

T.C.
İstanbul Üniversitesi
Sosyal Bilimler Enstitüsü
Ekonometri Anabilim Dalı

Yüksek Lisans Tezi

Ülkelerin Kalkınmışlık Düzeylerini Etkileyen
Faktörlerin Çok Seçenekli Tercih Modelleri ile
İncelenmesi

Burak Güriş
2501020256

Tez Danışmanı
Prof. Dr. Ahmet Gökçen

İstanbul, 2005

TEZ ONAY SAYFASI

ÖZ

Bu çalışmanın amacı ülkelerin kalkınmışlık durumlarına etki eden faktörleri incelemektir. Bu analiz farklı yöntemlerle yapılabilirse de bu çalışmada çok seçenekli bağımlı nitel değişkenli modeller ile inceleme yapılmıştır. Ekonometride yaygın olarak kullanılan regresyon modellerinde özel durumlarda özel değişkenler kullanılması gerekmektedir. Bu değişkenlerden biri de nitel değişkenlerdir. Nitel değişkenler modellerde bağımlı değişken olarak yer alabilirler. Nitel değişkenler daha yaygın olarak iki seçenekli olarak kullanılırlar. Bu tür modeller söz konusu olduğunda doğrusal olasılık, logit ve probit modelleri kullanılmaktadır. Kavram birliği sağlamak için bu modeller çalışmanın birinci bölümünü oluşturmaktadır. Nitel değişkenin ikiden fazla değer alması ve bağımlı değişken olması durumunda çok seçenekli doğrusal olasılık, logit ve probit modelleri söz konusu olmaktadır. Bu modeller iki seçenekli modellerden farklı şekilde incelenirler. Çalışmanın ikinci bölümünde bu modeller incelenmiştir. Çalışmanın son bölümünde ise yapılan uygulama ile ülkelerin gelişmişlik durumunu etkileyen faktörler çok seçenekli logit modeli ile belirlenmiştir. Buna göre etkin faktörler Eğitim çağındaki nüfusun öğrenim görme oranı, satınalma gücü paritesi ile düzeltilmiş kişi başına gayri safi milli hasıla, kadınların parlamentodaki oranı'dır.

Abstract

The aim of this study is to research the factors that effects the development level of countries. In this study, the multinomial dependent qualitative model is used, despite the fact that it is possible to use different methods in this kind of analysis. The regression models that are widely used in econometrics, necessitates to employ some special variables in special cases. One of those special variables is qualitative variable. The qualitative variables may take place in the models as dependent variable. The qualitative variables are usually employed as bienary variables. Linear probability models, logit and probit models are used ,when such models are discussed. Those kinds of models costitute the first section, in order to set a bridge to the next section. In case the qualitative variable is a dependent variable having more than two values, it is necessary to use the multinomial linear probability, logit and probit models. The mentioned models are evaluated in a way different from the bienary models. These models are employed and evaluated in the second section. In the last chapter of the thesis, the factors that effects the development level of countries are determined by using multinomial logit model. The results shows that the factors are literacy rate of the population at educational age, GNP per person adjusted by purchasing power parity, and the ratio of female members in the parliament.

ÖNSÖZ

Ülkelerin kalkınmışlık düzeyleri pek çok faktöre bağlıdır ve iktisadi açıdan incelenen önemli konulardan biridir. Kalkınmışlık düzeyi ile ilgili subjektif değerlendirmeler yapılabileceği gibi farklı ölçülerle objektif diyebileceğimiz değerlendirmeler de yapılabilir. Bunun yanında değerlendirme nasıl yapılırsa yapılsın, bu sonucu doğuran faktörler olayın nedeninin araştırılması açısından da önemlidir.

Ülkelerin kalkınmışlık düzeyine etki eden faktörlerin belirlenmesi için ekonometrik çalışma yapılacak ise seçilebilecek farklı yöntemler olabilir. Bu çalışmada ülkelerin gelişmişlik düzeyi üç farklı boyutta gelişmiş ülkeler, gelişmekte olan ülkeler ve az gelişmiş ülkeler olarak değerlendirilerek nitel bağımlı değişkenli modeller ile analizi amaçlanmıştır. Yöntem seçiminde ve ülkelerin kalkınmışlık düzeyinde net ölçüm veya değerlendirme yapılamaması bu çalışmada karşılaşılan sorunlardır.

Seçenek sayısı ikiden fazla olduğunda iki seçenekli nitel bağımlı değişkenler kullanılamayacağından çok seçenekli nitel bağımlı değişkenli modellerin kullanılması gereği ortaya çıkmıştır. Bu tür modeller özellikle bilgisayar teknolojisindeki gelişmeye paralel olarak gelişen modellerdendir.

Çalışmada nitel bağımlı değişkenli modeller konuya açıklık getirip kavram bütünlüğünün sağlanmak amacı ile yer almış, daha sonra çok seçenekli nitel bağımlı değişkenli modeller bu temel kavramlar ışığında, farklılıkları da ortaya koyacak şekilde ele alınmıştır.

Yapılan uygulama ile ülkelerin kalkınmışlık düzeylerini etkileyen faktörler araştırılmıştır. Uygulamada farklı modeller denenmiş yapılan çok sayıda deneme sonucunda uygun model bulunmuştur.

İÇİNDEKİLER

Öz	iii
Abstract	iv
Önsöz	v
İçindekiler	vi
Tablolar ve şekiller	viii
Kısaltmalar	ix
Giriş	1

1.BÖLÜM

İki Seçenekli Nitel Bağımlı Değişkenli Modeller

1.1 Doğrusal Olasılık Modeli	3
1.2 Logit Modeli	7
1.3 Probit Modeli	12
1.4 Doğrusal Olasılık Modeli, Logit Modeli ve Probit Modelinin Karşılaştırılması	16
1.5 Tahmin Sonuçlarının Test Edilmesi	18
1.5.1 Parametrelerin ayrı ayrı anlamlılığının testi	18
1.5.2 Parametrelerin birlikte anlamlılığının testi	18
1.6 Tahminin uygunluğunun ölçüsü	20
1.7 Varsayımların Testi	25
1.7.1 Çoklu doğrusal bağlantı	25
1.7.2 Hata terimlerinin analizi	31
1.7.2.1 Hata terimlerinin normalliliğinin analizi	32
1.7.2.2 Değişen varyans	33

2.BÖLÜM

Çok Seçenekli Nitel Bağımlı Değişkenli Modeller

2.1 Çok Seçenekli Doğrusal Olasılık Modeli	38
2.2 Sıralı Olmayan Çok Seçenekli Tercih Modelleri	41
2.2.1 Çok seçenekli Logit Modeli	42
2.2.2 Çok seçenekli Probit Model	47
2.3 Sıranın Önem Taşıdığı Modeller	50
2.3.1 Sıralı Çok seçenekli Logit Modeli	53
2.3.2 Sıralı Çok Seçenekli Probit Modeli	54
2.4 Ardışık Çok Seçenekli Tercih Modelleri	55
2.5 Çok Seçenekli Modellerin Yorumlanması	56
2.5.1 Fark Oranları	56
2.5.2 Nispi Risk Oranı	57
2.5.3 Marjinal Etkiler	57

3. BÖLÜM

Ülkelerin Kalkınmışlık Düzeylerini Etkileyen Faktörlerin Çok Seçenekli Logit Model ile Analizi

3.1 Kullanılan veri.....	59
3.2 Tahmin Edilen Modelin Sonuçları.....	60
SONUÇ.....	76
KAYNAKÇA	78
EK: Çalışmada kullanılan veriler	83

TABLolar VE ŐEKİLLER

Tablolar

Tablo1:	Normal Birikimli Dağılım Fonksiyonu ile Lojistik Birikimli Dağılım Fonksiyonunun Olasılık Değerleri	17
---------	--	----

Őekiller

Őekil 1:	Doğrusal olasılık modeli	4
Őekil 2:	Logit modeli	9
Őekil 3:	Probit modeli	13

KISALTMALAR

DOM	:	Doğrusal olasılık modeli
EKK	:	En küçük kareler
LR	:	Olabilirlik oranı
LM	:	Lagrange çarpanı
SSR	:	Sum Square of Regression
SST	:	Sum Square Total
VIF	:	Varyans büyütme faktörü
IIA	:	İlişkisiz Alternatiflerin Bağımsızlığı
GSYİH	:	Gayrisafi Yurt İçi Hasıla
GSMH	:	Gayrisafi Milli Hasıla

GİRİŞ

Teknolojik gelişme ve son yıllarda globalleşmenin etkileri ile ekonometrik araştırmalar daha da yaygınlaşmaya ve ekonometrik teknikler daha çok kullanılmaya başlanmıştır. Yaygın kullanım sonucu ekonometrik yöntemler ihtiyaçlara cevap verecek şekilde gelişme göstermektedir. Bilgisayarlar kullanılan ve sürekli geliştirilen paket programlar uygulanması güç yöntemleri uygulanır hale getirirken, diğer yandan teknolojik yöntemler diye adlandırabileceğimiz bilgisayarlar olmadan kullanılması hemen hemen imkansız olan yöntemlerin kullanılmasına da imkan vermektedir.

Son yıllarda ekonometrik araştırmalar ve geliştirilen ekonometrik yöntemler , klasik yöntemler yanında zaman serileri yöntemlerine yönelmiştir. Bunun yanında nitel değişkenler ile yapılan çalışmalarda da gelişmeler olmaktadır. Özellikle nitel bağımlı değişkenler ile ilgili kayda değer gelişmeler olmaktadır. Nitel bağımlı değişkenli modeller denince logit ve probit modelleri akla gelmektedir. Bu modellerin daha gelişmiş hali ise çok seçenekli logit ve probit modellerdir.

Ülkelerin gelişmişlik düzeylerini etkileyen faktörlerin belirlenmesi iktisadi açıdan önemli bir konudur. Bağımlı değişkeni ülkelerin gelişmişlik düzeylerini gösterecek nitel değişken olan logit ve probit modelleri ile ülkeler için gelişmiş, gelişmemiş gibi iki seçenekli ayırım yapılabilirken, çok seçenekli logit ve çok seçenekli probit modelleri ile bu ayırım gelişmiş, gelişmekte olan, az gelişmiş gibi daha çok seçeneğe göre yapılabilecektir. Bu çalışmanın amacı ülkelerin gelişmişlik düzeylerini etkileyen faktörleri çok seçenekli logit modeli ile belirlemektir.

Ülkelerin gelişmişlik düzeylerini etkileyen faktörleri tespit etmek amacı ile yapılan çalışmamızın ilk bölümü, teorik alt yapının oluşturulması ve dil birliğinin sağlanması amacı ile iki seçenekli logit ve probit modellerini ayrılmıştır. Modellerle ilgili temel açıklamaları yapabilmek için, her ne kadar kullanılsalar da önce

doğrusal olasılık modeli incelenmiştir. Çalışmanın ikinci bölümün “Çok Seçenekli Nitel Bağımlı Değişkenli Modeller” başlığını taşımaktadır. Bu bölümde çok seçenekli doğrusal olasılık modeli, çok seçenekli logit ve çok seçenekli probit modelleri açıklanmıştır. Bu bölümde modellerin tahmini ve anlamlılığının testi ile ilgili açıklamaların yanında, modellerin yorumu ile ilgili teorik açıklamalara da yer verilmiştir.

Çalışmanın son bölümü ülkelerin gelişmişlik düzeylerini etkileyen faktörlerin belirlenmesine yönelik uygulamaya ayrılmıştır. Bölümde tesadüfi olarak seçilen 159 ülke ile ilgili 3 değişken ile yapılan uygulamaya yer verilmiştir.

Çalışmamızın sonuç kısmında ise elde edilen teorik sonuçlar ile uygulamanın sonuçları özetlenmiştir.

1.BÖLÜM

İKİ SEÇENEKLİ NİTEL BAĞIMLI DEĞİŞKENLİ MODELLER

Regresyon modellerinde niteliksel değişimleri yansıtan ölçülemeyen fakat niteliksel değişimin meydana gelip gelmemesine göre sayısal olarak ifade edilebilen değişkenlere nitel değişkenler adı verilir. Nitel değişkenler iki durumu belirtebilecekleri gibi evet- hayır- belki gibi ikiden fazla değerde alabilmektedir.

Ekonometrik çalışmalarda nitel değişkenlere bağımsız değişken olarak yer verilebileceği gibi bağımlı değişken olarak yer verilebilmektedir. Nitel değişkenler modelde bağımsız değişken olarak yer aldığında gölge değişken olarak adlandırılır.

Bağımlı değişken konumunda regresyon modellerine dahil edilen nitel değişkenler, evet-hayır, satın alma-almama gibi tercih belirtirler. Bu nedenle bu tür modellere ikili tercih modelleri denir. İki değerden daha fazla değer alabilen nitel bağımlı değişkenli modeller ise çoklu tercih modelleri olarak ifade edilirler.

Uygulamada yaygın olarak kullanılan tercih modelleri doğrusal olasılık modeli, logit modeli ve probit modelidir.

1.1 DOĞRUSAL OLASILIK MODELİ

İki değer alabilen bağımlı değişkenin, bağımsız değişken veya değişkenlerin doğrusal bir fonksiyonu olarak gösterildiği modele doğrusal olasılık modeli denir.

Doğrusal regresyon modeli;

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

olarak tanımlandığında Y_i , 0 veya 1 değerlerini alacaktır.

$Y_i = 1$ ilk seçeneğin tercih edilmesi

$Y_i = 0$ ikinci seçeneğin tercih edilmesi

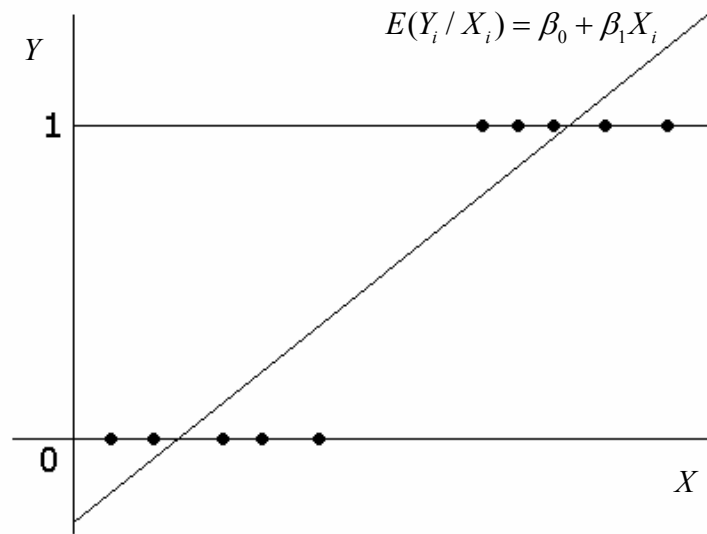
Modeldeki bağımlı değişken iki seçenekten birini tercih etme olasılığını gösterir. O halde ilk seçeneğin tercih edilme olasılığını P_i , ikinci seçeneğin tercih edilme olasılığını $1 - P_i$ olarak ifade ettiğimizde bağımlı değişkenin beklenen değeri

$$\begin{aligned} E(Y) &= 1.(P_i) + 0.(1 - P_i) \\ &= P(Y_i = 1) \\ &= P_i \end{aligned}$$

olacaktır. Böylece ,

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

şeklinde yazılabilir. Genel yapısı belirtilen doğrusal olasılık modeli Şekil 1 de gösterilebilir.



Şekil 1 : Doğrusal Olasılık Modeli

Şekilden de anlaşılabileceği gibi Y_i , 0 ve 1 gibi iki değer alabilmektedir. Değişkenin bu durumu tahmin edilen model açısından birtakım sorunlara neden olacaktır. Bu sorunlar şu şekilde açıklanabilir.

— Doğrusal olasılık modelinde hata terimleri bağımlı değişken gibi iki değer alacaktır.

$$\begin{aligned} Y_i = 1 \quad & \text{ise} \quad 1 = \beta_0 + \beta_1 X_i + \varepsilon_i \rightarrow \varepsilon_i = 1 - \beta_0 - \beta_1 X_i \\ Y_i = 0 \quad & \text{ise} \quad 0 = \beta_0 + \beta_1 X_i + \varepsilon_i \rightarrow \varepsilon_i = -\beta_0 - \beta_1 X_i \end{aligned}$$

Görüldüğü gibi hata terimleri iki değer aldığından normallik varsayımı geçerli olmayacaktır. Hata terimleri binom dağılıma sahiptir. fakat normallik varsayımının geçerli olmaması tahmincilerin sapmasızlığını etkilemeyecektir, yani parametre tahmincileri sapmasız olacaktır. Ayrıca büyük örneklerde binom dağılım normal dağılıma yaklaşmaktadır.

— Doğrusal olasılık modelinde hata terimleri iki değer alabildiğinden değişen varyansa sahiptir. Bunu aşağıdaki şekilde gösterebiliriz.

$$\begin{aligned} Var(\varepsilon_i) &= E(\varepsilon_i^2) = P(Y_i = 0)[- \beta_0 - \beta_1 X_i]^2 + P(Y_i = 1)[1 - \beta_0 - \beta_1 X_i]^2 \\ &= [1 - P(Y_i = 1)][P(Y_i = 1)]^2 + P(Y_i = 1)[1 - P(Y_i = 1)]^2 \\ &= [1 - P(Y_i = 1)][P(Y_i = 1)][P(Y_i = 1)[1 - P(Y_i = 1)]] \\ &= P(Y_i = 1)[1 - P(Y_i = 1)] \\ &= [-\beta_0 - \beta_1 X_i][1 - \beta_0 - \beta_1 X_i] \\ &= E(Y_i)[1 - E(Y_i)] \end{aligned}$$

Değişen varyans sorunu tartılı en küçük kareler yöntemi kullanılarak çözülebilir; fakat bu durumda bile hata terimleri normal dağılmadığı için hipotez testlerine güvenilemez.

Doğrusal olasılık modellerinde hata teriminin ortalaması sıfırdır. $E(\varepsilon_i)$,

$$\begin{aligned} E(\varepsilon_i) &= P(Y_i = 1)[1 - \beta_0 - \beta_1 X_i] + P(Y_i = 0)[- \beta_0 - \beta_1 X_i] \\ &= P(Y_i = 1)[1 - P(Y_i = 1)] + [1 - P(Y_i = 1)]P(Y_i = 1) \\ &= 0 \end{aligned}$$

şeklinde ifade edileceğinden bu varsayım doğrusal olasılık modelinde geçerlidir.

— Doğrusal olasılık modellerinde $E(Y_i/X_i)$ Y_i ‘nin gerçekleşmesinin koşullu olasılığını ölçtüğünden 0 ile 1 arasında değer alması gerekmektedir. Fakat tartılı en küçük kareler yöntemi kullanılsa bile $\hat{Y}_i$ ların 0 ile 1 arasında değer alması garanti edilemez. Bu durumun çözümü için iki yaklaşım ileri sürülmüştür. Bunlardan ilki 0 dan küçük değerleri 0, 1 den büyük değerleri 1 olarak kabullenmektir . Fakat bu durumda tahminler sapmalı olacaktır. İkinci yaklaşım ise modelin katsayılarının $0 \leq E(Y_i / X_i) \leq 1$ sınırlamasına dayalı olarak yeniden tahminidir. Fakat bu durumda tahminlerin sapmasız olacağı garanti edilemeyecektir.¹

Regresyon modellerinde tahminin uygunluğu ölçüsü olarak kullanılan belirlilik katsayısı (R^2) nitel bağımlı değişkenli modellerde genellikle 0.2 ile 0.6 arasında değerler almakta; yani 1’den çok küçük çıkma eğilimindedir. Y yalnızca iki değer aldığından, Y değerleri ya X eksenine yada 1 hizasında ki doğru üzerinde olacaktır. Belirlilik katsayısı hiçbir zaman bu şekilde dağılım gösteren verilere uyum sağlayamayacaktır.

¹ Hüseyin Özer, **Nitel Değişkenli Ekonometrik Modeller Teori ve Uygulama**, Ankara, Nobel Yayın Dağıtım, 2004, s.67.

Belirttiğimiz bütün bu sorunlar çözülebilir niteliktedir. Fakat doğrusal olasılık modelinin en büyük sorunu dayandığı mantıktan kaynaklanmaktadır. Doğrusal olasılık modeli P_i değerini X_i değerinin doğrusal bir fonksiyonu olduğunu varsaymaktadır. Yani marjinal etkiyi sabit varsaymaktadır. Bir başka deyişle X değişkeninde meydana gelecek her değişme Y değişkeni üzerinde aynı oranda etki yapacağını varsaymaktadır. Bu varsayım da gerçeklikten uzak niteliktedir. Bu sorun ise doğrusal olasılık modeli için aşılabilir bir sorundur ve doğrusal olasılık modelinin kullanılmasını anlamsız kılmaktadır.

