	0,72484541	0,27515459	176	1	0,33698142	0,66301858
132	1	0,65735172	0,34264828	177	1	0,44499138	0,55500862
133	1	0,65287377	0,34712623	178	1	0,30039137	0,69960863
134	1	0,80917841	0,19082159	179	1	0,93665344	0,06334656
135	1	0,85832291	0,14167709	180	1	0,68361416	0,31638584

Gözlem	Gözlenen	Tahmin	Kalıntı	Gözlem	Gözlenen	Tahmin	Kalıntı
181	1	0,73268030	0,26731970	226	0	0,20248353	-0,20248353
182	1	0,74037201	0,25962799	227	0	0,19930187	-0,19930187
183	1	0,74037201	0,25962799	228	0	0,32818312	-0,32818312
184	1	0,75896253	0,24103747	229	0	0,44010199	-0,44010199
185	0	0,13396703	-0,13396703	230	0	0,59954687	-0,59954687
186	0	0,72878062	-0,72878062	231	0	0,38361563	-0,38361563
187	0	0,79246207	-0,79246207	232	0	0,21895836	-0,21895836
188	0	0,30457295	-0,30457295	233	0	0,22581268	-0,22581268
189	0	0,70875553	-0,70875553	234	0	0,21225506	-0,21225506
190	0	0,69212442	-0,69212442	235	0	0,11356743	-0,11356743
191	0	0,69633155	-0,69633155	236	0	0,18998285	-0,18998285
192	0	0,84118151	-0,84118151	237	0	0,17519915	-0,17519915
193	0	0,18849231	-0,18849231	238	0	0,29835366	-0,29835366
194	0	0,21225506	-0,21225506	239	0	0,20248353	-0,20248353
195	0	0,47946335	-0,47946335	240	0	0,12503317	-0,12503317
196	0	0,73077322	-0,73077322	241	0	0,38258724	-0,38258724
197	0	0,63469632	-0,63469632	242	0	0,49684248	-0,49684248
198	0	0,12721757	-0,12721757	243	0	0,10396684	-0,10396684
199	0	0,56095090	-0,56095090	244	0	0,03099162	-0,03099162
200	0	0,64383686	-0,64383686	245	0	0,69633155	-0,69633155
201	0	0,67497913	-0,67497913	246	0	0,06787139	-0,06787139
202	0	0,70050606	-0,70050606	247	0	0,20248353	-0,20248353
203	0	0,55855683	-0,55855683	248	0	0,43522419	-0,43522419
204	0	0,75163683	-0,75163683	249	0	0,34809547	-0,34809547
205	0	0,56343790	-0,56343790	250	0	0,46464632	-0,46464632
206	0	0,85464488	-0,85464488	251	0	0,86358848	-0,86358848
207	0	0,03315022	-0,03315022	252	0	0,31303321	-0,31303321
208	0	0,20412013	-0,20412013	253	0	0,74416353	-0,74416353
209	0	0,27206500	-0,27206500	254	0	0,15604503	-0,15604503
210	0	0,73654418	-0,73654418	255	0	0,20570292	-0,20570292
211	0	0,75531821	-0,75531821	256	0	0,40625135	-0,40625135
212	0	0,60429579	-0,60429579	257	0	0,73268030	-0,73268030
213	0	0,14343459	-0,14343459	258	0	0,01683921	-0,01683921
214	0	0,23108687	-0,23108687	259	0	0,02925496	-0,02925496
215	0	0,74037201	-0,74037201	260	0	0,50179728	-0,50179728
216	0	0,74416353	-0,74416353	261	0	0,22581268	-0,22581268
217	0	0,41348906	-0,41348906	262	0	0,38258724	-0,38258724
218	0	0,19615792	-0,19615792	263	0	0,22236666	-0,22236666
219	0	0,76613938	-0,76613938	264	0	0,10072519	-0,10072519
220	0	0,75163683	-0,75163683	265	0	0,47451903	-0,47451903
221	0	0,90005341	-0,90005341	266	0	0,03411000	-0,03411000
222	0	0,13053314	-0,13053314	267	0	0,26045038	-0,26045038
223	0	0,76256963	-0,76256963	268	0	0,42550703	-0,42550703
224	0	0,21225506	-0,21225506	269	0	0,56095090	-0,56095090
225	0	0,94744392	-0,94744392	270	0	0,45971985	-0,45971985