Doğrusal olasılık modeli bu aşılabilir sorunu taşıdığından bu model yerine P_i 'nin X_i 'nin doğrusal bir fonksiyonu olmadığı ve P_i 'nin 0 ile 1 aralığının dışına çıkmayacak şekilde hareket ettiği olasılık modellerini kullanabiliriz. Lojistik dağılım fonksiyonu ve normal birikimli dağılım fonksiyonları bu özellikleri taşıyan ve uygulamada sıklıkla kullanılan fonksiyonlardır.

1.2 LOGİT MODELİ

Doğrusal olasılık modelinde karşılaşılan $P_i (\hat{Y})$ 'nin 0 ile 1 aralığı dışında değerler alması sorununu çözmek için kullanılan logit modeli Lojistik dağılım fonksiyonuna dayanmaktadır. Lojistik dağılım fonksiyonunu;

$$P_i = F(Z_i) = \frac{1}{1 + e^{-Z_i}}$$

şeklinde tanımlanabilir.

$$Z_i = \beta_1 + \beta_2 X_i$$

olarak ifade edildiğinde lojistik dağılım fonksiyonu;

$$P_i = \frac{1}{1 + e^{-(\beta_1 + \beta_2 X_i)}}$$

olarak gösterilebilir. P_i açıklayıcı değişken X_i hakkında bilgi veri iken i 'inci bireyin belirli bir seçimi yapma olasılığıdır.

Z_i , $-\infty$ ila $+\infty$ arasında değerler aldığı anda, olasılığı ifade eden P_i 'nin 0 ile 1 arasında değerler aldığı görülmekte ve bu sayede de DOM'nin en büyük problemlerinden biri çözülmektedir. Fakat, bu problem çözülmesine rağmen bu kez de karşımıza bir tahmin sorunu çıkar. Lojistik dağılım fonksiyonunu tanımladığımız yukarıdaki ifadede de görüldüğü gibi, P_i 'nin X ve β lar ile olan ilişkisi doğrusal değildir. Bu tahminlerimiz sırasında en küçük kareler metodunu kullanamayacağımız anlamına gelmektedir.

P_i bir olayın gerçekleşme olasılığı ise $1-P_i$ olayın gerçekleşmeme olasılığıdır. O halde

$$1 - P_i = \frac{1}{1 + e^{Z_i}}$$

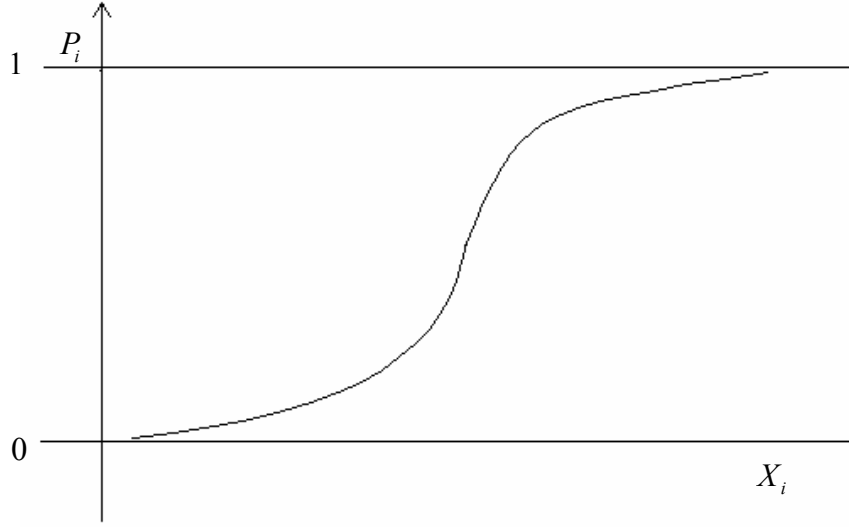
şeklinde ifade edilebilir. Dolayısıyla;

$$\frac{P_i}{1 - P_i} = \frac{1 + e^{Z_i}}{1 + e^{-Z_i}} = e^{Z_i}$$

$P_i / (1 - P_i)$ oranı bahis oranı olarak adlandırılır. Yukarıda ki ifadenin doğal logaritması alınırsa ;

$$L_i = \ln\left(\frac{P_i}{1 - P_i}\right) = Z_i = \beta_1 + \beta_2 X_i$$

olduğu görülmektedir. L_i , X_i ve β lara göre doğrusaldır. Burada L 'ye logit adı verilir. Logit modelin grafiği aşağıdaki şekilde gösterilebilir.



Şekil 2: Logit Model

Logit modeli tahmin edilirken P_i yerine 0 veya 1 değerleri konulduğunda logit değerlerinin anlamsız çıktığı görülmektedir. Bu sebepten verilerin tek tek alındığı basit frekanssız serilerde logit modelin EKK yöntemi ile tahmini mümkün değildir. Bu durumda en çok benzerlik yöntemi ile parametre tahminleri gerçekleştirilir.² Frekanslı serilerde ise lojistik modelin EKK yöntemi ile tahmin edilmesi mümkündür. Frekanslı seriler için tahmin aşamaları şu şekilde sıralanabilir.

Öncelikle $\hat{P}_i$ hesaplanır.

$$\hat{P}_i = \frac{n_i}{N_i} \text{ 'dir}$$

² Şahin Akkaya, **Ekonometri 2**, İzmir, Anadolu Matbaacılık, 1991, sf.84.

Burada N_i frekanstaki gözlem sayısını, n_i ise frekans içindeki istenen olayın gerçekleştiği gözlem sayısını ifade eder. Hesaplanan $\hat{P}_i$ değeri kullanılarak her frekans için $\hat{L}_i$ değerleri,

$$\hat{L}_i = \ln[\hat{P}_i / (1 - \hat{P}_i)]$$

şeklinde hesaplanır.

Logit modelinde hata terimleri binom dağılımlı olduğundan değişen varyans sorunu ile karşı karşıya kalınır. Bu sorunu gidermek için tartılı EKK yönteminden yararlanılır.

Tartı, $w_i = N_i \hat{P}_i (1 - \hat{P}_i)$ olarak hesaplanır. Hesaplanan tartı değişkenlerle çarpılarak

$$\sqrt{w_i} L_i = \beta_1 \sqrt{w_i} + \beta_2 \sqrt{w_i} X_i + \sqrt{w_i} u_i$$

şekline dönüştürülür. Bu dönüşüm işleminden sonra regresyon EKK yöntemi ile çözümlenir.

Tahmin edilen parametrelerden β_2 X ‘teki bir birimlik değişmeye karşılık L deki değişmeyi ölçer. Yani log-bahis oranının nasıl değiştiğini belirtir. Sabit parametre ise $X_i = 0$ olduğunda Y_i ’nin log bahis oranını verecektir. Frekanslı serilerde kullanılan bu yöntem minimum kare yöntemi olarak adlandırılır.

Frekanssız serilerde ise lojistik fonksiyonun tahmini en çok benzerlik yöntemi ile yapılacaktır.

Başarı olasılığı P_i başarısızlık oranı $1 - P_i$ olduğunda benzerlik fonksiyonu

$$L(Y / X) = P(Y_i / X) = \sum_{i=1}^n P_i^{Y_i} (1 - P_i)^{1-Y_i}$$

şeklinde ifade edilebilir. Burada benzerlik fonksiyonunu maksimum yapacak $\hat{\beta}$ katsayılar vektörünün belirlenmesi amaçlanmaktadır. Bunun gerçekleştirilmesi için öncelikle benzerlik fonksiyonunun logaritması alınacaktır.

$$L(Y / X, \beta) = \sum_{i=1}^n (Y_i \log P_i + (1 - Y_i) \log(1 - P_i))$$

Bu ifade üzerinde gerekli düzenlemeler yapılarak,

$$\begin{aligned} L &= \sum_{i=1}^n [Y_i (\log P_i - \log(1 - P_i)) + \log(1 - P_i)] \\ &= \sum_{i=1}^n [Y_i \log \left(\frac{P_i}{1 - P_i} \right) + \log(1 - P_i)] \\ &= \sum_{i=1}^n \left[Y_i (\beta_0 + \beta_1 X_i) + \log \left(\frac{1}{1 + e^{\beta_0 + \beta_1 X_i}} \right) \right] \\ &= \sum_{i=1}^n \left[Y_i (\beta_0 + \beta_1 X_i) + \log(1 + e^{\beta_0 + \beta_1 X_i})^{-1} \right] \\ &= \sum_{i=1}^n \left[Y_i (\beta_0 + \beta_1 X_i) - \log(1 + e^{\beta_0 + \beta_1 X_i}) \right] \end{aligned}$$

şeklinde yazılabilir.

Fonksiyonu maksimize edecek parametreleri tahmin etmek için en çok benzerlik fonksiyonunun birinci türevi alınmalıdır.

$$\frac{\partial L}{\partial \beta_0} = \sum \left(Y_i - \hat{P}_i \right) = 0$$

$$\frac{\partial L}{\partial \beta_1} = \sum \left(Y_i - \hat{P}_i \right) X_i = 0$$

Burada $\hat{P}_i$,

$$\hat{P}_i = \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 X_i}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_1 X_i}}$$

şeklinde ifade edilebilir. Bu denklemin çözümü ile $\hat{\beta}_0$ ve $\hat{\beta}_1$ tahmin edilebilir.

En çok benzerlik yöntemi tahminçileri küçük örneklerde sapmasız, etkin olmayan ve normal dağılmayan ; büyük örneklerde ise tutarlı asimptotik etkin ve asimptotik normal dağılımlı tahminçilerdir.³

1.3 PROBİT MODELİ

Doğrusal olasılık modelinin en büyük sorunlarından olan $P_i(\hat{Y})$ 'nin 0 ile 1 arasında değerler alamaması sorununu çözmek için kullanılan bir diğer model ise birikimli normal dağılım fonksiyonundan yararlanılarak oluşturulan probit modelidir.

Rasyonel seçim teorisine göre birimler karşılaştıkları alternatifler arasından kendileri için en çok fayda sağlayacak olanı seçerler. Probit modeli fayda teorisine ve rasyonel seçim yaklaşımına dayanmaktadır.⁴ Normal birikimli dağılım fonksiyonu;

³ Selahattin Güriş, Ebru Çağlayan, **Ekonometri Temel Kavramlar**, 2. bs., İstanbul, DER, 2005, s.687.

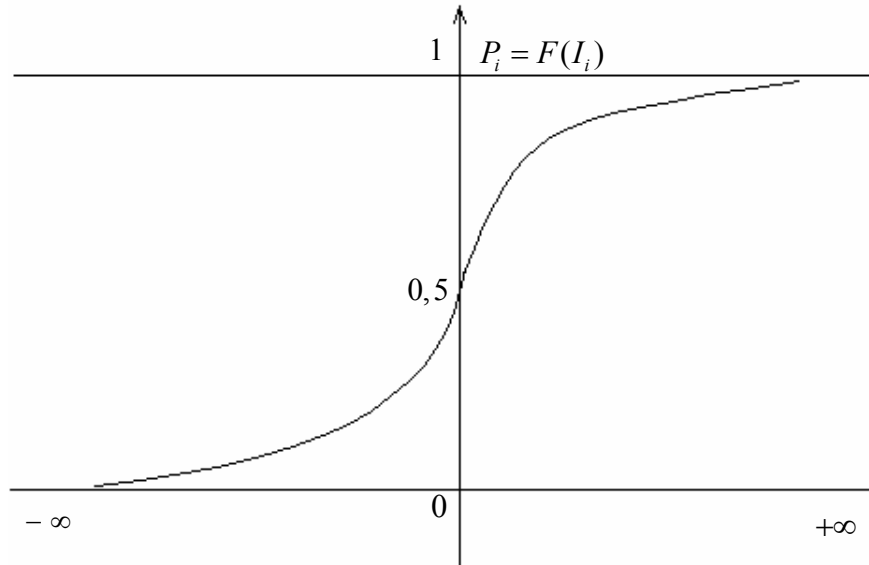
⁴ A.e., s.685.

$$P_i = F(I_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta_1 + \beta_2 X_i} e^{-t^2/2} dt$$

olarak tanımlanır. Burada t ortalaması 0 ve varyansı 1 olan normal dağılımlı bir tesadüfî değişken, I_i gözlenemeyen fayda endeksidir ve şu şekilde ifade edilir;

$$I_i = \beta_1 + \beta_2 X_i$$

$P_i < 0.5$ olduğunda I_i her zaman eksi değer aldığından uygulamada I_i 'ye 5 sayısı eklenir, sonuca da probit denir.⁵ Probit modelin grafiği aşağıdaki şekilde gösterilebilir.



Şekil 3: Probit Model

Logit modelde olduğu gibi frekanslı ve frekanssız serilerde probit modelin tahmini farklılık gösterecektir. Frekanslı serilerde probit modelin tahmin aşamaları şu şekilde sıralanabilir. Öncelikle logit modeldeki gibi $\hat{P}_i$ hesaplanır. Sonrasında bu $\hat{P}_i$,

⁵ Damador N. Gujarati, **Temel Ekonometri**, Çev. Ümit Şenesen, Gülay Günlük Şenesen, İstanbul, Literatür, 1999, s.566.

lere karşılık gelen I_i 'ler normal birikimli dağılım fonksiyonundan bulunur. Buradan bulunan $\hat{I}$ lere 5 eklenip probitlere çevrilerek , $I_i = \beta_1 + \beta_2 X_i + u_i$ regresyonunda bağımlı değişken olarak kullanılır. Model EKK aracılığıyla çözülür. Fakat modelin hata terimleri değişen varyanslıdır. Bu sorunun çözümü için değişkenlere

$$\sigma_u^2 = \frac{P_i(1-P_i)}{N_i f_i^2}$$

dönüşümü uygulanır. Burada f_i , $F^{-1}(P_i)$ ' de bulunan standart normal yoğunluk fonksiyonudur. Bu tahmin yöntemine tartılı en küçük kareler yöntemi denir.

Frekanssız serilerde ise probit modelin tahmini en çok benzerlik yöntemi ile yapılabilir. Benzerlik fonksiyonu

$$L = \sum [Y_i \log F(\beta_0 + \beta_1 X_i) + (1 - Y_i) \log [1 - F(\beta_0 + \beta_1 X_i)]]$$

olacaktır. Kısmi türevler alınarak logit modeldeki gibi L' yi maksimize edecek β_0 ve β_1 parametreleri tahmin edilecektir.

X_i ' nin farklı değerleri için Y' nin tekrarları söz konusu ise

$$F^{-1}(P_i) = F^{-1}(P_i + \varepsilon)$$

olarak yazılacaktır. Bu ifade

$$F^{-1}(P_i) = F^{-1}(P_i) + \frac{\partial F^{-1}(P_i)}{\partial P_i} + R_i$$

burada R_i Taylor açılımının kalan terimlerinin toplamıdır ve ihmal edilebilir niteliktedir. Sonrasında

$$\begin{aligned}\frac{\partial F^{-1}(P_i)}{\partial (P_i)} &= \frac{\partial (\beta_0 + \beta_1 X_i)}{\partial F(\beta_0 + \beta_1 X_i)} \\ &= \frac{1}{\partial F(\beta_0 + \beta_1 X_i) / \partial (\beta_0 + \beta_1 X_i)}\end{aligned}$$

paydada yer alan ifade $F(\beta_0 + \beta_1 X_i)$ standart normal dağılım fonksiyonuna eşittir.

Bu nedenle

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

büyük örnekler için

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

olacaktır. β_0 ve β_1 yerine en küçük kareler tahmincileri $\hat{\beta}_0, \hat{\beta}_1$ ve P_i değerleri kullanılarak tahmin yapılabilir. Elde edilen tahminler tutarlıdır. Hata teriminin beklenen değeri

$$E\left[\frac{\varepsilon_i}{f(\beta_0 + \beta_1 X_i)}\right] = 0$$

ve varyanslar

$$\begin{aligned}
& Var\left[\frac{\varepsilon_i}{f(\beta_0 + \beta_1 X_i)}\right] \\
&= Var\left[\frac{P_i}{f(\beta_0 + \beta_1 X_i)}\right] \\
&= \frac{P_i(1-P_i)}{n_i [f(\beta_0 + \beta_1 X_i)]^2}
\end{aligned}$$

olacaktır. En çok benzerlik tahminçileri bütün asimptotik özelliklere sahiptir.⁶

1.4 DOĞRUSAL OLASILIK MODELİ, LOGİT MODEL VE PROBİT MODELİN KARŞILAŞTIRILMASI

İkili tercih modellerinden en basit olanı doğrusal olasılık modelidir. Doğrusal olasılık modelinde ilişki doğrusal olduğundan en küçük kareler metodu ile çözülebilmesi bu basitliğin en önemli sebebidir. Doğrusal olasılık modelleri; hata terimlerinin iki değer almasından dolayı değişen varyansa sahip olması, $\hat{Y}_i$ ların 0 ile 1 aralığının dışına düşmesi, R^2 nin 1 den çok küçük çıkma eğiliminde olması, P_i değerinin X_i değerinin doğrusal bir fonksiyonu olarak kabul edilmesi gibi sorunları taşımaktadır. Özellikle ilişkinin doğrusal kabul edilmesi yani marjinal etkinin sabit olarak nitelendirilmesi bu sorunların en önemlisi ve aşılabilir nitelikte olanıdır.

Bu sorunun giderilmesi açısından logit ve probit modeller geliştirilmiştir. Logit model lojistik dağılım fonksiyonu temeline dayanırken probit model ise birikimli normal dağılım fonksiyonu temeline dayanmaktadır. Lojistik dağılım fonksiyonu ve normal birikimli dağılım fonksiyonu birbirine çok yakın fonksiyonlardır. Bu aşağıdaki tablodan daha net şekilde anlaşılabilir.

⁶ Jan Kmenta, **Elements of Econometrics**, Second Edition, Maxwell Macmillan International Edition, NewYork, 1990, pp.555.

Tablo1 : Normal Birikimli Dağılım Fonksiyonu ile Lojistik Birikimli Dağılım Fonksiyonunun Olasılık Değerleri

X	Normal Birikimli Dağılım Fonksiyonu	Lojistik Birikimli Dağılım Fonksiyonu
-3.00	0.0013	0.0474
-2.00	0.0228	0.1192
-1.50	0.0668	0.1824
-1.00	0.1587	0.2689
-0.50	0.3085	0.3775
0.00	0.5000	0.5000
0.50	0.6915	0.6225
1.00	0.8413	0.7311
1.50	0.9332	0.8176
2.00	0.9772	0.8808
3.00	0.9987	0.9526

İki fonksiyon birbirine yakın olduğu için Logit ve Probit modellerde birbirlerine oldukça benzemektedir. Normal birikimli dağılım fonksiyonu integraller içerdiği için uygulama sırasındaki hesaplamalarda zorluklar taşıdığından logit model, probit modele nazaran daha sık kullanılmaktadır.

Standart normal birikimli dağılımın ortalaması 0 ve varyansı 1 iken , lojistik dağılımın ortalaması 0 ve varyansı $\pi^2/3$ ’ tür. Yani logit ve probit modellerin sonuçlarını karşılaştırmak için logit modelden elde edilen katsayıların $\sqrt{3}/\pi$ ile çarpılması gerekecektir.⁷

⁷ Özer, a.g.e., s.80.

1.5 TAHMİN SONUÇLARININ TEST EDİLMESİ

Nitel bağımlı değişkene sahip modellerde, klasik regresyon modellerinde de olduğu gibi parametrelerin anlamlılığının test edilmesi gerekecektir.

1.5.1 Parametrelerin ayrı ayrı anlamlılığının testi

Parametrelerin ayrı ayrı anlamlılığının test edilmesi bilinen t testi ile gerçekleştirilmektedir.⁸ Test her parametre için tek tek uygulanacaktır. t test istatistiği,

$$t = \frac{\hat{\beta}_i}{\sigma_{\hat{\beta}_i}}$$

olarak hesaplanacaktır. Temel hipotez katsayının sıfıra eşit olduğu hipotezidir. Hesaplanan t istatistiği t tablo değerinden büyük ise temel hipotez reddedilir. Bu durumda ilgili katsayının istatistiksel olarak anlamlı olduğu sonucuna varılacaktır.