Gözlem	Gözlenen	Tahmin	Kalıntı	Gözlem	Gözlenen	Tahmin	Kalıntı
271	0	0,03612414	-0,03612414	305	0	0,80677859	-0,80677859
272	0	0,29422134	-0,29422134	306	0	0,01557577	-0,01557577
273	0	0,17808160	-0,17808160	307	0	0,19615792	-0,19615792
274	0	0,44748826	-0,44748826	308	0	0,53641666	-0,53641666
275	0	0,71686962	-0,71686962	309	0	0,44499138	-0,44499138
276	0	0,70875553	-0,70875553	310	0	0,03612414	-0,03612414
277	0	0,20570292	-0,20570292	311	0	0,16403646	-0,16403646
278	0	0,10963770	-0,10963770	312	0	0,06787139	-0,06787139
279	0	0,04167998	-0,04167998	313	0	0,18100107	-0,18100107
280	0	0,13628320	-0,13628320	314	0	0,03346301	-0,03346301
281	0	0,81963164	-0,81963164	315	0	0,57068854	-0,57068854
282	0	0,06540619	-0,06540619	316	0	0,45239341	-0,45239341
283	0	0,18849231	-0,18849231	317	0	0,21225506	-0,21225506
284	0	0,14837382	-0,14837382	318	0	0,22236666	-0,22236666
285	0	0,73465468	-0,73465468	319	0	0,04206975	-0,04206975
286	0	0,15604503	-0,15604503	320	0	0,04247885	-0,04247885
287	0	0,47956052	-0,47956052	321	0	0,76256963	-0,76256963
288	0	0,64836860	-0,64836860	322	0	0,07441564	-0,07441564
289	0	0,01619538	-0,01619538	323	0	0,62308763	-0,62308763
290	0	0,10582772	-0,10582772	324	0	0,12179608	-0,12179608
291	0	0,16133671	-0,16133671	325	0	0,21225506	-0,21225506
292	0	0,07791521	-0,07791521	326	0	0,03612414	-0,03612414
293	0	0,16677239	-0,16677239	327	0	0,04127824	-0,04127824
294	0	0,33915543	-0,33915543	328	0	0,11006320	-0,11006320
295	0	0,94440326	-0,94440326	329	0	0,43522419	-0,43522419
296	0	0,74416353	-0,74416353	330	0	0,18998285	-0,18998285
297	0	0,04538319	-0,04538319	331	0	0,25665101	-0,25665101
298	0	0,16954470	-0,16954470	332	0	0,07306198	-0,07306198
299	0	0,20735878	-0,20735878	333	0	0,56095090	-0,56095090
300	0	0,37428710	-0,37428710	334	0	0,59243554	-0,59243554
301	0	0,04669737	-0,04669737	335	0	0,64836860	-0,64836860
302	0	0,01820324	-0,01820324	336	0	0,97001907	-0,97001907
303	0	0,02435186	-0,02435186	337	0	0,19305160	-0,19305160
304	0	0,07966956	-0,07966956				

Ek 3 Sıralı Probit Modelde Alternatiflere İlişkin Olasılıklar

Gözlem	Alt.1	Alt.2	Alt.3	Alt.4	Alt.5	Gözlem	Alt.1	Alt.2	Alt.3	Alt.4	Alt.5
1	0,19882632	0,32478970	0,28634433	0,09472616	0,09531349	29	0,04201977	0,16333042	0,29299709	0,16690247	0,33475025
2	0,10232874	0,25584623	0,31733432	0,13672346	0,18776725	30	0,07998537	0,22851068	0,31641361	0,14827550	0,22681484
3	0,04484008	0,16933000	0,29639419	0,16576558	0,32367015	31	0,00958586	0,06575566	0,19274710	0,15749567	0,57441571
4	0,08249397	0,24456598	0,31763179	0,14175463	0,20355364	32	0,11290282	0,26683018	0,31605937	0,13144729	0,17276034
5	0,05139700	0,18239840	0,30290104	0,16284804	0,30045552	33	0,07699463	0,22435281	0,31579713	0,14984430	0,23301113
6	0,03027733	0,13532727	0,27353060	0,17028733	0,39057747	34	0,06355354	0,20389182	0,31101699	0,15684275	0,26469490
7	0,08033527	0,22898858	0,31647666	0,14809203	0,22610745	35	0,06907170	0,21266561	0,31341954	0,15398991	0,25085324
8	0,03069945	0,13643282	0,27441766	0,17022023	0,38822984	36	0,13018957	0,28257338	0,31234232	0,12316712	0,15172760
9	0,11847302	0,27218573	0,31505867	0,12873066	0,16555192	37	0,02160480	0,11047232	0,25073360	0,17004092	0,44714836
10	0,13323992	0,28509558	0,31151651	0,12175300	0,14839499	38	0,01109608	0,07238812	0,20306583	0,16073831	0,55271167
11	0,12423470	0,27743683	0,31381870	0,12596881	0,15854095	39	0,00816665	0,05911730	0,18166208	0,15357241	0,59748156
12	0,08463652	0,23472274	0,31710649	0,14584003	0,21769423	40	0,09249397	0,24456598	0,31763179	0,14175463	0,20355364
13	0,13633909	0,28758539	0,31063290	0,12033092	0,14511170	41	0,18544386	0,31864139	0,29229446	0,09973795	0,10388235
14	0,08337663	0,23306960	0,31694902	0,14649888	0,22010586	42	0,34955302	0,34840026	0,21143692	0,05207327	0,03852652
15	0,13018957	0,28257338	0,31234232	0,12316712	0,15172760	43	0,26654186	0,34438921	0,25346071	0,07275342	0,06285481
16	0,27875042	0,34619465	0,24728861	0,06932458	0,05844175	44	0,13633909	0,28758539	0,31063290	0,12033092	0,14511170
17	0,48670571	0,32161430	0,14618908	0,02853917	0,01695175	45	0,36863803	0,34688386	0,20192608	0,04810277	0,03444925
18	0,13323992	0,28509558	0,31151651	0,12175300	0,14839499	46	0,11847302	0,27218573	0,31505867	0,12873066	0,16555192
19	0,28081024	0,34645387	0,24624460	0,06876043	0,05773086	47	0,81745852	0,14745358	0,03080977	0,00317244	0,00110569
20	0,35486800	0,34806034	0,20878167	0,05094284	0,03734716	48	0,03899961	0,15662261	0,28888525	0,16801242	0,34748011
21	0,02620742	0,12420675	0,26402526	0,17061326	0,41494731	49	0,09488449	0,24740941	0,31764936	0,14052212	0,19953461
22	0,53645503	0,30405563	0,12473336	0,02240571	0,01235027	50	0,11566411	0,26952038	0,31558908	0,13009496	0,16913147
23	0,21168118	0,32991294	0,28039435	0,09013472	0,08787680	51	0,02160480	0,11047232	0,25073360	0,17004092	0,44714836
24	0,05838468	0,19514388	0,30809704	0,15945963	0,27891478	52	0,04201977	0,16333042	0,29299709	0,16690247	0,33475025
25	0,04331727	0,16612097	0,29460976	0,16639017	0,32956182	53	0,21755026	0,33201330	0,27761587	0,08810813	0,08471243
26	0,09873301	0,25184794	0,31755308	0,13855038	0,19331559	54	0,19486400	0,32305899	0,28813481	0,09618490	0,09775729
27	0,03599646	0,14963486	0,28424063	0,16897733	0,36115072	55	0,24169020	0,33920187	0,26588192	0,08020731	0,07301870
28	0,34428727	0,34867404	0,21408455	0,05321808	0,03973606	56	0,01643586	0,09317609	0,23126335	0,16746237	0,49166232