1.5.2 Parametrelerin birlikte anlamlılığının testi

t testinde parametrelerin ayrı ayrı anlamlılığı sözkonusu iken burada parametrelerin aynı anda birlikte anlamlı olup olmadığı test edilecektir. Klasik regresyonun aksine logit ve probit modelleri için F testi yerine Wald, lagrange çarpanı ve olabilirlik oranı testleri kullanılmaktadır.⁹ Wald testi sadece

⁸ William E. Griffiths, R. Carter Hill, George G. Judge, **Learning and Practicing Econometrics**, Canada, John Wiley and Sons, 1993, pp.748.

⁹ Özer, **a.g.e.**, s.159.

kısıtlandırılmamış modelin tahmin sonuçlarına dayanırken, lagrange çarpanı testi yalnızca kısıtlandırılmış modelin tahmin sonuçlarını dikkate almaktadır. Olabilirlik oranı testi ise hem kısıtlandırılmış hem de kısıtlandırılmamış modellerin sonuçlarına dayanmaktadır.

Logit ve probit modelleri için F testi yerine kullanılabilen olabilirlik oranı(LR)¹⁰ test istatistiği

$$OO = -2(l_R - l_U)$$

şeklinde gösterilebilir. Burada l_R , kısıtlı modelin en yüksek olabilirlik değerinin logaritmasını bir başka deyişle modelde bağımsız değişken olmaması durumunda benzerlik fonksiyonunun alacağı maksimum değerdir. l_U ise kısıtsız modelin en yüksek olabilirlik değerinin logaritması yani tahmin edilen model için benzerlik fonksiyonu değerini göstermektedir.

Bu test istatistiği h serbestlik derecesi ile ki kare dağılımına uyar. Buradaki h kısıtlama sayısıdır. Hesaplanan test istatistiği belirlenen anlamlılık düzeyinde ki kare tablo değerinden büyük olduğunda kısıtların geçerliliğini ifade eden temel hipotez reddedilecek yani katsayıların birlikte istatistiksel olarak anlamlı olacağına karar verilecektir.

Lagrange çarpanı testinin temelleri; kısıtlamaların geçerli olmaması durumunda lagrange çarpanını değeri yeterince büyükse, kısıtlamaların reddine dayanır. Lagrange çarpanı test istatistiği F testi gibi tüm bağımsız değişken katsayılarının sıfırdan farklı olup olmadığını belirlemek için kullanıldığında

$$LM = n R_U^2$$

¹⁰ Kenneth A. Small, Cheng Hsiao, "Multinomial Logit Specification Tests," **International Economic Review**, Vol:26 No:3, Ekim 1985, pp.622-623.

olarak hesaplanır. Burada R_U^2 kısıtsız modelin belirlilik katsayısını ifade eder. Test istatistiği (k-1) serbestlik derecesi ile ki kare dağılımına uyar. Hesaplanan test istatistiği ki kare tablo değerinden büyük olduğunda katsayıların birlikte istatistiksel olarak anlamlı olacağına karar verilecektir.

Wald testi istatistiği kısıtsız tahmini gerektirmektedir.¹¹ Wald test istatistiği

$$W = \frac{SSE_R - SSE_U}{\hat{\sigma}_e^2}$$

olarak gösterilebilir. Burada ki $\hat{\sigma}_e^2$

$$\hat{\sigma}_e^2 = \frac{SSE_U}{n}$$

şeklinde ifade edilebilir

Kısıt sayısı 1 olacağından test istatistiğinin dağılımı 1 serbestlik dereceli ki kare dağılımına uyacaktır. Test istatistiği, ki kare tablo değeri ile karşılaştırılarak parametrelerin anlamlılığı hakkında sonuca varılabilecektir. Hesaplanan wald testi istatistiği ki kare tablo değerinden büyük olduğunda katsayıların birlikte istatistiksel olarak anlamlı olacağına karar verilecektir.

1.6 TAHMİNİN UYGUNLUĞUNUN ÖLÇÜSÜ

Tahminin uygunluğunun ölçüsü modelin gözlenen veriye uygunluğunun ölçüsünü gösterir. Doğrusal regresyonda R^2 ölçüsüyle gösterilir ve bağımlı

¹¹ R.L. Thomas, **Modern Econometrics an İntruduction**, England, Addison Wesley, 1997, pp.258.

değişkenlerdeki değişmelerin bağımsız değişkenler tarafından açıklanma oranını verir. Aşağıdaki şekilde gösterilebilir.

$$R^2 = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} = \frac{SSR}{SST}$$

Burada SSR(Sum Square of Regression) açıklanan değişmeyi, SST (Sum Square Total) toplam değişmeyi ifade etmektedir.

Bu bölümde tahminin uygunluğu ölçüsünü bireysel ve gruplandırılmış veriler için ayrı ayrı inceleyebiliriz. Bireysel veriler için direkt R^2 veya likelihood Ratio Testten elde edilen pseudo R^2 ' si hesaplanabilir. Direkt R^2 sadece iki uçlu bağımlı değişkenlerin olduğu durumlarda kullanılabilir. Gerçek değer 0 veya 1 iken tahmin edilen olasılık değerini gösterir.¹²

İki uçlu bağımlı değişken ve olasılık tahmincileri arasındaki korelasyon Morrison(1972) tarafından çalışılmıştır. Morrison tahmincilerin beta dağılımına sahip olduğu varsayımı altında R^2 nin üst sınırını bulmuştur. R^2 nin üst sınırı aşağıdaki şekilde elde edilebilir.

Y , sırasıyla p ve $1-p$ olasılıklarla 0 ve 1 değerlerini alan tesadüfi değişken olsun. Burada p , açıklayıcı değişkenin bilinen fonksiyonu , $V(p)$ varyansı, $E(p)$ beklenen değeri , ve $f(p)$ yoğunluk fonksiyonu ile birim aralıkta dağılır. Y ve p arasındaki R^2 ,

$$R^2 = \frac{[Cov(Y, p)]^2}{Var(Y).Var(p)}$$

olarak ifade edilebilir. Burada $E(Y)=E(p)$ ve $E(Y^2)=E(p)$ dir. Buradan

¹² G.S. Maddala, **Limited Dependent and Qualitative Variables in Econometrics**, Cambridge, Cambridge University Press, 1983, pp.50.

$E(Y,p)=E(p^2)$ olur. Böylece

$$\begin{aligned} Cov(Y, p) &= E(Y, p) - E(Y).E(p) \\ &= E(p^2) - E^2(p) = V(p) \end{aligned}$$

olur. Böylece

$$\begin{aligned} R^2 &= \frac{[V(p)]^2}{V(p).Var(Y)} \\ &= \frac{V(p)}{V(Y)} \\ &= \frac{V(p)}{E(p).E^2(p)} \end{aligned}$$

Eğer $f(p)$ α ve β parametreleri ile beta dağılımına sahip ise

$$\begin{aligned} E(p) &= \frac{\alpha}{(\alpha + \beta)} \\ V(p) &= \frac{\alpha\beta}{[(\alpha + \beta + 1)(\alpha + \beta)^2]} \end{aligned}$$

olacaktır. Böylece

$$R^2 = \frac{1}{(\alpha + \beta + 1)}$$

olarak yazılabilir.

Şu ana kadar uygunluk ölçüsü ile ilgili yaptığımız açıklamalar sadece iki uçlu bağımlı değişkenli modeller için geçerlidir. Çok seçenekli logit model için benzerlik

oranı test istatistiğine dayalı olarak tahmin edilen ölçüyü kullanmak daha uygun olacaktır.¹³

Doğrusal regresyon modeli

$$y = \alpha + \sum_{j=1}^m \beta_j X_j + u$$

L_{Ω} bütün parametreler açısından maksimizasyon sağlandığında benzerlik fonksiyonunun maksimizasyonu olsun ve L_w sadece α açısından maksimizasyon sağlandığında benzerlik fonksiyonu maksimum olur. Daha sonra uygunluk ölçüsünü

$$R^2 = 1 - \left(\frac{L_w}{L_{\Omega}} \right)^{2/n}$$

şeklinde ifade edebiliriz. Buradaki n örnek büyüklüğüdür. Çok seçenekli logit model için L_{Ω} bütün parametreler açısından maksimizasyon sağlandığında benzerlik fonksiyonunun maksimumunu gösterir ve L_w sadece sabit parametre α_j açısından maksimizasyon sağlandığında maksimum olur.

n_j j inci kategorideki gözlem sayısı olduğunda

$$\exp(\hat{\alpha}_j) = n_j / n_m \quad j = 1, 2, \dots, m-1$$

olur. Böylece

¹³ A.e., s.39.

$$L_W = \prod_{j=1}^m \left(\frac{n_j}{n} \right)^{n_j}$$

olacaktır. $R^2 = 1 - \left(\frac{L_W}{L_\Omega} \right)^{2/n}$ ölçüsünün logitte kullanımı düşünülebilir fakat bu durumda $R^2 \rightarrow 0$ olsa bile R^2 nin üst sınırı 1 den çok küçük olacaktır. Bunun sebebi benzerlik fonksiyonunun 1 in tam maksimumunu göstermeyi başarmasıdır.

Böylece L nin maksimum değerini şu şekilde yazabiliriz.

$$L_{\max} = 1$$

böylece

$$\prod \left(\frac{n_j}{n} \right)^{n_j} = L_W \leq L_\Omega \leq L_{\max} = 1$$

veya

$$\frac{L_W}{L_{\max}} \leq \frac{L_W}{L_\Omega} \leq 1$$

olacaktır. Böylece R^2 ; $R^2 = 1 - \left(\frac{L_W}{L_\Omega} \right)^{2/n}$ tarafından sınırlandırıldığından bu aralık içinde bulunmalıdır.

$$0 \leq R^2 \leq 1 - (L_W)^{2/n}$$

Bu tahminin uygunluk ölçüsü olarak iyi bir ölçüm sağlamaz. Çünkü tahmin kusursuz olsa bile R^2 1 den küçük olacaktır.¹⁴

Daha iyi bir ölçü ise pseudo- R^2 'sidir ve aşağıdaki şekilde belirlenir.

$$\begin{aligned} pseudo-R^2 &= \frac{1 - (L_W / L_\Omega)^{2/n}}{1 - (L_W / L_{\max})^{2/n}} \\ &= \frac{L_\Omega^{2/n} - L_W^{2/n}}{1 - L_W^{2/n}} \end{aligned}$$

Burada L_W ; $L_W = \prod_{j=1}^m \left(\frac{n_j}{n} \right)^{n_j}$, ten sağlanmaktadır. $L_{\max}=1$ ve L_Ω benzerlik fonksiyonunun maksimumunu sağlanarak elde edilmektedir.

1.7 VARSAYIMLARIN TESTİ

Bu bölümde nitel bağımlı değişkenli modellerin varsayımlarının test edilmesi incelenecektir.

1.7.1 Çoklu Doğrusal Bağlantı

Regresyon modellerinde bağımsız değişkenlerden ikisi veya daha fazlası arasında doğrusal ilişki olması durumu çoklu doğrusal bağlantı olarak adlandırılır. Fakat uygulamada birçok değişkenin çoğunlukla birbiriyle ilişkili olduğu gözlenmektedir. Değişkenler arasındaki doğrusal ilişkinin derecesi doğrusal korelasyon katsayısı ile belirlenmektedir. Bu nedenle çoklu doğrusal bağlantı doğrusal korelasyon katsayısı ile ilişkilidir.

¹⁴ A.e., s.40.

Çoklu doğrusal bağlantının sonuçları değişkenler arasındaki ilişkinin derecesine göre farklılık gösterecektir. Tam çoklu doğrusal bağlantı olması durumunda

$$r_{23} = \frac{\sum (X_{i2} - \bar{X}_2)(X_{i3} - \bar{X}_3)}{\sqrt{\sum (X_{i2} - \bar{X}_2)^2 \sum (X_{i3} - \bar{X}_3)^2}}$$

$r_{23} = \pm 1$ olacaktır. Yani bağımsız değişkenler arasında

$$X_{i3} = \alpha X_{i2}$$

gibi bir ilişki varsa

$$\bar{X}_3 = \alpha \bar{X}_2$$

olacaktır. Bu durumda

$$\begin{aligned} r_{23} &= \frac{\sum (X_{i2} - \bar{X}_2)(\alpha X_{i2} - \alpha \bar{X}_2)}{\sqrt{\sum (X_{i2} - \bar{X}_2)^2 \sum (\alpha X_{i2} - \alpha \bar{X}_2)^2}} \\ &= \frac{\alpha \sum (X_{i2} - \bar{X}_2)^2}{\sqrt{\alpha^2 \left[\sum (X_{i2} - \bar{X}_2)^2 \right]^2}} \\ &= \frac{\alpha \sum (X_{i2} - \bar{X}_2)^2}{\alpha \sum (X_{i2} - \bar{X}_2)^2} \\ &= 1 \end{aligned}$$

olacaktır. Parametre tahmincilerinin durumlarına bakıldığında

$$\begin{bmatrix} \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} \sum (X_{i2} - \bar{X}_2)^2 & \sum (X_{i2} - \bar{X}_2)(X_{i3} - \bar{X}_3) \\ \sum (X_{i3} - \bar{X}_3)(X_{i2} - \bar{X}_2) & \sum (X_{i3} - \bar{X}_3)^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum (Y_i - \bar{Y})(X_{i2} - \bar{X}_2) \\ \sum (Y_i - \bar{Y})(X_{i3} - \bar{X}_3) \end{bmatrix}$$

eşitliğinde

$X_{i3} = \alpha X_{i2}$ ve $\bar{X}_3 = \alpha \bar{X}_2$ yerine konulduğunda

$$\sum (X_{i3} - \bar{X}_3)^2 = \sum (\alpha X_{i2} - \alpha \bar{X}_2)^2 = \alpha^2 \sum (X_{i2} - \bar{X}_2)^2$$

$$\begin{aligned} \sum (X_{i2} - \bar{X}_2)(X_{i3} - \bar{X}_3) &= \sum (X_{i2} - \bar{X}_2) \sum (\alpha X_{i2} - \alpha \bar{X}_2) \\ &= \alpha \sum (X_{i2} - \bar{X}_2)^2 \end{aligned}$$

bu durumda $X'X$ matrisinin tersi almak istediğimizde

$$\begin{aligned} (X'X)^{-1} &= \begin{bmatrix} \sum (X_{i2} - \bar{X}_2)^2 & \alpha \sum (X_{i2} - \bar{X}_2)^2 \\ \alpha \sum (X_{i2} - \bar{X}_2)^2 & \alpha^2 \sum (X_{i2} - \bar{X}_2)^2 \end{bmatrix}^{-1} \\ &= \sum (X_{i2} - \bar{X}_2)^2 \alpha^2 \sum (X_{i2} - \bar{X}_2)^2 - \alpha \sum (X_{i2} - \bar{X}_2) \alpha \sum (X_{i2} - \bar{X}_2)^2 \\ &= \alpha^2 \left[\sum (X_{i2} - \bar{X}_2)^2 \right]^2 - \alpha^2 \left[\sum (X_{i2} - \bar{X}_2)^2 \right]^2 \\ &= 0 \end{aligned}$$

sonucuna varılacaktır. Buradan da görülebildiği gibi tam çoklu doğrusallığın olduğu durumda matrisin tersi alınamadığının parametre tahmini yapılamayacaktır. Parametre tahmini yapılamamasının avantajı durumun uygulama sırasında anlaşılıp önlem alınmasına olanak tanınmasıdır.

Değişkenler arasında kuvvetli ilişki olması durumunu incelediğimizde, daha önce değişkenler arasında tam ilişkinin olması durumunda söz ettiğimiz $X_{i3} = \alpha X_{i2}$ eşitliği farklılaşacaktır. İlişkinin tam olmamasından dolayı eşitliği

$$X_{i3} = \alpha X_{i2} + u_i$$

şeklinde ifade etmemiz gerekmektedir. Bu durumda

$$\begin{aligned} \sum (X_{i3} - \bar{X}_3)^2 &= \sum (\alpha X_{i2} + u - \alpha \bar{X}_2 - \bar{u})^2 \\ &= \sum [\alpha (X_{i2} - \bar{X}_2) + (u_i - \bar{u})]^2 \end{aligned}$$

ve

$$\bar{u} = 0$$

ise

$$\alpha^2 \sum (X_{i2} - \bar{X}_2)^2 + \sum u_i^2 + 2\alpha \sum (X_{i2} - \bar{X}_2)u_i$$

olacaktır. Buradan

$$\sum (X_{i2} - \bar{X}_2)u_i = 0$$

olacağından

$$\sum (X_{i3} - \bar{X}_3)^2 = \alpha^2 \sum (X_{i2} - \bar{X}_2)^2 + \sum u_i^2$$

şeklinde ifade edilecektir.

$$\begin{aligned}
\sum (X_{i2} - \bar{X}_2)(X_{i3} - \bar{X}_3) &= \sum (X_{i2} - \bar{X}_2)(\alpha X_{i2} + u_i - \alpha \bar{X}_2 - \bar{u}) \\
&= \sum (X_{i2} - \bar{X}_2) [\alpha (X_{i2} - \bar{X}_2) + (u_i - \bar{u})] \\
&= \sum (X_{i2} - \bar{X}_2) [\alpha (X_{i2} - \bar{X}_2) + u_i] \\
&= \alpha \sum (X_{i2} - \bar{X}_2)^2 + \sum (X_{i2} - \bar{X}_2) u_i \\
&= \alpha \sum (X_{i2} - \bar{X}_2)^2
\end{aligned}$$

$$r_{23} = \frac{\alpha \sum (X_{i2} - \bar{X}_2)^2}{\sqrt{\sum (X_{i2} - \bar{X}_2)^2 [\alpha^2 \sum (X_{i2} - \bar{X}_2)^2 + \sum u_i^2]}}$$

olacaktır.

$$\sum u_i^2 = 0$$

ise $r_{23} = 1$ olacak, fakat $\sum u_i^2$ büyüdükçe r_{23} 1 den 0 a doğru küçülecektir.

Bu durumda değerleri yerine koyduğumuzda

$$(X'X) = \begin{bmatrix} \sum (X_{i2} - \bar{X}_2)^2 & \alpha \sum (X_{i2} - \bar{X}_2)^2 \\ \alpha \sum (X_{i2} - \bar{X}_2)^2 & \alpha^2 \sum (X_{i2} - \bar{X}_2)^2 + \sum u_i^2 \end{bmatrix}$$

matrisin determinanı

$$D = \sum (X_{i2} - \bar{X}_2)^2 [\alpha^2 \sum (X_{i2} - \bar{X}_2)^2 + \sum u_i^2] - [\alpha \sum (X_{i2} - \bar{X}_2)^2 \alpha \sum (X_{i2} - \bar{X}_2)^2]$$

olacağından

$$\sum u_i^2 = 0$$

ise $D=0$ olacaktır. $\sum u_i^2$ küçük ise r_{23} büyüyecek aynı zamanda $X'X$ matrisinin determinanı küçük olacağından parametre tahmincilerinin varyansları gerçek değerlerinden büyük tahmin edilecektir.

Çoklu doğrusal bağlantının sonuçları şu şekilde özetlenebilir. Tam çoklu doğrusal bağıllık durumunda parametreler tahmin edilemez. Kuvvetli çoklu doğrusal bağlantı durumunda ise tahmincilerin varyansları büyük çıkacaktır. Bu durum yapılacak t testlerinin olumsuz sonuç vermesine belirlilik katsayısının büyümesine F testinin olumlu sonuç vermesine yol açacaktır.

Çoklu doğrusal bağlantının belirlenmesinde varyans büyütme faktörü metodu uygulamada sıklıkla kullanılan yöntemlerdendir. Bu faktör ile parametre tahminlerinin ve varyanslarının çoklu doğrusal bağlantı nedeni ile gerçek değerlerinden ne derece uzaklaştığı belirlenir.¹⁵

Varyans büyütme faktörü(VIF) şu şekilde formülleştirilebilir.

$$VIF = \frac{1}{1 - R_i^2}$$

¹⁵ Güriş, Çağlayan, **a.g.e.**, s.567.

olarak hesaplanacaktır. R_i^2 incelenen modelde yeralan bağımsız değişkenlerin sırası ile bağımlı değişken olup diğer bağımsız değişkenlerin bağımsız değişkenlerin bağımsız değişken olduğu modelden elde edilen belirlilik katsayısıdır. Elde edilen VIF değeri 5 ten küçükse çoklu doğrusal bağlantının önemli olmadığına; 5 ten büyük ise çoklu doğrusal bağlantının önemli olduğuna karar verilecektir.