Gözlem	Alt. 1	Alt. 2	Alt. 3	Alt. 4	Alt. 5	Gözlem	Alt. 1	Alt. 2	Alt. 3	Alt. 4	Alt. 5
57	0,1705559	0,3107331	0,2985354	0,1056055	0,1145701	86	0,4933419	0,3194851	0,143251	0,0276579	0,0162642
58	0,2665419	0,3443892	0,2534607	0,0727534	0,0628548	87	0,0166757	0,0940315	0,2323047	0,1676464	0,4893418
59	0,0110961	0,0723881	0,2030658	0,1607383	0,5527117	88	0,3709489	0,3466458	0,2007812	0,0476392	0,033985
60	0,1271879	0,2800201	0,3131099	0,1245726	0,1551096	89	0,0781811	0,2260182	0,3160586	0,1492218	0,2305202
61	0,1075221	0,2613803	0,3168185	0,1341133	0,1801657	90	0,1023287	0,2558462	0,3173343	0,1367235	0,1877672
62	0,352592	0,348214	0,20992	0,0514253	0,0378487	91	0,1156641	0,2695204	0,3155891	0,130095	0,1691315
63	0,3060984	0,3486278	0,2334015	0,0621557	0,0497166	92	0,3161968	0,3490012	0,2282748	0,0596764	0,0468508
64	0,0948845	0,2474094	0,3176494	0,1405221	0,1995346	93	0,060803	0,199304	0,3095509	0,158244	0,2720982
65	0,4469549	0,3328971	0,1642545	0,0342633	0,0216301	94	0,1075221	0,2613803	0,3168185	0,1341133	0,1801657
66	0,1870848	0,3194424	0,2915805	0,0991103	0,1027821	95	0,0542219	0,1876871	0,3051934	0,1615048	0,2913928
67	0,4981049	0,3179153	0,1411563	0,0270379	0,0157856	96	0,4900839	0,3205389	0,1446906	0,028088	0,0165987
68	0,1037941	0,2574359	0,3172118	0,1359835	0,1855747	97	0,0440734	0,1677231	0,29551	0,166083	0,3266106
69	0,2672147	0,3445009	0,2531214	0,0725606	0,0626023	98	0,0643198	0,2051437	0,311392	0,1564495	0,2626949
70	0,1301896	0,2825734	0,3123423	0,1231671	0,1517276	99	0,0833766	0,2330696	0,316949	0,1464989	0,2201059
71	0,4469549	0,3328971	0,1642545	0,0342633	0,0216301	100	0,194864	0,323059	0,2881348	0,0961849	0,0977573
72	0,01958	0,1039645	0,2437868	0,1693382	0,4633304	101	0,0833766	0,2330696	0,316949	0,1464989	0,2201059
73	0,136787	0,2879393	0,3105016	0,1201266	0,1446454	102	0,1721127	0,311616	0,2979047	0,1049771	0,1133894
74	0,0551263	0,1893403	0,3058702	0,1610664	0,2885968	103	0,2835703	0,3467812	0,2448448	0,0680109	0,0567929
75	0,1129028	0,2668302	0,3160594	0,1314473	0,1727603	104	0,61469	0,2695819	0,0937829	0,0147249	0,0072203
76	0,1301896	0,2825734	0,3123423	0,1231671	0,1517276	105	0,1101889	0,2641164	0,3164693	0,132787	0,1764385
77	0,1301896	0,2825734	0,3123423	0,1231671	0,1517276	106	0,1726353	0,3119094	0,2976918	0,104767	0,1129966
78	0,0233671	0,115893	0,2561943	0,1704064	0,4341391	107	0,0558118	0,1905809	0,3063657	0,1607318	0,2865097
79	0,0635535	0,2038918	0,311017	0,1568428	0,2646949	108	0,118473	0,2721857	0,3150587	0,1287307	0,1655519
80	0,1156641	0,2695204	0,3155891	0,130095	0,1691315	109	0,1049022	0,2586232	0,3171069	0,1354257	0,183942
81	0,1156641	0,2695204	0,3155891	0,130095	0,1691315	110	0,1129028	0,2668302	0,3160594	0,1314473	0,1727603
82	0,073803	0,2197663	0,3149768	0,1515178	0,2399361	111	0,0824417	0,2318289	0,316818	0,1469882	0,2219232
83	0,1271879	0,2800201	0,3131099	0,1245726	0,1551096	112	0,0664	0,2084868	0,312341	0,1553768	0,2573955
84	0,224215	0,3342257	0,2744211	0,0858581	0,0812801	113	0,0821309	0,2314137	0,3167717	0,147151	0,2225328
85	0,3903897	0,3441983	0,1912206	0,0438805	0,0303109	114	0,1156641	0,2695204	0,3155891	0,130095	0,1691315