Çoklu doğrusal bağlantının tespiti modelin fonksiyonel formu ile ilgisiz olduğundan lojistik regresyon modellerinde çoklu doğrusal bağlantının tespiti doğrusal regresyon modelleri ile aynı şekilde yapılacaktır.¹⁶

1.7.2 Hata Terimlerinin Analizi

Regresyon denklemlerinde bağımlı değişkenin gerçek ve tahmini değerleri arasındaki fark istatistiksel hatayı ifade eder ve kalıntı, artık veya hata terimi olarak adlandırılır. Doğrusal regresyon modellerindeki varsayımların çoğu hata terimleri ile ilgilidir. Bunlar 0 ortalama homoskedasite, normal dağılım, otokorelasyon olmaması şeklindeki varsayımlardır. Sıfır ortalama ve normal dağılım varsayımlarının bozulması nispeten küçük sorunlara yol açacak iken diğer varsayımların bozulması daha büyük sorunlara sebep olacaktır. Bunlar standart hataları büyümesine ve bu nedenle istatistiksel çıkarımların doğruluğunun bozulmasına yol açabileceği gibi parametre tahmincilerinin eğilimli ve etkinsiz olmasına da neden olabilecektir.

Doğrusal regresyon modellerinde hata terimleri regresyon denkleminde hesaplanır. Lojistik regresyonda ise olasılıklar ve logit gibi çeşitli durumlara uygun farklı artıklar mümkündür. Doğrusal regresyonda hata terimlerinin Y'nin koşullu ortalamasından bağımsız olduğu varsayılır. Lojistik regresyonda ise hata

¹⁶ Scott Menard, **Applied logistic Regression Analysis**, 2. bs., California, Sage Publications, 2002, pp.76.

terimlerinin varyansı koşullu ortalamanın fonksiyonudur. Bu sebepten dolayı hata terimleri standart hatalar ile düzeltilerek standardize edilmelidir¹⁷. Bu bağlamda Pearson artıkları

$$r_j = \frac{P(Y_j = 1) - \hat{P}(Y_j = 1)}{\sqrt{\hat{P}(Y_j = 1)[1 - \hat{P}(Y_j = 1)]}}$$

şeklinde ifade edilecektir.

1.7.2.1 Hata Terimlerinin Normalliğinin Analizi

Standart regresyon modellerinde hata terimlerinin normal dağıldığı varsayılır. Küçük örnekler için bu varsayım bozulduğunda istatistiksel çıkarımlar doğru sonuç vermeyecektir. Büyük örneklerde ise merkezi limit teoremine göre regresyon katsayılarının dağılımı normal dağılıma yaklaşacağından istatistiksel çıkarımların kesin olmayışının önemli olmadığı düşünülür.¹⁸

Lojistik regresyonda ise hata terimleri binom dağılıma sahiptir. Sadece büyük örnekler için hata terimlerinin dağılımını normal dağılıma yaklaştıracaktır. Normalliği test etmek için standardize edilmiş kalıntıların şeklinin çizilmesi gerekmektedir

¹⁷ A.e., pp.82.

¹⁸ A.e., pp.83.

1.7.2.2 Değişen Varyans

Lojistik regresyonda değişen varyansın tespiti için Goldfeld-Quandt(LM) testi uygulanabilmektedir.¹⁹

LM testinde hipotezler

H_0 : Sabit varyans varsayımı geçerlidir

H_1 : Sabit varyans varsayımı geçerli değildir

Test istatistiğinin hesaplanmasında X^* değişkeni büyüklüğüne göre sıralanır ve ortadaki p gözlem seçilerek atılacaktır. Buradaki X^* değişkeni σ_i^2 ile ilişkili olduğu düşünülen değişkendir.

Testin uygulanmasında ortadaki p gözlem atılacağı için p sayısının belirlenmesi önem taşıyacaktır. Atılacak p gözlem sayısı belirlenmeden önce serbestlik derecesinin

$$(n - p) / 2 > k$$

koşulunun sağlanması gerekmektedir. Burada k parametre sayısıdır. Atılacak gözlem sayısı olan p

$$p = \frac{n}{5}$$

şeklinde belirlenir.²⁰

¹⁹ William H Grene, **Econometric Analysis**, 5. bs., New Jersey, Prentice Hall, 2003, pp. 679.

²⁰ Güriş, Çağlayan, **a.g.e**, s.532.

P belirlendikten sonra sırasıyla baştan ve sondan $(n-p)/2$ gözlem atılarak iki ayrı model tahmin edilir. Tahmin edilen bu modellerin artıklarının kareleri toplamı hesaplanır. Bu artıkların kareleri toplamı aracılığı ile test istatistiği şu şekilde hesaplanabilir.

$$F = \frac{SSE_2}{SSE_1}$$

Bu formüldeki

SSE_1 : İlk grup için hesaplanan artıkların kareleri toplamı

SSE_2 : Son grup için hesaplanan artıkların kareleri toplamı

şeklinde tanımlanabilir. Hesaplanan bu test istatistiğinin dağılımı

$$SD_1 = \frac{1}{2}(n-p-2k)$$

$$SD_2 = \frac{1}{2}(n-p-k)$$

serbestlik dereceli F dağılımıdır ve tablo değeri ilgili serbestlik dereceleri ile F dağılımı tablosundan bulunacaktır.

Hesaplanan test istatistiği tablo değerinden büyük ise H_1 , hesaplanan test istatistiği tablo değerinden küçük ise H_0 hipotezi kabul edilecektir.

Bu test hangi deęiřkenin deęiřen varyansa neden olduęunun belirlenemedięi durumlarda tm baęımsız deęiřkenler iin ayrı ayrı uygulanır ve bu sayede sabit varyans varsayımının geersizlięine hangi deęiřken veya deęiřkenlerin neden olduęu belirlenmiř olur.

2. BÖLÜM

ÇOK SEÇENEKLİ NİTEL BAĞIMLI DEĞİŞKENLİ MODELLER

Nitel bağımlı değişkenler birinci bölümde bahsedildiği üzere iki değer alabileceği gibi ikiden fazla değer de alabilmektedir. Örneğin evinden işine gidecek olan kişi otobüs, minibüs veya tren seçeneklerinden birini tercih etme şansına sahiptir. Veya kişilerin bir ürünü alıp almayacaklarına ilişkin soruya evet, hayır veya belki cevabını verebileceklerdir. Bu şekilde ikiden fazla değer alabilen nitel değişkenlere çok seçenekli nitel değişkenler adı verilir.

Çok seçenekli nitel değişkenler bağımsız değişken olarak regresyon modellerine dahil edilebileceği gibi bağımlı değişken olarak ta regresyon modellerine dahil edilebilmektedir.

Çok seçenekli nitel bağımlı değişkenler kategorik ve kategorik olmayan değişkenler biçiminde sınıflandırılabilir.¹

Kategorik değişkenler, değişkenin aldığı değerlerin belirli aralıklar arasında olup olmamasına göre sınıflandırılabilen değişkenlerdir ve bu durumda nitel değişken 1,2,3,4 gibi değerler alacaktır.

Kategorik olmayan değişkenlere ise ailenin sahip olduğu otomobil sayısı örnek olarak verilebilir ve 0 ,1,2,3 gibi değerler alır.

Kategorik değişkenler sıralı, sıralı olmayan, ardışık değişkenler olarak sınıflandırılabilir.

¹ G.S. Maddala, **Limited Dependent and Qualitative Variables in Econometrics**, Cambridge, Cambridge University Press, 1983, pp.14.

Sıralı değişken; belirli bir değerden başlayarak aralıklar halinde birbirini takip eden ve seçenekler arasında sıralamanın önem taşıdığı değişkenlerdir². Kişilerin harcama düzeyleri bu tür değişkenlere örnek olarak verilebilir.

$Y = 1$	500 milyon TL den az harcama
$Y = 2$	500 – 599 milyon TL arası harcamı
$Y = 3$	600 – 699 milyon TL arası harcamı
$Y = 4$	700 – 799 milyon TL arası harcamı
$Y = 5$	800 milyon TL den fazla harcama

şeklinde ifade edilebilen değişkenler sıralı kategorik değişkenlerdir.

Seçenekler arasında sıranın önem taşımadığı yani tercihin subjektif olduğu değişkenlere sıralı olmayan kategorik değişkenler denir.³ Sıralı olmayan kategorik değişkenlere taşıt tercihi örnek verilebilir.

$Y = 1$	Tren ile seyahat
$Y = 2$	Otobüs ile seyahat
$Y = 3$	Minibüs ile seyahat
$Y = 4$	Taksi ile seyahat
$Y = 5$	Vapur ile seyahat

Ardışık değişkenler seçeneklerin birbirini sırası ile takip ettiği değişkenlerdir. Ardışık değişkenlere eğitim durumları örnek olarak verilebilir.

² Michael R. Frone, “Regression Models for Discrete and Limited Dependent Variables,” **Research Methods Forum**, No.2, Summer 1997, pp.2.

³ A.e., pp.2.

- $Y = 1$ *İlköğretim mezunu*
 $Y = 2$ *Ortaogretim mezunu*
 $Y = 3$ *Üniversite mezunu*
 $Y = 4$ *Yükseklisans mezunu*
 $Y = 5$ *Doktora mezunu*

Bu değişkenler nitel bağımlı değişkenler olarak regresyon modellerinde kullanılabilir. Bu tür değişkenlerin nitel bağımlı değişken olarak kullanıldığı regresyon modellerine çoklu tercih modelleri denir.

2.1 ÇOK SEÇENEKLİ DOĞRUSAL OLASILIK MODELİ

Çok seçenekli doğrusal olasılık modeli, bağımlı değişkenin iki değer alabildiği doğrusal olasılık modelinin genişletilmiş şeklidir.

Doğrusal regresyon modeli;

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

şeklinde ifade edilebilir. Modeli matris gösterimi ile

$$Y_i = \sum \beta_k X_{ik} + u_i$$

şeklinde yazabiliriz. Modeldeki bağımlı değişkenlerin j tane değer aldığını varsaydığımızda,

$$Y_{ij} = \sum \beta_{kj} X_{ik} + u_{ij} \quad j=1,2,\dots,J$$

şeklinde ifade edilecektir. Bu durumda model seçenek sayısı kadar farklı denklem haline getirilerek çözümlenir. Bunu bir örnek üzerinde açıklayalım. Bir çalışanın evden çıkıp işyerine gelmesi için üç seçeneği olsun. Bunların sırasıyla tren, otobüs, minibüs olduğunu varsayalım. Bu durumda model, çok seçenekli doğrusal olasılık modeline göre, iki değer alabilen üç parçada değerlendirilecektir. Yani;

1. Parça Tren tercih etme	1
Diğer	0

2. Parça Otobüs tercih etme	1
Diğer	0

3. Parça Minibüs tercih etme	1
Diğer	0

şeklinde parçalara ayrılacaktır.

İki seçeneğe dönüştürülen modeller önceki bölümde belirtildiği gibi En Küçük Kareler yöntemi ile çözülebilir. Bu modellerin doğrusal olasılık modelinden farkı seçenek sayısı kadar denklemin çözümlenmesi ile çözüme ulaştırılabilmesidir.

İki seçenekli doğrusal olasılık modelinde karşılaştığımız en büyük sorun tahmin ettiğimiz olasılık değerinin 0 ile 1 arasında olmamasıydı. Aynı sorun burada da devam etmektedir. Ayrıca burada denklemlerden ayrı ayrı elde ettiğimiz olasılık değerlerinin toplamının 0 ile 1 arasında olmaması ilave sorunuyla da karşılaşırız. Bu ilave sorun her bir denklemin ayrı ayrı sabit terim içermesi durumunda regresyonun normal denklemlerle çözümlenmesi olarak ifade edilen cebirsel yöntem ile çözülebilir. Bu yöntem bize bir ilave fayda sağlar. Bu fayda J tane denklemi tahmin etmek yerine J-1 tane denklemin çözülmesi diğer denklemin cebirsel olarak çözülmesidir.⁴

⁴ John. H. Aldrich, Forrest. D Nelson, **Linear Probability, Logit and Probit Models**, California, Sage Publications, 1984, pp.23.

Tahmin edilen olasılıkların toplamı 1'e eşit olması gerektiğinden tahmin edilen sabit terimlerin toplamı 1'e, eğim katsayılarının toplamı da 0'a eşit olacaktır.⁵

Bunu, aşağıdaki şekilde gösterebiliriz.

P_{1i} = i inci bireyin 1. seçeneği tercih etmesi

P_{2i} = i inci bireyin 2. seçeneği tercih etmesi

P_{3i} = i inci bireyin 3. seçeneği tercih etmesi

olarak ifade edildiğinde

$$P_{1i} + P_{2i} + P_{3i} = 1$$

$$\bar{P}_1 + \bar{P}_2 + \bar{P}_3 = 1$$

$$p_{1i} + p_{2i} + p_{3i} = 0$$

olacaktır. Yani olasılıklar toplamı ve olasılıkların ortalamalarının toplamı 1 iken olasılıklarının ortalama farkları toplamı da sıfırdır. Bu durumda

$$\begin{aligned} \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 &= \frac{\sum P_{1i} X_i}{\sum X_i^2} + \frac{\sum P_{2i} X_i}{\sum X_i^2} + \frac{\sum P_{3i} X_i}{\sum X_i^2} \\ &= \frac{\sum (P_{1i} + P_{2i} + P_{3i}) X_i}{\sum X_i^2} \end{aligned}$$

⁵ Hüseyin Özer, **Nitel Değişkenli Ekonometrik Modeller Teori ve Uygulama**, Ankara, Nobel Yayın Dağıtım, 2004, s.84.

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

olacaktır. Sonuçta,

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

olarak elde edilecektir. Buradan da anlaşılacağı gibi 3 seçeneğe ait her bir denklemin çözülmesi yerine 2 tane denklem çözülüp 3. denklem buradan hesaplanabilecektir.

J-1 sayıdaki denklemin EKK ile çözümlenmesi sapmasızdır. Hata terimleri önceki bölümde de belirtildiği gibi değişen varyansa sahiptir. Bu yüzden Tartılı EKK yöntemi kullanılmalıdır.

2.2 SIRALI OLMAYAN ÇOK SEÇENEKLİ TERCİH MODELLERİ

Seçenekler arasındaki tercihin subjektif olduğu modellerdir. Bu bölümde uygulamada sıklıkla kullanılan çok seçenekli logit ve probit modellerden bahsedilecektir.

2.2.1 Çok seçenekli Logit Modeli

İkiden fazla seçeneğe sahip seçenekler arasında sıranın önem taşımadığı ve lojistik dağılım temellerine dayanan modele çok seçenekli logit modeli denir.⁶ Bu modele multinomial logit model adı da verilmektedir. Çok seçenekli doğrusal olasılık modelinde karşılaştığımız en büyük sorun olan tahmin edilen olasılık değerinin 0 ile 1 arasında değerler almaması sorununun çözülebilmesi için kullanılan modellerdendir.

Bağımlı değişkenin m tane değer aldığını varsayalım.

P_{ij} = i inci bireyin j inci seçeneği seçme olasılığıdır.

O halde;

$$P_{i1} + P_{i2} + P_{i3} + \dots + P_{im} = 1 \text{ olacaktır.}$$

Önceki bölümde lojistik dağılım fonksiyonu

$$P_i = \frac{1}{1 + e^{-(\beta_1 + \beta_2 X_i)}}$$

veya

$$P_i = \frac{e^{(\beta_1 + \beta_2 X_i)}}{1 + e^{(\beta_1 + \beta_2 X_i)}}$$

şeklinde ifade edilmişti. Bu ifadeyi iki seçenekli logit modeli için kullanmıştık. Çok seçenekli logit modeli ise iki seçenekli logit modelin genelleştirilmiş halidir ve

⁶ J.S. Cramer, **An Introduction The Logit Model for Economists**, New York, Edward Arnold, 1991, pp.43.

$$P_j = \frac{e^{\beta_j' X}}{1 + \sum_{k=1}^{m-1} e^{\beta_k' X}} \quad j = 1, 2, 3, \dots, m-1$$

şeklinde ifade edilebilir. Bu ifade çok seçenekli logit olarak adlandırılır.

Çok seçenekli logit modelin parametreleri en çok benzerlik yöntemi ile tahmin edilebilir.

$y_{ij} = 1$ i inci bireyin j inci kategoriye seçmesi
 $y_{ij} = 0$ diğer durumlar

olarak yazıldığında çok seçenekli logit modelin benzerlik fonksiyonu;

$$L = \prod_{i=1}^n P_{i1}^{y_{i1}} P_{i2}^{y_{i2}} \dots P_{im}^{y_{im}}$$

veya

$$\log L = \sum_{i=1}^n \sum_{j=1}^m y_{ij} \log P_{ij}$$

olarak yazılabilir⁷. Burada,

$$P_{ij} = \frac{e^{(x_i' \beta_j)}}{1 + \sum_{k=1}^{m-1} e^{(x_i' \beta_k)}} \quad (j = 1, 2, \dots, m-1)$$

ve

⁷ Maddala, **a.g.e.**, s.36.

$$P_{im} = \frac{1}{1 + \sum_{k=1}^{m-1} e^{(x'_i \beta_k)}}$$

olacaktır. Buradan

$$\frac{\partial P_{ij}}{\partial \beta_j} = P_{ij} (1 - P_{ij}) x_i$$

$$(j, k = 1, 2, \dots, m-1)$$

$$\frac{\partial P_{ij}}{\partial \beta_k} = -P_{ij} P_{ik} x_i$$

ve

$$\frac{\partial P_{im}}{\partial \beta_j} = -P_{ij} P_{im} x_i$$

şeklinde yazılabilir. Böylece;

$$\begin{aligned} \frac{\partial \log L}{\partial \beta_k} &= \sum_{i=1}^n \left[\frac{y_{ik}}{P_{ik}} P_{ik} (1 - P_{ik}) + \sum_{\substack{j=1 \\ j \neq k}}^m \frac{y_{ij}}{P_{ij}} (-P_{ij} P_{ik}) \right] x_i \\ &= \sum_{i=1}^n \left[y_{ik} - P_{ik} \left(\sum_{j=1}^m y_{ij} \right) \right] x_i \\ &= \sum_{i=1}^n (y_{ik} - P_{ik}) x_i \end{aligned}$$

olacaktır. Çünkü

$$\sum_{j=1}^m y_{ij} = 1 \text{ ' dir.}$$

Böylece, elde edilen En Çok Benzerlik tahmincileri için çözümde kullanılan denklem aşağıdaki gibidir.

$$\sum_{i=1}^n (y_{ik} - P_{ik}) x_i = 0 \quad (k = 1, 2, \dots, m-1)$$

Buradan logaritmik fark oranlarını şu şekilde gösterebiliriz.

$$\ln\left(\frac{P_{ij}}{P_{io}}\right) = \beta'_j X_i$$

$$\ln\left(\frac{P_{ij}}{P_{ik}}\right) = X'_i(\beta_j - \beta_k)$$

Çok durumlu logit modelde iki alternatifin olabilirlik oranları belirlenirken diğer alternatifler dikkate alınmamaktadır. Bu durum Mc Fadden tarafından ilişkisiz alternatiflerin bağımsızlığı olarak adlandırılmıştır.⁸ Alternatifler arası bağımsızlık iki seçeneğin nispi olasılıklarının üçüncü seçeneği etkilememesi olarak tanımlanabilir.⁹ Bu çok seçenekli logit modelin dezavantajıdır.¹⁰

Çok seçenekli logit model yapısının verilere uygun olup olmadığını belirlemek için alternatifler arasındaki bağımsızlığı Hausman testi ile test edebiliriz¹¹. Eğer katsayılar istatistiksel olarak 1 den farklı iseler bu İlişkisiz Alternatiflerin Bağımsızlığı(IIA) özelliğinin yerine getirilmediğinin kanıtıdır.¹²

⁸ William H Greene, **Econometric Analysis**, 5. bs., New Jersey, Prentice Hall, 2003, pp.724.

⁹ Michael J. Hilmer, "A Comparison of Alternative Specifications of the Collage Attendance Equation with an Extension to two-stage Selectivity-Correction Models," **Economics of Education Review**, Vol:20, 2001, pp. 264.

¹⁰ Peter Kennedy, **A Guide to Econometrics**, 4. bs., Oxford, Blackwell Publishers, 1998, pp.235.

¹¹ Hilmer, **a.g.e.**, pp. 265.

¹² Maddala, **a.g.e.**, pp.77.

Seenekler arasından bir seenek diğerklerinden ilişkisiz ise onu modelden çıkarmak parametre tahminlerini değıştirmeyecektir. Fakat fark oranları bu alternatiflerden bağımsız ise elde edilen parametre tahminleri tutarsız olacaktır.¹³

Hausman testi bağımlı değışkendeki seeneklerden birini çıkartarak sınırlandırılmamış model ile sınırlandırılmış model tahminlerinin karşılaştırılması şeklinde tanımlanabilir. Test istatistiğı

$$\chi^2 = (\beta_s - \beta_f)' [V_s - V_f]^{-1} (\beta_s - \beta_f)$$

şeklinde ifade edilebilir.

β_s : “s” sayıda tercihlerin sınırlandırılmış modelinden elde edilen parametre tahminlerinin vektörü,

β_f : “f” sayıda tüm tercihlerin olduğu modelden elde edilen parametre tahminlerinin vektörü,

V_s : Sınırlandırılmış modelden elde edilen kovaryans matrisi tahmini,

V_f : Sınırlandırılmamış modelden elde edilen kovaryans matrisi tahminidir.

Bu test istatistiğı $V_s - V_f$ matrisinin rankı büyüklüğünde serbestlik derecesine sahip olan ki kare dağılımına uymaktadır.¹⁴

¹³ Şenay Üçdoğruk, **Multinomial Logit Model: Bölgeler Arası İşgücü Tercihleri Analizi**, İzmir, Dokuz Eylül Üniversitesi Yayınları, 2002, s.19.

¹⁴ Maddala, **a.g.e.**, pp.78.

Adonis Yatchew, Zvi Griliches, “Specification Error in Probit Models,” **The Review of Economic and Statistics**, Vol. 67, No.1, Şubat 1985 pp.138.

Çok seçenekli logit modelin avantajı probit model gibi çoklu integraller içermemesi nedeniyle hesaplanmasının daha kolay olmasıdır¹⁵.

2.2.2 Çok seçenekli Probit Model

Çok seçenekli, seçenekler arasında sıranın önem taşımadığı nitel bağımlı değişkenli modellere bir örnek de hata terimlerinin çok değişkenli normal dağılıma sahip olduğu probit modellerdir.

Çok seçenekli probit modellerin kullanım alanı çok seçenekli logit modellere göre daha kısıtlıdır. Çok seçenekli probit modeller seçenek sayısının az olduğu durumlarda kullanılabilir.¹⁶ Bunun sebebi çok seçenekli probit modelin iç içe geçmiş integrallerden oluşması ve bunu çözümlemede yaşanan zorluklardır.

Çok seçenekli probit modeli şu şekilde açıklanabilir. Üç seçeneğin olduğu durumda bu seçenekleri,

$$Y_1^* = V_1 + \varepsilon_1$$

$$Y_2^* = V_2 + \varepsilon_2$$

$$Y_3^* = V_3 + \varepsilon_3$$

şeklinde ifade edebiliriz.

$\varepsilon_1, \varepsilon_2, \varepsilon_3$ ' ün normal dağılıma sahip oldukları varsayıldığında kovaryans matrisi

¹⁵ Piyush Tiwari, Tetsu Kawakami, “ Modes of Commuting in Mumbai: A Discrete Choice Analysis,” **The Applied Regional Science Conference**, Vol.13, No.1, March 2000, pp.38.

¹⁶ Allan P. Layton, Masaki Katsuura, “Comparison of Regime Switching, Probit and Logit Models in Dating and Forecasting US Business Cycles,” **International Journal of Forecasting**, Vol.17, pp.407.

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

şeklinde yazılacaktır. İlk seçeneğin seçilme olasılığını

$$\Pr ob(Y_1^* > Y_2^*, Y_1^* > Y_3^*) = \Pr ob(\varepsilon_2 - \varepsilon_1 < V_1 - V_2, \varepsilon_3 - \varepsilon_1 < V_1 - V_3)$$

olarak ifade ettiğimizde;

$$n_{21} = \varepsilon_2 - \varepsilon_1, \quad n_{31} = \varepsilon_3 - \varepsilon_1; \quad V_{12} = V_1 - V_2 \quad \text{ve} \quad V_{13} = V_1 - V_3$$

n_{21} ve n_{23} normal dağılıma sahiptirler ve kovaryans matrisi

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

şeklinde olacaktır. Bu bilgiler ışığında birinci seçeneğin seçilme olasılığı,

$$P_1 = \int_{-\infty}^{V_{12}} \int_{-\infty}^{V_{13}} f(n_{21}, n_{31}) dn_{21} dn_{31}$$

olacaktır. Burada $f(n_{21}, n_{31})$ normal dağılıma sahiptir. İkinci ve üçüncü seçeneğin seçilme olasılığı da buna uygun şekilde yazılabilir.¹⁷

İkinci ve üçüncü seçeneklerin seçilme olasılığı da

¹⁷ Maddala, **a.g.e.**, pp.63.

Denic Bolduc, "A Practical Technique to Estimate Multinomial Probit Models in Transportation," **Transportation Research Part B**, Vol. 33, 1999, pp.68.

$$P_2 = \int_{-\infty}^{V_{12}} \int_{-\infty}^{V_{13}} f(n_{12}, n_{32}) dn_{12} dn_{32}$$

$$P_3 = \int_{-\infty}^{V_{12}} \int_{-\infty}^{V_{13}} f(n_{13}, n_{23}) dn_{12} dn_{23}$$

şeklinde ifade edilebilir.

Çok seçenekli probit modelin tahmini En çok benzerlik yöntemi ile yapılabilir. Bu yöntem alternatif olarak iki yöntem daha önerilmiştir. Bunlardan ilki Lerman ve Manski(1982) tarafından kullanılan Monte Carlo metodudur. V_i' ye değerler verilerek kovaryans matrisleri oluşturulur ve i' inci seçenekte faydanın maksimize edilmesi prosedürüne dayanır.¹⁸

Çok seçenekli probit modelin çözümü için önerilen ikinci yaklaşım Clark(1961) tarafından önerilen tahmin prosedürüne dayanır¹⁹. Bu tahmin prosedürü eşit varyansların negatif olmayan korelasyonlu değişimleri için uygundur. Fakat negatif korelasyonlu veya eşit olmayan varyanslılar için uygun değildir. Clark metodu X_1, X_2, X_3 gibi üç'lü normal dağılan üçlü tesadüfi değişkenin dağılımını ve $[(X_1, \max(X_2, X_3))]$ nin birinci ve ikinci momentleri aynı olan ikili normal dağılımlarına yaklaştırma temeline dayanır. m seçenek için çok seçenekli probit modeli $m-2$ kere Clark tahminlerinin tekrarlanmasıyla elde edilir.²⁰

¹⁸ Maddala, **a.g.e.**, pp. 63.

¹⁹ Pradeep K. Chintagunta, "Estimating A Multinomial Probit Model of Brand Choice Using the Method of Simulated Moments," **Marketing Science**, Vol.11 No.4 Fall 1992, pp. 393.

²⁰ Maddala, **a.g.e.**, pp.64.

2.3 SIRANIN ÖNEM TAŞIDIĞI MODELLER

Çok seçenekli ve seçenekler arasında sıranın önem taşıdığı modellere sıralı çok seçenekli tercih modelleri denir. Uygulamada sıklıkla kullanılan sıralı tercih modelleri sıralı çok seçenekli logit ve sıralı çok seçenekli probit modelleridir.

Bu modelleri açıklayabilmek için bir bölgede yaşayan N tane kişi olduğunu ve her kişinin gelir düzeyinin D_i değişkeni ile ifade edildiğini varsayalım. Gelir düzeyi D_i ' yi etkileyen K tane faktör olduğunu ve D_i ' nin K tane faktörün doğrusal fonksiyonu olduğunu varsayalım. i ' inci birey için K faktörün değeri X_{ik} $k=1, \dots, K$ ' dır. Bunu şu şekilde ifade edebiliriz.

$$D_i = \sum_{k=1}^K \beta_k X_{ik} + \varepsilon_i = Z_i + \varepsilon_i$$

Burada β_k k' inci değişkenin katsayısı ve $Z_i = \sum_{k=1}^K \beta_k X_{ik}$ dır. Burada $\beta_k > 0$ ise k inci faktörün değerindeki artış D_i değerinde bir artışa; $\beta_k < 0$ ise k inci faktörün değerindeki bir azalış D_i değerinde azalışa neden olur. Buradaki hata payı da k faktör tarafından açıklanamayan değişmeyi ifade edecektir.

Örneğin kişilerin gelir düzeyleri belirli eşik değerlere göre sınıflandırılabilir. $Y_i = 1$ düşük gelir düzeyi, $Y_i = 2$ orta gelir düzeyi, $Y_i = 3$

yüksek gelir düzeyi şeklinde sınıflandırılabilir. Bu sınıflandırma , δ_1 ve δ_2 olarak belirlenen eşik değerlerine bağlıdır²¹.

$$\begin{aligned} Y_i = 1 & \quad \text{eğer} \quad D_i \leq \delta_1 \\ Y_i = 2 & \quad \text{eğer} \quad \delta_1 \leq D_i \leq \delta_2 \\ Y_i = 3 & \quad \text{eğer} \quad D_i \geq \delta_2 \end{aligned}$$

Kişilerin hangi sınıfta olduğu kişinin gelir düzeylerinin eşik değerlerine göre olan durumuna bağlıdır.

Y_i 'ye 1, 2 ve 3 değerlerinin verildiği durumlardaki olasılıklar;

$$\Pr(Y_i = 1) = \Pr(Z_i + \varepsilon_i \leq \delta_1) = \Pr(\varepsilon_i \leq \delta_1 - Z_i)$$

$$\Pr(Y_i = 2) = \Pr(\delta_1 \leq Z_i + \varepsilon_i \leq \delta_2) = \Pr(\delta_1 - Z_i < \varepsilon_i \leq \delta_2 - Z_i)$$

$$\Pr(Y_i = 3) = \Pr(Z_i + \varepsilon_i \geq \delta_2) = \Pr(\varepsilon_i \geq \delta_2 - Z_i)$$

şeklinde ifade edilebilir.

Her N gözlem bu 3 seçenekten birine sahiptir. Bu N tane kişinin, N_1 düşük gelir düzeyi , N_2 orta gelir düzeyi, N_3 yüksek gelir düzeyi olduğu varsayıldığında;

$$\begin{aligned} L &= [\Pr(Y_i = 1)]^{N_1} [\Pr(Y_i = 2)]^{N_2} [\Pr(Y_i = 3)]^{N_3} \\ &= [F(\delta_1 - Z_i)]^{N_1} [F(\delta_2 - Z_i) - F(\delta_1 - Z_i)]^{N_2} [1 - F(\delta_1 - Z_i)]^{N_3} \end{aligned}$$

olacaktır. Burada $F(x) = \Pr(\varepsilon_i < x)$ hata terimlerinin toplam olasılık dağılımıdır.

Eğer hata terimlerinin olasılık dağılımları biliniyorsa β_k ' lerin tahminleri seçilebilir.

Bilginin yokluğunda hata terimlerinin özel olasılık dağılımlarına sahip olduğu

²¹ Mohamed A. Abdel-Aty, "Using Ordered Probit Modeling to Study the Effect of ATIS on Transit Ridership," **Transportation Research Part C**, Vol. 9 , 2001, pp. 271.

varsayılır. Sıralı Logit ve Sıralı Probit arasındaki fark ε_i 'nin dağılımlarından kaynaklanır. Sıralı Logit modelde hata terimlerinin lojistik dağıldığı varsayımına dayanır. Sıralı Probit modelde ise hata terimlerinin normal dağıldığı varsayılır. Dağılımlar arasında fark çok fazla olmadığından tercih yapmak zordur.²²

β_k 'nin katsayılarının tahmini kullanılarak örnekteki her bir birim için $\hat{Z}_i = \sum_{k=1}^K \hat{\beta}_k X_{ik}$ hesaplanır.

$\hat{Z}_i$, δ_1 ve δ_2 kullanılarak örnekteki her bir birim için olasılıklar şu şekilde hesaplanır.

$$\begin{aligned}\hat{P}_{i1} &= \Pr(\varepsilon_i \leq \delta_1 - \hat{Z}_i) = F(\delta_1 - \hat{Z}_i) \\ \hat{P}_{i2} &= \Pr(\delta_1 - \hat{Z}_i < \varepsilon_i \leq \delta_2 - \hat{Z}_i) = F(\delta_2 - \hat{Z}_i) - F(\delta_1 - \hat{Z}_i) \\ \hat{P}_{i3} &= \Pr(\varepsilon_i \geq \delta_2 - \hat{Z}_i) = 1 - F(\delta_2 - \hat{Z}_i)\end{aligned}$$

Burada $\sum_{j=1}^3 \hat{P}_{ij} = 1$ olacaktır.

Yukarıda açıklanan model oransal ihtimaller modeli olarak bilinir. Oransal ihtimaller modeli;

$$OR(m) = \frac{\Pr(Y_i \leq m)}{\Pr(Y_i > m)}$$

şeklinde gösterilebilir

²² Vani K. Borooah, **Logit and Probit Ordered and Multinomial Models**; California, Sage Publications, 2002, pp.9.

2.3.1 Sıralı Çok Seçenekli Logit Modeli

Çok seçenekli seçenekler arası sıralamanın önem taşıdığı modeller biri sıralı çok seçenekli logit modelidir. Tesadüfi değişken X 'in Lojistik dağılım fonksiyonu

$$\begin{aligned}\Pr(X \leq x) = \Lambda(x) &= \frac{e^x}{[1 + e^x]} \\ &= \frac{1}{[1 + e^{-x}]}\end{aligned}$$

ve eğer hata terimleri lojistik dağılıma sahip ise

$$\Pr(Y_i = 1) = \Lambda(\delta_1 - Z_i) = \frac{1}{[1 + e^{(Z_i - \delta_1)}]}$$

$$\begin{aligned}\Pr(Y_i = 2) &= \Lambda(\delta_2 - Z_i) - \Lambda(\delta_1 - Z_i) \\ &= \frac{1}{[1 + e^{(Z_i - \delta_2)}]} - \frac{1}{[1 + e^{(Z_i - \delta_1)}]}\end{aligned}$$

$$\Pr(Y_i = 3) = 1 - \Lambda(\delta_2 - Z_i) = 1 - \frac{1}{[1 + e^{(Z_i - \delta_2)}]}$$

β_k ' lerin tahmini $F(.)$ nin yerine lojistik dağılım fonksiyonu $\Lambda(.)$ kullanılarak

$$\begin{aligned}L &= [\Pr(Y_i = 1)]^{N_1} [\Pr(Y_i = 2)]^{N_2} [\Pr(Y_i = 3)]^{N_3} \\ &= [F(\delta_1 - Z_i)]^{N_1} [F(\delta_2 - Z_i) - F(\delta_1 - Z_i)]^{N_2} [1 - F(\delta_1 - Z_i)]^{N_3}\end{aligned}$$

denkleminin maksimum benzerlik fonksiyonundan elde edilebilir.²³

²³ A.e., pp.12.

2.3.2 Sıralı Çok Seçenekli Probit Modeli

Çok seçenekli seçenekler arası sıralamanın önem taşıdığı modeller bir diğeri sıralı çok seçenekli probit modelidir. Standart normal kümülatif dağılım X

$$\Pr(X < x) = \Phi(x) = \int_0^x (1/2\pi) e^{(-x^2/2)} dx$$

ve hata terimlerinin standart normal kümülatif dağılımlı olduğu varsayılırsa

$$\begin{aligned}\Pr(Y_i = 1) &= \Phi(\delta_1 - Z_i) \\ \Pr(Y_i = 2) &= \Phi(\delta_2 - Z_i) - \Phi(\delta_1 - Z_i) \\ \Pr(Y_i = 3) &= 1 - \Phi(\delta_2 - Z_i)\end{aligned}$$

β_k 'nın tahmini F(.)nin yerine Normal dağılım fonksiyonu $\Phi(\cdot)$ kullanılarak

$$\begin{aligned}L &= [\Pr(Y_i = 1)]^{N_1} [\Pr(Y_i = 2)]^{N_2} [\Pr(Y_i = 3)]^{N_3} \\ &= [F(\delta_1 - Z_i)]^{N_1} [F(\delta_2 - Z_i) - F(\delta_1 - Z_i)]^{N_2} [1 - F(\delta_2 - Z_i)]^{N_3}\end{aligned}$$

denkleminin maksimum benzerlik fonksiyonunun çözümünden elde edilir.²⁴

Sıralı probit model ile çok seçenekli sıranın önem taşımadığı logit modeli arasındaki temel fark, sıralı probit modelin sıralı yapısı gereği ilişkisiz alternatiflerin bağımsızlığı(IIA) özelliğinin sağlanmasına gerek olmamasıdır.²⁵

²⁴ A.e., pp.12.

²⁵ Hilmer, a.g.e., pp. 265.

Kara Maria Kockelman, Young-Jun Kweon, Driver Injury severity: An Application of Ordered Probit models, **Accident Analysis & Prevention**, vol.34, 2002, pp.316

2.4 ARDIŞIK ÇOK SEÇENEKLİ TERCİH MODELLERİ

Seçeneklerin birbirini ardışık olarak takip ettiği çok seçenekli tercih modellerine ardışık çok seçenekli tercih modelleri denir.

Ardışık çok seçenekli nitel bağımlı değişkeni

$Y = 1$ *İlköğretim mezunu*

$Y = 2$ *Ortaöğretim mezunu*

$Y = 3$ *Üniversite mezunu*

$Y = 4$ *Yüksek lisans veya doktora mezunu*

biçiminde tanımlarsak, olasılıklar

$$P_1 = F(\beta'_1 X)$$

$$P_2 = [1 - F(\beta'_1 X)] F(\beta'_2 X)$$

$$P_3 = [1 - F(\beta'_1 X)] [1 - F(\beta'_2 X)] F(\beta'_3 X)$$

$$P_4 = [1 - F(\beta'_1 X)] [1 - F(\beta'_2 X)] [1 - F(\beta'_3 X)]$$

şeklinde ifade edilecektir. Burada β_1 katsayısı tüm örnekteki bireylerin ilköğretim mezunu olup olmamalarına göre iki gruba ayrılarak tahmin edilebilir. β_2 katsayısı ise ilköğretim mezunu olan bireylerin oluşturduğu alt grupta bireyler bu defa ortaöğretim mezunu olup olmamalarına göre ikiye ayrılarak tahmin edilebilir. β_3 katsayısı orta öğretim mezunu olan bireylerin oluşturduğu alt grupta bireyler yükseköğretim mezunu olup olmamalarına göre ikiye ayrılarak tahmin edilebilir. Her

bir durumda oluşturulan ikili modeller logit veya probit yöntemi ile tahmin edilebilir.²⁶

Çok seçenekli ardışık modelleri analiz etmek bu modeller için ikili benzerlik fonksiyonlarının maksimizasyonunu gerektirdiğinden kolaydır.²⁷

2.5 ÇOK SEÇENEKLİ MODELLERİN YORUMLANMASI

Çok seçenekli modellerin yorumlanmasında Fark oranları, nisbi risk oranları ve marjinal etkilerden yararlanılır.

2.5.1 Fark Oranları

Çok seçenekli modelleri yorumlarken kullandığımız önemli göstergelerden birisi Fark oranıdır.

$Y=0$ seçeneğini temel seçenek olarak varsayalım İndisler hangi seçeneğin temel seçenek ile karşılaştırıldığını göstermektedir. X in a ve b değerlerini alması durumunda fark oranı

$$OR_j(a,b) = \frac{P(Y = j / x = a) / P(Y = 0 / x = a)}{P(Y = j / x = b) / P(Y = 0 / x = b)}$$

şeklinde ifade edilecektir. Çok seçenekli modellerde fark oranı modelin tahmin edilen parametresinin e sayısı üssü olarak alınması şeklinde hesaplanabilir.²⁸

²⁶ Maddala, **a.g.e.**, pp.38.

²⁷ A.e., pp.49.

²⁸ David W.Hosmer, Stanley Lemashof , **Applied Logistic Regression**; New York, John Wiley and Sons Publication, pp.266.

Fark oranı seçenekler arasındaki nispi farklılığı ve bu farklılığın yönü hakkında bilgi vermektedir.