Gözlem	Alt. 1	Alt. 2	Alt. 3	Alt. 4	Alt. 5	Gözlem	Alt. 1	Alt. 2	Alt. 3	Alt. 4	Alt. 5
115	0,0617924	0,2009715	0,3101005	0,1577419	0,2693936	143	0,0948845	0,2474094	0,3176494	0,1405221	0,1995346
116	0,1363391	0,2875854	0,3106329	0,1203309	0,1451117	144	0,1757821	0,3136442	0,2963967	0,1035099	0,1106672
117	0,0216048	0,1104723	0,2507336	0,1700409	0,4471484	145	0,0671927	0,2097399	0,312677	0,1549662	0,2554242
118	0,3911755	0,344083	0,190837	0,0437337	0,0301707	146	0,5615055	0,2938857	0,1144427	0,0197036	0,0104625
119	0,0378084	0,1538906	0,2871138	0,1684136	0,3527737	147	0,0420198	0,1633304	0,2929971	0,1669025	0,3347502
120	0,037141	0,1523372	0,2860812	0,1686281	0,3558125	148	0,0551263	0,1893403	0,3058702	0,1610664	0,2885968
121	0,1012405	0,2546509	0,317413	0,1372747	0,1894208	149	0,4762532	0,3248289	0,1508629	0,029969	0,018086
122	0,0490416	0,177837	0,3007671	0,1639327	0,3084215	150	0,0833766	0,2330696	0,316949	0,1464989	0,2201059
123	0,1242347	0,2774368	0,3138187	0,1259688	0,158541	151	0,074952	0,2214356	0,3152923	0,1509156	0,2374045
124	0,08559	0,2359596	0,3172114	0,145342	0,215897	152	0,1129028	0,2668302	0,3160594	0,1314473	0,1727603
125	0,0454217	0,1705373	0,2970463	0,1655209	0,3214738	153	0,0355151	0,1484833	0,283439	0,1691168	0,3634458
126	0,0201696	0,1058923	0,2458905	0,1695778	0,4584697	154	0,276698	0,3459235	0,2483282	0,0698908	0,0591595
127	0,1363391	0,2875854	0,3106329	0,1203309	0,1451117	155	0,0520133	0,1835688	0,3034251	0,1625586	0,2984342
128	0,0812065	0,230171	0,3166257	0,1476333	0,2243614	156	0,1205712	0,274131	0,31463	0,1277192	0,1629486
129	0,0635535	0,2038918	0,311017	0,1568428	0,2646949	157	0,016092	0,0919393	0,2297423	0,1671849	0,4950415
130	0,0422037	0,1637292	0,293231	0,166831	0,3340051	158	0,1763902	0,3139733	0,296144	0,1032685	0,110224
131	0,1049022	0,2586232	0,3171069	0,1354257	0,183942	159	0,1877171	0,3197474	0,2913041	0,0988694	0,102362
132	0,0690717	0,2126656	0,3134195	0,1539899	0,2508532	160	0,0407539	0,160556	0,291337	0,1673827	0,3399704
133	0,0671927	0,2097399	0,312677	0,1549662	0,2554242	161	0,0257147	0,1227995	0,2627446	0,1706068	0,4181344
134	0,1052746	0,2590194	0,3170693	0,1352386	0,1833981	162	0,0878471	0,2388398	0,3174128	0,1441652	0,2117351
135	0,159907	0,3043208	0,3026891	0,1099992	0,123084	163	0,1757821	0,3136442	0,2963967	0,1035099	0,1106672
136	0,0690717	0,2126656	0,3134195	0,1539899	0,2508532	164	0,0144449	0,0858391	0,221966	0,1656081	0,5121419
137	0,0257147	0,1227995	0,2627446	0,1706068	0,4181344	165	0,0460093	0,1717473	0,2976893	0,1652706	0,3192836
138	0,239784	0,3387148	0,266823	0,0808067	0,0738715	166	0,1129028	0,2668302	0,3160594	0,1314473	0,1727603
139	0,0782254	0,2260799	0,316068	0,1491986	0,2304281	167	0,0741322	0,2202467	0,3150696	0,1513453	0,2392062
140	0,0144449	0,0858391	0,221966	0,1656081	0,5121419	168	0,08559	0,2359596	0,3172114	0,145342	0,215897
141	0,1156641	0,2695204	0,3155891	0,130095	0,1691315	169	0,052668	0,1848019	0,3039669	0,1622488	0,2963143
142	0,1909501	0,3212762	0,2898805	0,0976465	0,1002467	170	0,0205713	0,10719	0,2472846	0,169724	0,4552301