2.5.2 Nispi Risk Oranı

Nispi risk oranı, fark oranlarının hesaplanması sırasında kullanılan mantığa dayanmaktadır. Fark oranlarından farkı çok seçenekli modelin hesaplanması sırasında kullanılan temel seçeneğe göre nispi oran vermesi ve dolayısı ile yorumun bu temel seçenek baz alınarak yapılmasıdır. Nispi risk oranı

$$RRR_{x_1} = \frac{e^{(\beta_0 + (x_{1ki}+1)\beta_1 + x_{2ki}\beta_2)}}{e^{(\beta_0 + x_{1ki}\beta_1 + x_{2ki}\beta_2)}} \\ = e^{\beta_1}$$

şeklinde ifade edilebilir. Nispi risk oranı e sayısı üzeri katsayı şeklinde hesaplanarak yorumlanır.²⁹

2.5.3 Marjinal Etkiler

Çok seçenekli logit model doğrusal olmayan bir model olduğundan her değişkenin farklı seçenekler üzerine etkisi sabit olmayacaktır.

J=1 kategorisini temel kategori olarak varsaydığımızda marjinal etkiler

²⁹ James Hardin, Joseph Hilbe, **Generalized Linear Models and Extensions**, Texas, Stata Pres, 2001, pp.178.

$$\frac{\partial P_{ij}}{\partial X_{ik}} = P_{ij} \left(\beta_j - \sum_{j=2}^j P_{ij} \beta_j \right)$$

şeklinde hesaplanabilir. Bağımsız değişkendeki bir birimlik değişmenin ilgili seçeneğin seçilme olasılığını yüzde kaç değiştirdiği hakkında bilgi verecektir.

3. BÖLÜM

ÜLKELERİN KALKINMIŞLIK DÜZEYLERİNİ ETKİLEYEN FAKTÖRLERİN ÇOK SEÇENEKLİ LOGİT MODEL İLE ANALİZİ

Bu bölümde Birleşmiş Milletlerin yapmış olduğu ülkelerin kalkınmışlık sınıflandırmaları kullanılarak, bir ülkenin gelişmiş, gelişmekte olan ve az gelişmiş ülkeler kategorisinde olma durumunu etkileyen faktörler araştırılacaktır. Analizimizde çok seçenekli logit modeli kullanılacaktır.

3.1 KULLANILAN VERİ

Birleşmiş Milletler yapmış olduğu araştırmalarda ülkeleri gelişmiş, gelişmekte olan az gelişmiş ülkeler olarak sınıflandırmaktadır. Bu sınıflandırmayı etkileyen çok çeşitli faktörler vardır. Bu faktörlerin belirlenmesinde Birleşmiş Milletlerin konuyla ilgili en yeni verileri olan 2002 yılının verileri kullanılarak, tesadüfi olarak seçilen 169 ülke ile analiz yapılmıştır. Ülkelerin gelişmişliklerini etkileyen faktörler olarak

Var1: Ülkenin gelişmiş ülke olması:1

Ülkenin gelişmekte olan ülke olması:2

Ülkenin az gelişmiş ülke olması:3

Var2: Yaşam beklentisi

Var3: Yetişkinlerin okuma oranı

Var4: Eğitim çağındaki nüfusun öğrenim görme oranı

Var5: Satınalma gücü paritesi ile düzeltilmiş kişi başına gayri safi milli hasıla

Var6: İnternet kullanım oranı

Var7: Kadınların parlamentodaki oranı

Var8: Kadınların ekonomik aktivitelere katılma oranının erkeklerin ekonomik aktivitelere katılma oranına göre yüzdesi

Var9: 15 Yaş altı nüfusun toplam nüfusa oranı

Var10: 65 yaş üstü nüfusun toplam nüfusa oranı

değişkenlerine ulaşılmıştır. Bu değişkenler aracılığı ile STATA programı kullanılarak yaklaşık 200 farklı model kombinasyonu denenmiş ve bunlardan en anlamlı olan model seçilerek sunulmuştur. Modelimizde az gelişmiş ülkeler ile gelişmiş ülkeler arasındaki farkı daha iyi vurgulayabilmek için temel seçenek olarak 3. seçenek olan az gelişmiş ülkeler seçilmiştir.

3.2 TAHMİN EDİLEN MODELİN SONUÇLARI

Eğitim çağındaki nüfusun öğrenim görme oranı, satın alma gücü paritesine göre düzeltilmiş kişi başına gayri safi milli hasıla ve kadınların parlamentodaki oranı değişkenlerinin açıklayıcı değişken olarak kullanıldığı model en uygun ve anlamlı model olarak seçilmiştir.

```
mlogit var1 var4 var5 var7 , base(3)
```

```
Iteration 0: log likelihood = -176.54843  
Iteration 1: log likelihood = -89.157967  
Iteration 2: log likelihood = -70.801583  
Iteration 3: log likelihood = -60.406891  
Iteration 4: log likelihood = -53.565073  
Iteration 5: log likelihood = -50.05823  
Iteration 6: log likelihood = -48.289751  
Iteration 7: log likelihood = -47.731193  
Iteration 8: log likelihood = -47.686829  
Iteration 9: log likelihood = -47.686535  
Iteration 10: log likelihood = -47.686535
```

Multinomial logistic regression

Number of obs = 169

LR chi2(6) = 257.72

Prob > chi2 = 0.0000

Pseudo R2 = 0.7299

Log likelihood = -47.686535

	var1	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
1	var4	.2331567	.0565092	4.13	0.000	.1224008	.3439127
	var5	.0028155	.0007704	3.65	0.000	.0013055	.0043255
	var7	-.0132013	.0739595	-0.18	0.858	-.1581593	.1317567
	_cons	-24.33412	4.576028	-5.32	0.000	-33.30297	-15.36527
2	var4	.1034743	.0335828	3.08	0.002	.0376532	.1692954
	var5	.002408	.0007652	3.15	0.002	.0009082	.0039077
	var7	-.128306	.0595726	-2.15	0.031	-.2450662	-.0115459
	_cons	-8.269564	2.214887	-3.73	0.000	-12.61066	-3.928466

(Outcome var1==3 is the comparison group)

Çok seçenekli modelin veri setine, bir başka değişle model yapısına uygunluğunun test edilmesi yani ilişkisiz alternatiflerin bağımsızlığının test edilmesi gerekmektedir. Bu uygunluk testi Hausman tarafından geliştirilmiştir. Analizimizde bulunan anlamlı modelimize her seçenek için ayrı ayrı bu test uygulanmış ve uygulanan test sonuçları aşağıda verilmiştir.

mlogit var1 var4 var5 var7 , base(3)

Iteration 0: log likelihood = -176.54843

Iteration 1: log likelihood = -89.157967

Iteration 2: log likelihood = -70.801583

Iteration 3: log likelihood = -60.406891

Iteration 4: log likelihood = -53.565073

Iteration 5: log likelihood = -50.05823

Iteration 6: log likelihood = -48.289751

Iteration 7: log likelihood = -47.731193

Iteration 8: log likelihood = -47.686829

Iteration 9: log likelihood = -47.686535

Iteration 10: log likelihood = -47.686535

Multinomial logistic regression Number of obs = 169
 LR chi2(6) = 257.72
 Prob > chi2 = 0.0000
 Log likelihood = -47.686535 Pseudo R2 = 0.7299

	var1	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
1	var4	.2331567	.0565092	4.13	0.000	.1224008	.3439127
	var5	.0028155	.0007704	3.65	0.000	.0013055	.0043255
	var7	-.0132013	.0739595	-0.18	0.858	-.1581593	.1317567
	_cons	-24.33412	4.576028	-5.32	0.000	-33.30297	-15.36527
2	var4	.1034743	.0335828	3.08	0.002	.0376532	.1692954
	var5	.002408	.0007652	3.15	0.002	.0009082	.0039077
	var7	-.128306	.0595726	-2.15	0.031	-.2450662	-.0115459
	_cons	-8.269564	2.214887	-3.73	0.000	-12.61066	-3.928466

(Outcome var1==3 is the comparison group)

. est store all

. mlogit var1 var4 var5 var7 if var1!=1, base(3)

Iteration 0: log likelihood = -70.667882
 Iteration 1: log likelihood = -35.475077
 Iteration 2: log likelihood = -28.865992
 Iteration 3: log likelihood = -24.566322
 Iteration 4: log likelihood = -22.164567
 Iteration 5: log likelihood = -21.095967
 Iteration 6: log likelihood = -20.908521
 Iteration 7: log likelihood = -20.903286
 Iteration 8: log likelihood = -20.903281

Multinomial logistic regression Number of obs = 115
 LR chi2(3) = 99.53
 Prob > chi2 = 0.0000
 Log likelihood = -20.903281 Pseudo R2 = 0.7042

	var1	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
2	var4	.1021432	.0336087	3.04	0.002	.0362712	.1680151
	var5	.0023858	.0007608	3.14	0.002	.0008948	.0038769
	var7	-.1259932	.0594295	-2.12	0.034	-.2424729	-.0095136
	_cons	-8.19335	2.207853	-3.71	0.000	-12.52066	-3.866038

(Outcome var1==3 is the comparison group)

. est store partial

. hausman partial all, alleqs constant

---- Coefficients ----				
	(b)	(B)	(b-B)	sqrt(diag(V_b-V_B))
	partial	all	Difference	S.E.
var4	.1021432	.1034743	-.0013311	.0013199
var5	.0023858	.002408	-.0000221	.
var7	-.1259932	-.128306	.0023128	.
_cons	-8.19335	-8.269564	.0762145	.

b = consistent under Ho and Ha; obtained from mlogit
 B = inconsistent under Ha, efficient under Ho; obtained from mlogit

Test: Ho: difference in coefficients not systematic

$$\begin{aligned} \chi^2(4) &= (b-B)'[(V_b-V_B)^{-1}](b-B) \\ &= 0.03 \end{aligned}$$

$$\text{Prob}>\chi^2 = 0.9999$$

. mlogit var1 var4 var5 var7 , base(3)

Iteration 0: log likelihood = -176.54843
 Iteration 1: log likelihood = -89.157967
 Iteration 2: log likelihood = -70.801583
 Iteration 3: log likelihood = -60.406891
 Iteration 4: log likelihood = -53.565073
 Iteration 5: log likelihood = -50.05823
 Iteration 6: log likelihood = -48.289751
 Iteration 7: log likelihood = -47.731193
 Iteration 8: log likelihood = -47.686829
 Iteration 9: log likelihood = -47.686535
 Iteration 10: log likelihood = -47.686535

Multinomial logistic regression

Number of obs = 169

LR chi2(6) = 257.72

Prob > chi2 = 0.0000

Pseudo R2 = 0.7299

Log likelihood = -47.686535

	var1	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
1	var4	.2331567	.0565092	4.13	0.000	.1224008	.3439127
	var5	.0028155	.0007704	3.65	0.000	.0013055	.0043255
	var7	-.0132013	.0739595	-0.18	0.858	-.1581593	.1317567
	_cons	-24.33412	4.576028	-5.32	0.000	-33.30297	-15.36527
2	var4	.1034743	.0335828	3.08	0.002	.0376532	.1692954
	var5	.002408	.0007652	3.15	0.002	.0009082	.0039077
	var7	-.128306	.0595726	-2.15	0.031	-.2450662	-.0115459
	_cons	-8.269564	2.214887	-3.73	0.000	-12.61066	-3.928466

(Outcome var1==3 is the comparison group)

. est store all

. mlogit var1 var4 var5 var7 if var1!=2, base(3)

Iteration 0: log likelihood = -59.646316
 Iteration 1: log likelihood = -19.078011
 Iteration 2: log likelihood = -10.792533
 Iteration 3: log likelihood = -6.5113315
 Iteration 4: log likelihood = -4.3860514
 Iteration 5: log likelihood = -3.3603914
 Iteration 6: log likelihood = -2.7037216
 Iteration 7: log likelihood = -2.1318525
 Iteration 8: log likelihood = -1.5602494
 Iteration 9: log likelihood = -.88198073
 Iteration 10: log likelihood = -.40812497
 Iteration 11: log likelihood = -.17474269
 Iteration 12: log likelihood = -.06356529
 Iteration 13: log likelihood = -.02297519
 Iteration 14: log likelihood = -.0083956
 Iteration 15: log likelihood = -.0030808
 Iteration 16: log likelihood = -.0011323
 Iteration 17: log likelihood = -.00041641
 Iteration 18: log likelihood = -.00015317
 Iteration 19: log likelihood = -.00005634
 Iteration 20: log likelihood = -.00002073
 Iteration 21: log likelihood = -7.625e-06
 Iteration 22: log likelihood = -2.805e-06
 Iteration 23: log likelihood = -1.032e-06

Iteration 24: log likelihood = -3.796e-07
 Iteration 25: log likelihood = -1.397e-07
 Iteration 26: log likelihood = -5.138e-08
 Iteration 27: log likelihood = -1.890e-08
 Iteration 28: log likelihood = -6.953e-09
 Iteration 29: log likelihood = -2.558e-09
 Iteration 30: log likelihood = -9.410e-10
 Iteration 31: log likelihood = -3.462e-10
 Iteration 32: log likelihood = -1.274e-10
 Iteration 33: log likelihood = -4.685e-11
 Iteration 34: log likelihood = -1.724e-11
 Iteration 35: log likelihood = -6.342e-12
 Iteration 36: log likelihood = -2.334e-12
 Iteration 37: log likelihood = -8.582e-13
 Iteration 38: log likelihood = -3.164e-13
 Iteration 39: log likelihood = -1.151e-13
 Iteration 40: log likelihood = -4.330e-14
 Iteration 41: log likelihood = -1.688e-14
 Iteration 42: log likelihood = -7.661e-15
 Iteration 43: log likelihood = -3.109e-15
 Iteration 44: log likelihood = -2.554e-15
 Iteration 45: log likelihood = -6.661e-16
 Iteration 46: log likelihood = -3.331e-16
 Iteration 47: log likelihood = -2.220e-16

Multinomial logistic regression	Number of obs	=	89
	LR chi2(3)	=	119.29
	Prob > chi2	=	0.0000
Log likelihood = -2.220e-16	Pseudo R2	=	1.0000

	var1	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
1							
	var4	18.07648	4336346	0.00	1.000	-8499064	8499100
	var5	.1166185	86882.44	0.00	1.000	-170286.3	170286.6
	var7	-6.991577	3.31e+07	-0.00	1.000	-6.49e+07	6.49e+07
	_cons	-1377.314	.	.	.	.	.

(Outcome var1==3 is the comparison group)

.
 . est store partial

. hausman partial all, alleqs constant

---- Coefficients ----				
	(b) partial	(B) all	(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
var4	18.07648	.2331567	17.84332	4336346
var5	.1166185	.0028155	.113803	86882.44
var7	-6.991577	-.0132013	-6.978376	3.31e+07

b = consistent under Ho and Ha; obtained from mlogit
B = inconsistent under Ha, efficient under Ho; obtained from mlogit

Test: Ho: difference in coefficients not systematic

$$\begin{aligned}\text{chi2}(3) &= (b-B)'[(V_b-V_B)^{-1}](b-B) \\ &= 0.00 \\ \text{Prob}>\text{chi2} &= 1.0000\end{aligned}$$

Test istatistiğinin Prob değeri nin 0,05'ten büyük olduğu tespit edilmiştir. Bu sonuç model yapımızın çok seçenekli logit modele elverişli olduğunu yani ilişkisiz alternatiflerin bağımsızlığı testinin olumlu sonuç verdiğinin kanıtı niteliğindedir.

Çok seçenekli logit modelinde değişkenlerin yorumu katsayılar üzerinden yapılamamaktadır. Bu yüzden modelimiz için marjinal etkiler ve nispi risk oranları hesaplanmıştır. Yorumlamalar bu göstergeler üzerinden yapılmıştır. Marjinal etkiler ve nispi risk oranları aşağıdaki gibi hesaplanmıştır.

mlogit var1 var4 var5 var7, base(3)

Iteration 0: log likelihood = -176.54843
Iteration 1: log likelihood = -89.157967
Iteration 2: log likelihood = -70.801583
Iteration 3: log likelihood = -60.406891
Iteration 4: log likelihood = -53.565073
Iteration 5: log likelihood = -50.05823
Iteration 6: log likelihood = -48.289751
Iteration 7: log likelihood = -47.731193
Iteration 8: log likelihood = -47.686829
Iteration 9: log likelihood = -47.686535
Iteration 10: log likelihood = -47.686535

Multinomial logistic regression Number of obs = 169
 LR chi2(6) = 257.72
 Prob > chi2 = 0.0000
 Log likelihood = -47.686535 Pseudo R2 = 0.7299

	var1	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
1							
	var4	.2331567	.0565092	4.13	0.000	.1224008	.3439127
	var5	.0028155	.0007704	3.65	0.000	.0013055	.0043255
	var7	-.0132013	.0739595	-0.18	0.858	-.1581593	.1317567
	_cons	-24.33412	4.576028	-5.32	0.000	-33.30297	-15.36527
2							
	var4	.1034743	.0335828	3.08	0.002	.0376532	.1692954
	var5	.002408	.0007652	3.15	0.002	.0009082	.0039077
	var7	-.128306	.0595726	-2.15	0.031	-.2450662	-.0115459
	_cons	-8.269564	2.214887	-3.73	0.000	-12.61066	-3.928466

(Outcome var1==3 is the comparison group)

mfx compute, predict(outcome(1))

Marginal effects after mlogit
 $y = \Pr(\text{var1}=1) (\text{predict}, \text{outcome}(1))$
 = .14084657

variable	dy/dx	Std. Err.	z	P> z	[95% C.I.]		X
var4	.0156927	.00417	3.77	0.000	.007526	.02386	68.8047
var5	.0000493	.00002	2.41	0.016	9.2e-06	.000089	9122.95
var7	.0139287	.00647	2.15	0.031	.001248	.02661	14.0367

. mfx compute, predict(outcome(2))

Marginal effects after mlogit
 $y = \Pr(\text{var1}=2) (\text{predict}, \text{outcome}(2))$
 = .85915343

variable	dy/dx	Std. Err.	z	P> z	[95% C.I.]		X
var4	-.0156927	.00417	-3.77	0.000	-.02386	-.007526	68.8047
var5	-.0000493	.00002	-2.41	0.016	-.000089	-9.2e-06	9122.95
var7	-.0139287	.00647	-2.15	0.031	-.02661	-.001248	14.0367

```
. mfx compute, predict(outcome(3))
Marginal effects after mlogit
      y = Pr(var1==3) (predict, outcome(3))
      = 4.734e-09
```

variable	dy/dx	Std. Err.	z	P> z	[95% C.I.]	X
var4	-5.76e-10	.	.	.	-5.8e-10 -5.8e-10	68.8047
var5	-1.17e-11	.	.	.	-1.2e-11 -1.2e-11	9122.95
var7	5.31e-10	.	.	.	5.3e-10 5.3e-10	14.0367

Nispi Risk Oranı tahminleri

```
mlogit var1 var4 var5 var7, base(3) rrr
```

```
Iteration 0: log likelihood = -176.54843
Iteration 1: log likelihood = -89.157967
Iteration 2: log likelihood = -70.801583
Iteration 3: log likelihood = -60.406891
Iteration 4: log likelihood = -53.565073
Iteration 5: log likelihood = -50.05823
Iteration 6: log likelihood = -48.289751
Iteration 7: log likelihood = -47.731193
Iteration 8: log likelihood = -47.686829
Iteration 9: log likelihood = -47.686535
Iteration 10: log likelihood = -47.686535
```

Multinomial logistic regression	Number of obs	=	169
	LR chi2(6)	=	257.72
	Prob > chi2	=	0.0000
Log likelihood = -47.686535	Pseudo R2	=	0.7299

	var1	RRR	Std. Err.	z	P> z	[95% Conf. Interval]
1	var4	1.262579	.0713473	4.13	0.000	1.130207 1.410455
	var5	1.002819	.0007726	3.65	0.000	1.001306 1.004335
	var7	.9868854	.0729896	-0.18	0.858	.8537138 1.140831
2	var4	1.109017	.0372439	3.08	0.002	1.038371 1.18447
	var5	1.002411	.0007671	3.15	0.002	1.000909 1.003915
	var7	.8795841	.0523991	-2.15	0.031	.7826527 .9885205

(Outcome var1==3 is the comparison group)

Yapılan çok seçenekli logit model sonuçlarına göre eğitim çağındaki nüfusun öğrenim görme oranındaki artışın, ülkenin gelişmiş ülkeler grubunda olma olasılığını arttırdığı sonucuna varılmıştır. Bu çerçevede eğitim çağındaki nüfusun öğrenim görme oranında meydana gelecek %1 lik artış sözkonusu ülkenin gelişmiş ülkeler kategorisinde olma olasılığını %1.56 arttıracaktır.

Analiz sonuçları gelişmekte olan ülkeler açısından incelendiğinde eğitim çağındaki nüfusun öğrenim görme oranında meydana gelecek artış karşısında olasılığın negatif yönlü olarak değiştiğini; bu çerçevede de değişkenimizdeki %1'lik artışın ülkenin gelişmekte olan ülke olma olasılığını %-1.56 oranında düşüreceğini görüyoruz. Çünkü gelişmekte olan ülkeler kategorisinde olma olasılığının azalması, gelişmiş ülkeler kategorisinde olma olasılığının artmasının işareti olarak yorumlanabilmektedir.

Az gelişmiş ülkeler açısından analiz sonuçları incelendiğinde eğitim çağındaki nüfusun öğrenim görme oranındaki değişimin ülkenin az gelişmiş ülke olma olasılığını neredeyse hiç etkilemediği sonucuyla karşılaşmaktayız.

Eğitim kurumları bazı istisnalar dışında büyük ölçüde kar amacıyla organize edilerek yönetilmemeleri ve öğrencilerinin de tüketici konumunun genel gereği olan maliyetlere katlanmalarının söz konusu olmamasına karşılık ekonominin tam da içerisinde bir konumda yer almaktadır. Çünkü büyük ölçüde kamusal finansmana dayanan harcamalar, öğrencilerin gelecekteki verimliliklerinin artmasına yönelik yatırım konumundadır.¹

Eğitimin ve beraberinde de bilginin Sosyo-Ekonomik ve Sosyo-Kültürel yaşamda ne kadar önemli bir konumda olduğunu anlamak için günümüzün komplike yapısına bakmamız da şart değildir. Günümüzün bilgi toplumunun düşünülemediği

¹ Amiran Kurtkan, **Sosyolojik Açıdan Eğitim Yolu ile Kalkınmanın Esasları**, İstanbul, İ.Ü. İktisat Fakültesi Yayınları, No:1742/312, 1972, s.54.

bir dönemde, MÖ. 6. yüzyılda Atinaya yönelik beyin göçüyle karşılaşılması² son derece düşündürücüdür. Günümüze doğru gelindiğinde, elbette önem seviyesinin artacağı bir gerçektir. Örneğin ABD'ye 1907-23 döneminde göç eden 6.9 milyon kişinin sadece %2.6'sı mesleki vasıflara sahipken, 1947-65 döneminde gelen yabancı göçmenlerin %22'sinin mesleki ve teknik işgücünü oluşturduğu görülmüştür.³

Dolayısıyla sosyo-ekonomik ve sosyo-kültürel yaşamda eğitimin öneminin zaman ve mekandan bağımsız olduğu, buna karşılık söz konusu önemin her geçen gün arttığını savunmak yanlış olmayacaktır.

Günümüzün teknolojik gelişim eksenine paralel olarak nitelikli işgücü talebinin artışının yaşanmasını gözlemlemekteyiz. Dolayısıyla hem yaşanan sosyo-ekonomik dönüşümü iş gücü niteliğini tetiklemesini, hem de işgücü niteliğinin sosyo-ekonomik dönüşümden tetiklenmesi ile karşılaşıyoruz.

Günümüzde teknolojik gelişim son derece yüksek ivme kazanmış durumdadır. Gelişmiş ülkelerde işletmeler bünyesinde, görevi sadece teknoloji geliştirmek olan çok sayıda araştırmacı görev almaktadır. Gelişmişlik düzeyi düştükçe de bu düzeyde ciddi düşüşlerin yaşandığı gözlemlenmektedir. Sanayi ihracatı içerisinde ileri teknoloji ürünlerinin ağırlığının da arttığı görülmektedir. Bu çerçevede vasıfsal artış gün geçtikçe mevcut yüksek önemini daha da arttırmaktadır. ABD eski çalışma bakanı olan Prof. Dr. Robert Reich'in, sorunları saptayarak çözümler üretebilen veya çözümlerin üretilmesine aracılık eden elit zümre olarak tanımladığı işgücünün, Amerikan işgücü içerisindeki payı 1950'de %8 iken bu oranın günümüzde %20'ye çıktığı düşünüldüğünde⁴, eğitimin neden gelişmişlikte bu denli ayırt edici bir özelliğe sahip olduğunu anlayabiliriz.

² Berrak Kurtuluş, "Beyin Göçü: Geçmişte Günümüzde ve Gelecekte", **İstanbul Üniversitesi İktisat Fakültesi Mecmuası**, 46. cilt, 38/13-93, 1996, s.172.

³ A.e., s.173-175.

⁴ Cem Kozlu, **Türkiye Mucizesi için Vizyon Arayışları ve Asya Modelleri**, 3. Baskı, Ankara, Türkiye İş Bankası Kültür Yayınları, 1995, s.38.

Günümüzde eğitime verilen önem bir zorunluluk halini almıştır. Bilgi çağı olarak adlandırılan 21. yüzyılda okulda edinilen bilgiler de teknolojik gelişimin süratliliğinden dolayı hızla eskiyip yetersiz kalmakta olup bu sebepten değişen her teknolojinin yeniden öğrenilmesi amacını güden ömür boyu iş yaşamında eğitim kavramı türetilmiştir.⁵

Bütün bu göstergeler ışığında eğitim çağındaki nüfusun öğrenim görme oranlarının, ülkelerin gelişmişlik düzeyleri açısından açık bir referans oluşturduğunu ana hatlarıyla görebiliyoruz. Bu açıdan yaptığımız çok seçenekli logit model analizinden elde ettiğimiz sonuçlarla örtüşür niteliktedir.

Yaptığımız analiz sonuçlarını satınalma gücü paritesine göre düzeltilmiş fert başı GSYİH seviyesindeki artışın, ülkenin gelişmiş ülke olma olasılığını oldukça düşük düzeyde arttırdığı sonucu ile karşılaşmaktayız. Bu çerçevede satın alma gücü paritesine göre düzeltilmiş fert başına GSYİH seviyesinde meydana gelecek %1'lik artış karşısında, ülkenin gelişmiş ülkeler kategorisinde olma olasılığı %0.05 oranında arttıracaktır.

Gelişmekte olan ülkeler açısından analiz sonuçlarını incelediğimizde satınalma gücü paritesine göre düzeltilmiş fert başı GSYİH seviyesinde meydana gelecek %1'lik artış karşısında negatif yönlü olasılık etkisine yol açtığını; bu çerçevede de değişkendeki %1'lik artışı karşısında, gelişmekte olan ülkeler kategorisinde olma olasılığını %0.05 oranında düşüreceğini görmekteyiz. Negatif yönlü ilişki doğaldır çünkü gelişmekte olan ülkeler kategorisinde olma olasılığının azalması, gelişmiş ülkeler kategorisinde olma olasılığının artışının işareti olarak görülebilir.

Az gelişmiş ülkeler açısından analiz sonuçları incelendiğinde satınalma gücü paritesine göre düzeltilmiş fert başı GSYİH seviyesindeki değişimin, bu kategoride olma olasılığını hiç etkilemediğini görüyoruz.

⁵ Cafer Tayyar Sadıklar, **2000'li Yıllar Dünya ve Türkiye**, 2. Baskı, Ankara, TBMM Vakfı Yayınları, 1995, s.115.

On yıl arayla tüm dünya ülkelerini içererek hazırlanmış aynı formatlı iki araştırma sonucunda fert başına GSMH seviyesinin gelişmişlik düzeyi üzerinde Bedriye Saraçoğlu tarafından 1990 yılında yapılan analiz sonucuna göre ilk önem sırasında, Karpat-Açıkgöz tarafından 2000 yılında yapılan diskriminant analizi sonucuna göre 2. önem sırasına sahip olduğu tespit edilmiştir⁶. Bu çalışmalarda satın alma gücü paritesine göre düzeltme işleminin yapılmadığı çalışmalarda benzer karakterli değişkenin önem sırasının bu denli yüksek olması oldukça ilgi çekicidir.

Yaptığımız çok seçenekli logit model sonuçlarının bu analiz sonuçları ile arasında ters yönlü bir sonuç ile karşılaşılmaktadır. Sonuçlar birleşmiş milletler ile IMF-Dünya Bankası eksenli gelişmişlik düzeyi algılamasındaki farkı yansıtmaktadır.

Burada iktisadi büyüme ile iktisadi kalkınma arasındaki fark önem kazanmaktadır. İktisadi büyüme genel olarak gelişmiş sanayileşmiş ülkelerde ulusal gelirin bir dönemden diğerine artma oranıdır ve ekonominin üretim kapasitesinin büyümesidir. İktisadi kalkınma ise genellikle az gelişmiş ülkeler için kullanılan ve bir ülkenin toplumsal ve iktisadi yapısının dönüşümünü ifade eder. Kalkınma ulusal gelir ve üretimin sayısal olarak artması yanında kurumlardaki köklü değişiklikleri, iktisadi ve toplumsal yapının yeniden düzenlenmesini, halkın değer yargılarında, dünya görüşünde ve davranış kalıplarındaki değişiklikleri kapsar.⁷

Buradan hareketle çok seçenekli logit modeli ile yaptığımız kalkınmışlık düzeyi analizinde Birleşmiş Milletlerin kullandığı satın alma gücü paritesine göre düzeltilmiş fert başına düşen GSYİH seviyelerinin cari GSYİH veya cari GSMH düzeylerine nazaran tercih edilmesinin daha komplike ve sosyal içerikli olduğunu söylemek mümkündür.

⁶ Bedriye Saraçoğlu, “Ülkelerin Ekonomik Kalkınmışlık Düzeyleri Açısından İncelenmesi”, **Hacettepe Üniversitesi İİBF Dergisi**, Cilt 10, 1992, s.31.

Gaye Karpat, Şenay Açıkgöz, “Ülkelerin Kalkınmışlık Düzeyleri Açısından Değerlendirilmesi”, **V. Ulusal Ekonometri ve İstatistik Sempozyumu 2001** www.idari.cu.edu.tr/sempozyum/

⁷ Rona Turanlı, Tamer İşgüden, **Ansiklopedik Ekonomi Sözlüğü**, 2.baskı, İstanbul, Bilim Teknik Yayınevi, t.y., s.151-152.

Bu önemli bulgular çerçevesinde yaptığımız çok seçenekli logit model bulguları ile ülkelerin kalkınmışlık düzeylerini etkileyen unsurları saptamak yanında değişken seçiminde de önemli bulgular elde etmiş olmaktadır.

Yapılan çok seçenekli logit modeli sonuçlarına göre kadın parlamenterlerin parlamentodaki temsil oranlarındaki artışın, ülkenin gelişmiş ülkeler grubunda olması olasılığını arttırdığı sonucuyla karşılaşmaktayız. Bu çerçevede kadın parlamenterlerin parlamentodaki temsil oranda meydana gelecek %1'lik artış karşısında, ülkenin gelişmiş ülkeler kategorisinde olma olasılığını %1.39 oranında arttıracaktır.

Gelişmekte olan ülkeler kategorisi açısından analiz sonuçlarına bakıldığında kadın parlamenterlerin parlamentodaki temsil oralarında meydana gelecek artış karşısında negatif yönlü olasılık varlığı sözkonusudur. Yani bu çerçevede değişkenimizde %1'lik artış karşısında, gelişmekte olan ülkeler kategorisinde olma olasılığını %1.39 oranında düşüreceğini görüyoruz. Bu sonuçta iktisaden mantıklıdır. Çünkü gelişmekte olan ülkeler kategorisinde olma olasılığının azalması gelişmiş ülkeler kategorisinde olma olasılığının arttığının işareti olarak yorumlanabilir.

Az gelişmiş ülkeler açısından analiz sonuçlarına baktığımızda kadın parlamenterlerin parlamentodaki temsil oranlarındaki değişimin bu kategoride olma olasılığını etkilemediği sonucuyla karşılaşyoruz.

Kadın nüfusun eğitimi ile ilgili gelişmelere bakıldığında çarpıcı sonuçlarla karşılaşmaktadır. ABD'nin bütünü için lisans düzeyinde derece veren yüksek öğrenim kurumlarında okuyan kadın nüfusun 2010 yılına gelindiğinde erkek nüfusun 4/3'üne ulaşacağı tahmin edilmektedir. Erkek öğrenci nüfusunun kadın öğrenci nüfusu tarafından geçildiği yapının daha 1890'da aynı oranda fakat bu defa kadın

öğrenci nüfusu aleyhinde olduğunu ABD milli eğitim bakanlığının istatistiklerinden açıkça görebiliyoruz.⁸

Kadınların genel olarak gerek iş yaşamında, gerekse de eğitim alma konusunda istikrarlı bir şekilde artmasıyla paralel olarak siyasal yaşamda da ağırlıkları gün geçtikçe artmıştır. 1932 İngilteresinde yapılan reformla sadece 500 bine ulaşabilen seçmen sayısının 1884'te 7 milyona sıçramasının ardından birinci dünya savaşından sonra 30 yaşından büyük tüm kadınlara ve 21 yaşını geçen tüm erkeklere seçme hakkı tanınmış, söz konusu yaş eşitsizliği de 1928'de yapılan değişiklikle kaldırılarak her iki cinsiyet içinde 21 olarak kabul edilmiştir.⁹

Günümüz gelişmiş ülkelerinde çalışma sürelerinin giderek kısalıyor olması, yarım günlük çalışmanın yaygınlaşması bu ülkelerdeki kadın iş gücünün ağırlığının giderek artmasına destek olmaktadır.¹⁰ Nitekim OECD ülkelerinde 1980-1992 döneminde kadın istihdamın, nüfus artışının çok üzerinde bir seviyede %2 oranında yıllık artış kaydettiğini görüyoruz.¹¹

1998 yılında AB genelindeki bölgesel hükümetler içerisinde kadın nüfus ağırlığı %24,5'e ulaşmış, bu oran bölgesel parlamentolarda %27,8 olarak gerçekleşmiş, ulusal parlamentolara katılım oranı da %17,5 oranında gerçekleşmiştir.¹² 2004 yılında yapılan son seçimlerde de Avrupa Birliği Parlamentosundaki kadın milletvekili oranı %30,3'e çıkmıştır.¹³ Dolayısı ile kadınların parlamentolardaki temsil oranlarının, ülkelerin gelişmişlik düzeyleri açısından açık bir referans oluşturduğunu ana hatları ile görebiliyoruz.

⁸ World Almanac Books, **The world almanac and book of facts**, New York, 2002, pp.237.

⁹ Esat Çam, **Devlet Sistemleri**, İstanbul, DER, 1993, s.25-26.

¹⁰ Sabahat Avşar, Şule Özkan, "Kadının çalışma hayatına ilişkin görüşleri Erzurum'da uygulama", **İ.Ü. İktisat Fakültesi Mecmuası Sabahattin Zaim Armağan**, 1994/ B-3 C.1-4, s.522-523.

¹¹ Hacer Ansal, "Ekonomik yeniden yapılanma sürecinde kadın emeği", **Türk-iş 1997 yılı**, 2. cild, İstanbul, 1997, s.165.

¹² İktisadi Kalkınma Vakfı, **AB ve eşitlik**, İstanbul, 2002, s.6.

¹³ İktisadi Kalkınma Vakfı, **İKV Bülteni**, 16-30 Haziran 2004, s.4.

Yapılan çok seçenekli model sonuçları çerçevesinde durumu ampirik bir zemine oturtmuş olduk. Bu nedenle çok seçenekli logit analiz sonuçlarının sosyo-ekonomik yapı tanımlamalarında ve genel kuramların geçerliliklerinin test edilmesinde anlamlı olduğu sonucuna bir kez daha varabiliriz.

Yapılan çok seçenekli analiz sonuçlarına göre eğitim çağındaki nüfusun öğrenim görme oranı itibarı ile gelişmiş ülkelerde az gelişmiş ülkelere göre 1.26 kat ve gelişmekte olan ülkelere 1.1 kat daha fazla etkilidir.

Satınalma gücü paritesine göre düzeltilmiş fert başı GSYİH düzeyleri itibarı ile gelişmiş ülkelerde az gelişmiş ülkelere göre 1 kat, gelişmekte olan ülkeler kategorisindeki ülkelere de yine 1 kat daha fazla etkilidir.

Kadın parlamenterlerin parlamentodaki temsil oranları itibarı ile az gelişmiş ülkelere göre gelişmiş olan ülkelere 0.9 kat etkilidir.

SONUÇ

Ülkelerin kalkınmışlık düzeylerini etkileyen faktörleri belirlemek için yaptığımız çalışmada, ülkelerin gelişmişlik düzeyleri nitel bağımlı değişken olarak belirtilmiştir. Nitel değişken bağımlı değişken olarak alınarak kurulacak modeller söz konusu olduğunda doğrusal olasılık, logit ve probit modelleri hatırlanmaktadır.

Nitel bağımlı değişken iki değer alırsa, iki seçenekli doğrusal olasılık, logit ve probit modelleri kullanılabilir. Bu tür modellerden doğrusal olasılık modelinin sorunları vardır. Bu nedenle logit ve probit modelleri geliştirilerek doğrusal olasılık modelinde karşılaşılan sorunlar çözülmüştür.

İki değer alan bağımlı nitel değişkenli modellerde bağımlı değişken 0 – 1 aralığında değer aldığından, modelden tahmin edilen teorik değerler olasılıklar olarak ele alınabilirler. Bu tür modeller tahmin amacı ile kullanılabilirler, tahmin edilen bağımsız değişken katsayıları klasik regresyon modellerine benzer anlamlar ifade ederler.

Nitel değişkenler ikiden fazla değer de alabilirler. Bu durum çok seçeneklilik olarak adlandırılır. Çok seçenekli nitel değişkenler bağımlı değişken olarak regresyon modellerinde yer alırsa çok seçenekli doğrusal olasılık, çok seçenekli logit ve çok seçenekli probit modelleri söz konusu olacaktır. Çok seçenekli doğrusal olasılık modellerinde de iki seçenekli doğrusal olasılık modellerinin sorunlarına benzer sorunlar vardır. Bu nedenle çok seçenekli logit ve probit modelleri kullanılacaktır.

Çok seçenekli logit ve probit modellerinde bağımlı değişkenin teorik değerleri, iki seçenekli bağımlı değişkenli modellerde olduğu gibi olasılık olarak ele alınıp yorumlanamaz. Burada teorik değerler farklı değerler olacağından bu tür modeller tahmin için kullanılamaz. Çok seçenekli modellerin yorumlanması marjinal

etkiler ve nispi risk oranları üzerinden yapılabilecektir.çok seçenekli logit ve probit modellerinde de iki seçenekli logit ve probit modellerinde olduğu gibi istatistiksel anlamlılığın testi, belirlilik katsayısı gibi ölçülerin tahmini farklı olarak yapılır.

Çok seçenekli logit modelde model yapısının verilere uygunluğunun test edilmesi gerekir. Bu test Hausman testidir. Sıralı çok seçenekli modellerde ise veri yapısı gereği bu testin uygulanmasına gerek olmayacaktır.

Çalışmanın uygulamasında BM'nin gelişmiş, gelişmekte olan ve az gelişmiş ülke ayrımından yararlanılarak çok seçenekli logit model uygulaması yapılmıştır. Denenen 200'e yakın model arasından en anlamlı olan model tercih edilmiştir. Bu model sonucuna göre ülkelerin gelişmişlik düzeylerinin eğitim çağındaki nüfusun öğrenim görme oranı, gayri safi yurt içi hasıla ve kadınların parlamentodaki oranı faktörleri tarafından etkilendiği tespit edilmiştir. Modelin parametrelerin tek tek anlamlılıklarının testi t testi ile, toplu olarak anlamlılıklarının testi ise LR testi ile test edilmiş ve anlamlı sonuçlara ulaşılmıştır. Ayrıca tahminin uygunluğu ölçüsü olarak pseudo-R² kullanılmıştır. Modelde ilişkisiz alternatiflerin bağımsızlığının testi için hausman tarafından geliştirilen test uygulanmış ve test sonucunda çok seçenekli logit modelin uygulanabilirliği tespit edilmiştir. Bu durum model yapısının verilere uygunluğunun kanıtı niteliğindedir. Sonrasında modelin yorumlanabilmesi için nispi risk oranları ve marjinal etkiler hesaplanmıştır.

Model sonuçlarından elde edilen bir başka çarpıcı sonuçta IMF ve Dünya Bankasının yapmış olduğu sınıflandırma ile BM'nin yapmış olduğu sınıflandırmada kullandığı GSYİH'nin, satın alma gücü paritesi tarafından düzeltilip düzeltilmemesine göre çok büyük oranda farklılık sergilediğinin tespit edilmiş olmasıdır.