Gözlem	Alt.1	Alt.2	Alt.3	Alt.4	Alt.5	Gözlem	Alt.1	Alt.2	Alt.3	Alt.4	Alt.5
171	0,01437	0,0855544	0,2215917	0,1655257	0,5129581	199	0,0542219	0,1876871	0,3051934	0,1615048	0,2913928
172	0,0661365	0,2080678	0,3122262	0,155513	0,2580564	200	0,0635535	0,2038918	0,311017	0,1568428	0,2646949
173	0,0327352	0,1416492	0,2784668	0,1688244	0,3773245	201	0,0769946	0,2243528	0,3157971	0,1498443	0,2330111
174	0,4810096	0,3233874	0,1487291	0,0293119	0,0175619	202	0,0901483	0,241709	0,3175529	0,1429692	0,2076206
175	0,0339481	0,1446703	0,2807106	0,1695373	0,3711338	203	0,0373073	0,1527259	0,2863413	0,1685754	0,3550501
176	0,0178975	0,0983031	0,2373778	0,1684715	0,4779502	204	0,1242347	0,2774368	0,3138187	0,1259688	0,158541
177	0,038315	0,1550587	0,2878781	0,1682458	0,3505025	205	0,0384856	0,15545	0,2881319	0,1681883	0,3497442
178	0,0201696	0,1058923	0,2458905	0,1695778	0,4584697	206	0,1988263	0,3247897	0,2863443	0,0947262	0,0953135
179	0,2442465	0,3398343	0,2646167	0,0794099	0,0718925	207	0,0003885	0,0066402	0,043758	0,0630648	0,8861485
180	0,0812065	0,230171	0,3166257	0,1476353	0,2243614	208	0,0077221	0,056944	0,1778533	0,1521209	0,6053597
181	0,1101889	0,2641164	0,3164693	0,132787	0,1764385	209	0,0157663	0,0907565	0,2282705	0,1669066	0,4983
182	0,1156641	0,2695204	0,3155891	0,130095	0,1691315	210	0,1129028	0,2668302	0,3160594	0,1314473	0,1727603
183	0,1156641	0,2695204	0,3155891	0,130095	0,1691315	211	0,1271879	0,2800201	0,3131099	0,1245726	0,1551096
184	0,1301896	0,2825734	0,3123423	0,1231671	0,1517276	212	0,0698889	0,2139188	0,3137196	0,1535641	0,2489085
185	0,0044519	0,0391186	0,1425839	0,1361801	0,6776655	213	0,0052552	0,0438577	0,1527466	0,141239	0,6569016
186	0,1075221	0,2613803	0,3168185	0,1341133	0,1801657	214	0,0105142	0,0698804	0,1992481	0,1595853	0,5607719
187	0,092884	0,2450346	0,3176389	0,1415532	0,2028894	215	0,1156641	0,2695204	0,3155891	0,130095	0,1691315
188	0,0208768	0,1081685	0,2483242	0,1698263	0,4528042	216	0,118473	0,2721857	0,3150587	0,1287307	0,1655519
189	0,0948845	0,2474094	0,3176494	0,1405221	0,1995346	217	0,0209798	0,1084968	0,2486709	0,1698591	0,4519934
190	0,08559	0,2359596	0,3172114	0,145342	0,215897	218	0,0110961	0,0723881	0,2030658	0,1607383	0,5527117
191	0,0878471	0,2388398	0,3174128	0,1441652	0,2117351	219	0,1363391	0,2875854	0,3106329	0,1203309	0,1451117
192	0,1042185	0,2578921	0,3171729	0,1357697	0,1849468	220	0,1242347	0,2774368	0,3138187	0,1259688	0,158541
193	0,0063291	0,0497926	0,164626	0,1466724	0,6325799	221	0,2391485	0,3385495	0,2671363	0,0810074	0,0741583
194	0,013359	0,0816439	0,2163416	0,1643086	0,5243468	222	0,0068181	0,0523664	0,1695147	0,1487598	0,622541
195	0,0474048	0,1745813	0,2991544	0,1646639	0,3141955	223	0,1332399	0,2850956	0,3115165	0,121753	0,148395
196	0,0802901	0,228927	0,3164686	0,1481157	0,2261986	224	0,013359	0,0816439	0,2163416	0,1643086	0,5243468
197	0,0600696	0,1980552	0,309127	0,1586144	0,2741338	225	0,291225	0,3475711	0,2409587	0,0659695	0,0542758
198	0,0039233	0,0358309	0,1351331	0,1322258	0,6928869	226	0,0119578	0,0760012	0,2083963	0,1622549	0,5413898

Gözlem	Alt.1	Alt.2	Alt.3	Alt.4	Alt.5	Gözlem	Alt.1	Alt.2	Alt.3	Alt.4	Alt.5
227	0,0115199	0,0741798	0,2057337	0,161511	0,5470555	255	0,01241	0,0778523	0,2110526	0,1629693	0,5357157
228	0,0166757	0,0940315	0,2323047	0,1676464	0,4893418	256	0,0297216	0,1338586	0,2723363	0,1703671	0,3937163
229	0,037141	0,1523372	0,2860812	0,1686281	0,3558125	257	0,1101889	0,2641164	0,3164693	0,132787	0,1764385
230	0,0679927	0,2109932	0,3130023	0,154551	0,2534608	258	0,0002157	0,0042453	0,0317458	0,0498361	0,913957
231	0,0170302	0,0952855	0,2338159	0,1679046	0,4859638	259	0,000513	0,0081851	0,0507169	0,0700915	0,8704934
232	0,01437	0,0855544	0,2215917	0,1655257	0,5129581	260	0,0378084	0,1538906	0,2871138	0,1684136	0,3527737
233	0,0154462	0,0895834	0,2267938	0,1666178	0,5015587	261	0,0154462	0,0895834	0,2267938	0,1666178	0,5015587
234	0,013359	0,0816439	0,2163416	0,1643086	0,5243468	262	0,0252119	0,1213488	0,2614056	0,1705882	0,4214455
235	0,0049253	0,0419462	0,1487249	0,1392827	0,6651208	263	0,0148998	0,0875541	0,2241994	0,1660876	0,5072592
236	0,0102889	0,0688938	0,1977187	0,1591079	0,5639907	264	0,0022271	0,024023	0,1050023	0,1139975	0,7547502
237	0,0084911	0,0606738	0,1843338	0,1545591	0,5919422	265	0,0460093	0,1717473	0,2976893	0,1652706	0,3192836
238	0,012945	0,0800051	0,2140802	0,1637497	0,52922	266	0,0007649	0,0110251	0,0623849	0,0810191	0,844806
239	0,0119578	0,0760012	0,2083963	0,1622549	0,5413898	267	0,0141481	0,0847069	0,220471	0,1652755	0,5153986
240	0,0037601	0,0347841	0,1326857	0,1308807	0,6978894	268	0,0337941	0,1442902	0,2804324	0,1695757	0,3719077
241	0,0252119	0,1213488	0,2614056	0,1705882	0,4214455	269	0,0542219	0,1876871	0,3051934	0,1615048	0,2913928
242	0,0366471	0,1511772	0,285298	0,1687818	0,3580959	270	0,0420198	0,1633304	0,2929971	0,1669025	0,3347502
243	0,0039952	0,0362867	0,1361872	0,132798	0,6907328	271	0,0008857	0,0122886	0,0672064	0,0852593	0,83436
244	0,0005968	0,0091657	0,0548918	0,0741167	0,8612289	272	0,012476	0,07812	0,2114325	0,1630693	0,5349023
245	0,0878471	0,2388398	0,3174128	0,1441652	0,2117351	273	0,0088269	0,0622593	0,1870085	0,1555209	0,5863845
246	0,0024175	0,0254706	0,1090321	0,1166539	0,7464259	274	0,0265632	0,1252142	0,2649311	0,1706111	0,4126805
247	0,0119578	0,0760012	0,2083963	0,1622549	0,5413898	275	0,0998015	0,2530508	0,3175006	0,1380059	0,1916412
248	0,0359965	0,1496349	0,2842406	0,1689773	0,3611507	276	0,0948845	0,2474094	0,3176494	0,1405221	0,1995346
249	0,0291743	0,1323977	0,2711302	0,1704357	0,3968621	277	0,01241	0,0778523	0,2110526	0,1629693	0,5357157
250	0,0433173	0,166121	0,2946098	0,1663902	0,3295618	278	0,0045323	0,0396064	0,1436604	0,1367341	0,6754667
251	0,1293926	0,2819025	0,3125506	0,1235389	0,1526154	279	0,0007129	0,0104636	0,0601754	0,0790247	0,8496234
252	0,0223542	0,1128036	0,2531175	0,1702216	0,4415032	280	0,0046417	0,0402647	0,1451015	0,1374691	0,6725231
253	0,118473	0,2721857	0,3150587	0,1287307	0,1655519	281	0,1511325	0,2985213	0,3058722	0,1137467	0,1307274
254	0,0064391	0,0503779	0,165751	0,1471604	0,6302715	282	0,0022107	0,0238966	0,1046452	0,1137587	0,7554888

Gözlem	Alt.1	Alt.2	Alt.3	Alt.4	Alt.5	Gözlem	Alt.1	Alt.2	Alt.3	Alt.4	Alt.5
283	0,0063291	0,0497926	0,164626	0,1466724	0,6325799	311	0,0072577	0,0546202	0,1736754	0,1504683	0,6139783
284	0,0057033	0,0463847	0,1579129	0,1436651	0,6463341	312	0,0024175	0,0254706	0,1090321	0,1166539	0,7464259
285	0,0824417	0,2318289	0,316818	0,1469882	0,2219232	313	0,0091742	0,0638739	0,1896854	0,1564572	0,5808093
286	0,0084391	0,0503779	0,165751	0,1471604	0,6302715	314	0,0007282	0,0106294	0,0608322	0,079621	0,8481892
287	0,0221383	0,1121361	0,2524404	0,1701735	0,4431117	315	0,0574383	0,1934823	0,3074832	0,1599303	0,2816658
288	0,0653535	0,206815	0,311876	0,1559173	0,2600382	316	0,0274541	0,127706	0,2671327	0,1705815	0,4071258
289	0,0001936	0,0039078	0,0298925	0,0476481	0,918358	317	0,013359	0,0816439	0,2163416	0,1643086	0,5243468
290	0,0041675	0,0373684	0,1386609	0,1341243	0,6856789	318	0,0148998	0,0875541	0,2241994	0,1660876	0,5072592
291	0,0069752	0,0531781	0,1710258	0,1493876	0,6194334	319	0,0012987	0,0162724	0,0812676	0,0968228	0,8043385
292	0,0019814	0,0220926	0,0994516	0,1102206	0,7662539	320	0,000749	0,010854	0,0617165	0,0804192	0,8462614
293	0,0075502	0,0560907	0,176332	0,1515265	0,6085007	321	0,1332399	0,2850956	0,3115165	0,121753	0,148395
294	0,0273249	0,1273475	0,2668193	0,1705878	0,4079203	322	0,0030128	0,0297737	0,1204118	0,1237819	0,7230197
295	0,276698	0,3459235	0,2483282	0,0698908	0,0591595	323	0,0778825	0,225601	0,315995	0,1493785	0,231143
296	0,118473	0,2721857	0,3150587	0,1287307	0,1655519	324	0,0058036	0,0469399	0,159026	0,144175	0,6440556
297	0,0015653	0,0186385	0,0889505	0,1026833	0,7881624	325	0,013359	0,0816439	0,2163416	0,1643086	0,5243468
298	0,0078531	0,0575897	0,1789945	0,1525613	0,6030014	326	0,0008857	0,0122886	0,0672064	0,0852593	0,83436
299	0,008031	0,0584594	0,1805188	0,1531422	0,5998486	327	0,0012388	0,0157218	0,0794144	0,0953628	0,8082622
300	0,0158594	0,0910958	0,2286945	0,1669877	0,4973625	328	0,0027804	0,0281312	0,1161678	0,1211867	0,7317339
301	0,0009558	0,0129982	0,0698305	0,0875052	0,8287102	329	0,0359965	0,1496349	0,2842406	0,1689773	0,3611507
302	0,0002672	0,0049986	0,0357209	0,0543868	0,9046265	330	0,0102889	0,0688938	0,1977187	0,1591079	0,5639907
303	0,0005799	0,0089714	0,0540779	0,0733426	0,8630281	331	0,0136415	0,0827495	0,2178465	0,1646691	0,5210934
304	0,0012235	0,0155791	0,0789298	0,0949779	0,8092897	332	0,0028841	0,0288697	0,1180904	0,1223715	0,7277843
305	0,0764515	0,2235833	0,3156697	0,1501292	0,2341662	333	0,0542219	0,1876871	0,3051934	0,1615048	0,2913928
306	0,0001736	0,0035943	0,0281254	0,0455195	0,9225871	334	0,0462071	0,1721524	0,2979022	0,1651856	0,3185526
307	0,0110961	0,0723881	0,2030658	0,1607383	0,5527117	335	0,0653535	0,206815	0,311876	0,1559173	0,2600382
308	0,0468029	0,1733657	0,298533	0,1649276	0,3163709	336	0,3825789	0,3452758	0,195046	0,0453607	0,0317387
309	0,038315	0,1550587	0,2878781	0,1682458	0,3505025	337	0,0106859	0,0706261	0,2003938	0,1599371	0,5583571
310	0,0008857	0,0122886	0,0672064	0,0852593	0,83436						

ÖZGEÇMİŞ**Kişisel Bilgi**

- Adı Soyadı : Özgür Sönmez
- Doğum Yeri : Bakırköy/İstanbul
- Doğum Tarihi : 10-06-1980

Eğitim

- Avcılar Anadolu Teknik Lisesi Makine Bölümü (1995-1998)
- Yıldız Teknik Üniversitesi Fen-Edebiyat Fakültesi, İstatistik Bölümü (1998-2003)
- Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü İstatistik Ana Bilim Dalı, İstatistik Programı (2003-2006)

İş deneyimi

- 01.02.2002 – 01.03.2002 tarihleri arasında Tacirler Menkul Değerler A.Ş Genel Müdürlük Müşteri Hizmetleri Bölümünde stajyerlik yaptım.
- 31.03.2003 – 11.04.2003 tarihleri arasında İ.M.K.B (İstanbul Menkul Kıymetler Borsası) de stajyerlik yaptım.
- Dışbank A.Ş Operasyon Merkezi Kredi Kartları bölümünde 27.08.2002 – 01.04.2005 tarihleri arasında bilgi giriş,bilgi kontrol ve müşteri özlük eşleştirme elemanı olarak görev aldım.