KAYNAKÇA

- Abdel-Aty, Mohamed A.: “Using Ordered Probit Modeling to Study the Effect of ATIS on Transit Ridership,” **Transportation Research Part C**, Vol. 9 , 2001, s. 265-277.
- Akkaya, Şahin: **Ekonometri 2**, İzmir, Anadolu Matbaacılık, 1991.
- Aldrich John. H., Nelson Forrest. D: **Linear Probability, Logit and Probit Models**, California, Sage Publications, 1984.
- Ansal, Hacer: “Ekonomik Yeniden yapılanma sürecinde kadın emeği”, **Türk-iş 1997 yılı**, 2. cild, İstanbul, 1997, s.165-184.
- Avşar Sabahat, Özkan Şule: “Kadının çalışma hayatına ilişkin görüşleri Erzurum’da uygulama”, **İ.Ü. İktisat Fakültesi Mecmuası Sabahattin Zaim Armağanı**, 1994\ B-3 C.1-4, İstanbul, s.521-534.
- Bolduc, Denic: “A Practical Technique to Estimate Multinomial Probit Models in Transportation,” **Transportation Research Part B**, Vol. 33, 1999, s.63-79.
- Borooah, Vani K: **Logit and Probit Ordered and Multinomial Models**; California, Sage Publications, 2002.
- Chintagunta, Pradeep K.: “Estimating A Multinomial Probit Model of Brand Choice Using the Method of Simulated Moments,” **Marketing Science**, Vol.11 No.4 Fall 1992, s. 386-407.

- Cramer, J.S.: **An Introduction The Logit Model for Economists**, New York, Edward Arnold, 1991, s.43.
- Çam, Esat: **Devlet Sistemleri**, İstanbul, Der yayınevi, 1993.
- Frone, Michael R.: “Regression Models for Discrete and Limited Dependent Variables,” **Research Methods Forum**, No.2, Summer 1997, s. 1-5.
- Greene, William H: **Econometric Analysis**, 5. bs., New Jersey, Prentice Hall, 2003.
- Gujarati, Damador N.: **Temel Ekonometri**, Çev. Ümit Şenesen, Gülay Günlük Şenesen, İstanbul, Literatür, 1999.
- Güriş, Selahattin, Çağlayan, Ebru: **Ekonometri Temel Kavramlar**, 2. bs., İstanbul, DER, 2005.
- Hardin James, Hilbe Joseph: **Generalized Linear Models and Extensions**, Texas, Stata Pres, 2001.
- Hilmer, Michael J.: “A Comparison of Alternative Specifications of the Collage Attendance Equation with an Extension to two-stage Selectivity-Correction Models,” **Economics of Education Review**, Vol:20, 2001. s. 263-278.
- Hilmer, Michael J.: “A Comparison of Alternative Specifications of the Collage Attendance Equation with an Extension to two-stage Selectivity-Correction Models,” **Economics of Education Review**, Vol:20, 2001, s. 263-278.
- Hosmer, David W., Lemashof, Stanley: **Applied Logistic Regression**; New York, John Wiley and Sons Publication, s.266.

- İktisadi Kalkınma Vakfı: **AB ve eşitlik**, İstanbul, 2002.
- İktisadi Kalkınma Vakfı: İKV Bülteni, İstanbul, 16-30 Haziran 2004.
- Kara Maria, Kockelman, Young-Jun, Kweon: “Driver Injury severity: An Application of Ordered Probit Models”, **Accident Analysis & Prevention**, vol.34, 2002, s.313-321.
- Karpat Gaye, Açıkgöz Şenay: “Ülkelerin Kalkınmışlık Düzeyleri Açısından Değerlendirilmesi”, **V. Ulusal Ekonometri ve İstatistik Sempozyumu**, 2001
www.idari.cu.edu.tr/sempozyum/
- Kennedy, Peter: **A Guide to Econometrics**, 4. bs., Oxford, Blackwell Publishers, 1998.
- Kmenta, Jan: **Elements of Econometrics**, Second Edition, Maxwell Macmillan International Edition, NewYork, 1990.
- Kozlu, Cem: **Türkiye Mucizesi için Vizyon Arayışları ve Asya Modelleri**, 3. bs. , Ankara Türkiye İş Bankası Kültür Yayınları, 1995.
- Kurtkan, Amiran: **Sosyolojik Açıdan Eğitim Yolu ile Kalkınmanın Esasları**, İstanbul, İ.Ü. İktisat Fakültesi Yayınları, No:1742/312, 1972.
- Kurtuluş, Berrak: “Beyin Göçü: Geçmişte Günümüzde ve Gelecekte”, **İstanbul Üniversitesi İktisat Fakültesi Mecmuası**, 46. cilt, 38/13-93, İstanbul, 1996, s.171-186.
- Layton Allan P., Katsuura Masaki: “Comparison of Regime Switching, Probit and Logit Models in Dating and Forecasting US Business Cycles,” **International Journal of Forecasting**, Vol.17, s.403-417.

- Maddala, G.S.: **Limited Dependent and Qualitative Variables in Econometrics**, Cambridge, Cambridge University Press, 1983.
- Menard, Scott: **Applied logistic Regression Analysis**, 2. bs., California, Sage Publications, 2002.
- Özer, Hüseyin: **Nitel Değişkenli Ekonometrik Modeller Teori ve Uygulama**, Ankara, Nobel Yayın Dağıtım, 2004.
- Sadıklar, Cafer Tayyar: **2000’li Yıllar Dünya ve Türkiye**, 2. Baskı, Ankara, TBMM Vakfı Yayınları, 1995.
- Saraçoğlu, Bedriye: “Ülkelerin Ekonomik Kalkınmışlık Düzeyleri Açısından İncelenmesi”, **Hacettepe Üniversitesi İİBF Dergisi**, Cilt 10, 1992 , s.17-54.
- Thomas, R.L.: **Modern Econometrics an İntruduction**, Addison Wesley, England, 1997.
- Tiwari Piyush, Kawakami Tetsu: “Modes of Commuting in Mumbai: A Discrete Choice Analaysis,” **The Applied Regional Science Conference**, Vol.13, No.1, March 200, s.34-45.
- Turanlı Rona, İşgüden Tamer: **Ansiklopedik Ekonomi Sözlüğü**, 2.bs. , İstanbul Bilim Teknik Yayınevi, t.y.
- Üçdoğruk, Şenay: **Multinomial Logit Model: Bölgeler Arası İşgücü Tercihleri Analizi**, İzmir, Dokuz Eylül Üniversitesi Yayınları, 2002.

William E.Griffiths, R.Carter Hill, **Learning and Practicing Econometrics**,
George G. Judge: Canada, John Willey and Sons, 1993.

World Almanac Books: **The world almanac and book of facts**, New
York, 2002.

Yatchew Adonis, Griliches Zvi: “Specification Error in Probit Models,” **The
Review of Economic and Statistics**, Vol. 67,
No.1, Şubat 1985 s.134-139.

EK: Çalışmada Kullanılan Veriler

		yaşam beklentisi	yetişkinlerin okuma oranı	Eğitim Çağındaki Nüfusun Öğrenim Görme Oranı	KBGSYİH
ÜLKELER	VAR1	VAR2	VAR3	VAR4	VAR5
Norveç	1	78,9	100	98	36600
İsveç	1	80	100	100	26050
Avusturalya	1	79,1	100	100	28260
Kanada	1	79,3	100	95	29480
Hollanda	1	78,3	100	99	29100
Belçika	1	78,7	100	100	27570
İzlanda	1	79,7	100	90	29750
Amerika	1	77	100	92	25752
Japonya	1	81,5	100	84	26940
İrlanda	1	76,9	100	90	36360
İsviçre	1	79,1	100	88	30010
İngiltere	1	78,1	100	100	26150
Finlandiya	1	77,9	100	100	26190
Avusturya	1	78,5	100	91	29220
Lüksemburg	1	78,3	100	75	61190
Fransa	1	78,9	100	91	26920
Danimarka	1	76,6	100	96	30940
Yeni Zelanda	1	78,2	100	100	21740
Almanya	1	78,2	100	88	27100
İspanya	1	79,2	97,7	92	21460
İtalya	1	78,7	98,5	82	26430
İsrail	1	79,1	95,3	92	19530
Çin	1	79,9	93,5	72	26910
Yunanistan	1	78,2	97,3	86	18720
Singapur	1	78	92,5	87	24040
Portekiz	1	76,1	92,5	93	18280
Slovenya	1	76,2	99,7	90	18540
Kore	1	75,4	97,9	92	16950
Barbodos	1	77,1	99,7	88	15290
Kıbrıs Rum Kesimi	1	78,2	96,8	74	18360
Malta	1	78,3	92,6	77	17640
Çek Cumhuriyeti	1	75,3	100	78	15780
Brunei Darussalam	1	76,2	93,9	73	19210
Arjantin	1	74,1	97	94	10880
Seychelles	1	72,7	91,9	85	18232
Estonya	1	71,6	99,8	96	12260
Pollanda	1	73,8	99,7	90	10560

Macaristan	1	71,7	99,3	86	13400
Saint Kitts and Nevis	1	70	97,8	97	12420
Bahreyn	1	73,9	88,5	79	17170
Letonya	1	72,5	99,6	90	10320
Slovakya	1	73,6	99,7	74	12840
Şili	1	76	95,7	79	9820
Kuveyt	1	76,5	82,9	76	16240
Kostarika	1	78	95,8	69	8840
Uruguay	1	75,2	97,7	85	7830
Katar	1	72	84,2	82	19844
Hırvatistan	1	74,1	98,1	73	10240
Birleşik Arap Emirlikleri	1	74,6	77,3	68	22420
Litvanya	1	70,9	99,7	87	9210
Bahama	1	67,1	95,5	74	17280
Küba	1	76,7	96,9	78	5259
Meksika	1	73,3	90,5	74	8970
Trinidad Tobago	1	71,4	98,5	64	9430
Antigua Barbuda	1	73,9	85,8	69	1920
Bulgaristan	2	70,9	98,6	76	7130
Rusya	2	66,7	99,6	88	8230
Libyan Arab Jamahiriya	2	72,6	81,7	97	7570
Malezya	2	73	88,7	70	9120
Makedonya	2	73,5	96	70	6470
Panama	2	74,6	92,3	73	6170
Belarus	2	69,9	99,7	88	5520
Tonga	2	68,4	98,8	82	6850
Mauritius	2	71,9	84,3	69	10810
Albania	2	73,6	98,7	69	4830
Bosna Hersek	2	74	94,6	64	5970
Suriname	2	71	94	74	6590
Venezuela	2	73,6	93,1	71	5380
Romanya	2	70,5	97,3	68	6560
Ukrayna	2	69,5	99,6	84	4870
Saint Lucia	2	72,4	94,8	74	5300
Brazilya	2	68	86,4	92	7770
Kolombiya	2	72,1	92,1	68	6370
Oman	2	72,3	74,4	63	13340
Samoa	2	69,8	98,7	69	5600
Tayland	2	69,1	92,6	73	7010
Sudi Arabistan	2	72,1	77,9	57	12650
Kazakistan	2	66,2	99,4	81	5870
Jameika	2	75,6	87,6	75	3980
Lübnan	2	73,5	86,5	78	4360
Fiji	2	69,6	92,9	73	5440
Ermenistan	2	72,3	99,4	72	3120
Filipinler	2	69,8	92,6	81	4170
Maldivler	2	67,2	97,2	78	4798
Peru	2	69,7	85	88	5010

Türkmenistan	2	66,9	98,8	81	4300
Saint Vincent	2	74	83,1	64	5460
Türkiye	2	70,4	86,5	68	6390
Paraguay	2	70,7	91,6	72	4610
Jordan	2	70,9	90,9	77	4200
Azerbaycan	2	72,1	97	69	3210
Tunus	2	72,7	73,2	75	6760
Grenada	2	65,3	94,4	65	7280
China	2	70,9	90,9	68	4580
Dominica	2	73,1	76,4	74	5640
Srilanka	2	72,5	92,1	65	3570
Georgia	2	73,5	100	69	2260
Dominican Republic	2	66,7	84,4	77	6640
Belize	2	70,5	76,9	71	6080
Ekvator	2	70,7	91	72	3580
İran	2	70,1	77,1	69	6690
Occupied Palestinian Territories	2	72,3	90,2	79	
El Salvador	2	70,6	79,7	66	4890
Guyana	2	63,2	96,5	75	4260
Cape Verde	2	70	75,7	73	5000
Syrian	2	71,7	82,9	59	3620
Özbekistan	2	69,5	99,3	76	1670
Algeria	2	69,5	68,9	70	5760
Equatorial Guinea	2	49,1	84,2	58	30130
Kırgızistan	2	68,4	97	81	1620
Endonezya	2	66,6	87,9	65	3230
Vietnam	2	69	90,3	64	2300
Moldova	2	68,8	99	62	1470
Bolivya	2	63,7	86,7	86	2460
Honduras	2	68,8	80	62	2600
Tacikistan	2	68,6	99,5	73	980
Mongolia	2	63,7	97,8	70	1710
Nicaragua	2	69,4	76,7	65	2470
Kuzey Afrika	2	48,8	86	77	10070
Mısır	2	68,6	55,6	76	3810
Guatamala	2	65,7	69,9	56	4080
Gabon	2	56,6	71	74	6590
Sao Tome and Principe	2	69,7	83,1	62	1317
Solomon adaları	2	69	76,6	50	1590
Morocco	2	68,5	50,7	57	3810
Namibya	2	45,3	83,3	71	6210
İndia	2	63,7	61,3	55	2670
Botswana	2	41,4	78,9	70	8170
Vanuatu	2	68,6	34	59	2890
Kamboçya	2	57,4	69,4	59	2060
Gana	2	57,8	73,8	46	2130
Myamnar	2	57,2	85,3	48	1027
Papoa yeni gine	2	57,4	64,6	41	2270

Bhutan	2	63	47		1969
Lao People's Dem. Rep.	2	54,3	66,4	59	1720
Comoros	2	60,6	56,2	45	1690
Swaziland	2	35,7	80,9	61	4550
Bangladeř	2	61,1	41,1	54	1700
Sudan	2	55,5	59,9	36	1820
Nepal	2	59,6	44	61	1370
Kamerun	2	46,8	67,9	56	2000
Pakistan	3	60,8	41,5	37	1940
Togo	3	49,9	59,6	67	1480
Kongo	3	48,3	82,8	48	980
Lesotho	3	36,3	81,4	65	2420
Uganda	3	45,7	68,9	71	1390
Zimbabwe	3	33,9	90	58	2400
Kenya	3	45,2	84,3	53	1020
Yemen	3	59,8	49	53	870
Madagaskar	3	53,4	67,3	45	740
Nijerya	3	51,6	66,8	45	860
Mauritanya	3	52,3	41,2	44	2220
Haiti	3	49,4	51,9	52	1610
Djibouti	3	45,8	65,5	24	1990
Gambia	3	53,9	37,8	45	1690
Eritrea	3	52,7	56,7	33	890
Senegal	3	52,7	39,3	38	1580
Timur Leste	3	49,3	58,6	75	
Rwanda	3	38,9	69,2	53	1270
Guinea	3	48,9	41	29	2100
Benin	3	50,7	39,8	52	1070
Tanzanya	3	43,5	77,1	31	580
Cote d'Ivoire	3	41,2	49,7	42	1520
Zambiya	3	32,7	79,9	45	840
Malawi	3	37,8	61,8	74	580
Angola	3	40,1	42	30	2130
Chad	3	44,7	45,8	35	1020
Congo Dem. Rep.	3	41,4	62,7	27	650
Central Africa	3	39,8	48,6	31	1170
Etyopya	3	45,5	41,5	34	780
Mozambik	3	38,5	46,5	41	1050
Guinea-Bissau	3	45,2	39,6	37	710
Burundi	3	40,8	50,4	33	630
Mali	3	48,5	19	26	930
Burkina Faso	3	45,8	12,8	22	1100
Niger	3	46	17,1	19	800
Sierra Leone	3	34,3	36	45	520

	İNT. KUL (%)	KADINLARIN PARLEMENT ODAKİ ORANI	KADINLARIN EKONOMİK AKTİVİTELERE KATILMA ORANININ ERKEKLERE GÖRE YÜZDESİ (15 YAŞ ÜSTÜ)	15 yaş altı nüfus/toplam nüfus	65 yaş üstü nüfus/ toplam nüfus
ÜLKELER	VAR6	VAR7	VAR8	VAR9	VAR10
Norveç	50,26	36,4	85	19,7	15,2
İsveç	57,31	45,3	89	17,9	17,4
Avusturalya	48,17	26,5	78	20,1	12,5
Kanada	51,28	23,6	83	18,4	12,8
Hollanda	50,63	35,1	67	18,4	13,8
Belçika	32,83	33,9	67	17,2	12,3
İzlanda	64,79	30,2	83	23	11,5
Amerika	55,14	14	82	21,6	12,2
Japonya	44,89	9,9	68	14,3	18,2
İrlanda	27,09	14,2	53	20,9	11,3
İsviçre	35,1	24,8	66	16,2	16,4
İngiltere	42,31	17,3	75	18,7	15,9
Finlandiya	50,89	37,5	87	17,8	15,3
Avusturya	40,94	30,6	66	16,2	15,8
Lüksemburg	37	16,7	58	19	13,4
Fransa	31,38	11,7	77	18,6	16,2
Danimarka	51,28	38	84	18,5	15
Yeni Zelanda	48,44	28,3	80	22,6	11,9
Almanya	41,19	31,4	70	15,2	17,1
İspanya	15,63	26,6	57	14,3	17
İtalya	35,24	10,3	59	14,1	18,7
İsrail	30,14	15	69	27,9	9,9
Çin	43,01		65	15,7	11
Yunanistan	15,47	8,7	59	14,7	18,2
Singapur	50,44	16	64	21,1	7,6
Portekiz	19,35	19,1	72	16,6	16
Slovenya	37,58	12,2	81	15	14,6
Kore	55,19	5,9	71	20,3	7,8
Barbodos	11,15	17,6	79	20	10
Kıbrıs Rum Kesimi	29,37	10,7	62	22,1	11,8
Malta	30,3	9,2	38	19,4	12,5
Çek Cumhuriyeti	25,63	15,7	83	15,7	13,9
Brunei Darussalam	10,23	0	63	30,6	2,9
Arjantin	11,2	31,3	48	27,3	9,9
Seychelles	14,52	29,4			
Estonya	32,77	18,8	82	16,7	15,8
Pollanda	23	20,7	80	17,9	12,5
Macaristan	15,76	9,8	72	16,4	14,8

Saint Kitts and Nevis	21,28	13,3			
Bahreyn	24,56	7,5	40	29,2	2,7
Letonya	14,44	10,6	80	19	14,5
Slovakya	16,04	19,3	84	18,4	11,5
Şili	23,75	10,1	50	27,8	7,5
Kuveyt	10,58	0	48	26,1	1,4
Kostarika	19,31	35,1	47	30,4	5,5
Uruguay	11,9	11,5	67	24,6	13,1
Katar	11,34	0	47	26,6	1,5
Hırvatistan	18,04	17,8	73	16,9	16,3
Birleşik Arap Emirlikleri	31,32	0	37	25,8	1,3
Litvanya	13,31	21	80	16,5	15,8
Bahama	19,23	26,8	84	29	5,5
Küba	1,07	36	66	20,3	10,2
Meksika	9,85	21,2	48	32,8	5
Trinidad Tobago	10,6	25,4	60	23,3	6,9
Antigua Barbuda	12,82	8,3			
Bulgaristan	8,08	26,3	86	14,8	16,3
Rusya	4,09	8	82	16,5	13,2
Libyan Arab Jamahiriya	2,25		34	31,3	3,7
Malezya	31,97	16,3	62	33,2	4,3
Makedonya	4,84	18,3	72	22	10,4
Panama	4,14	9,9	56	31,2	5,7
Belarus	8,16	18,4	82	17,1	14,2
Tonga	2,92	0		37,1	5,9
Mauritius	9,91	5,7	48	25,3	6,2
Albania	0,39	5,7	74	28,5	6,2
Bosna Hersek	2,62	12,3	60	17,6	10,8
Suriname	4,16	17,6	50	31	5,3
Venezuela	5,06	9,7	54	33	4,6
Romanya	10,15	9,3	76	17,1	13,9
Ukrayna	1,8	5,3	80	16,5	14,6
Saint Lucia	8,24	20,7		30,4	5,4
Brazilya	8,22	9,1	52	28,3	5,4
Kolombiya	4,62	10,8	61	32,1	4,9
Oman	7,09	0	26	37,2	2,1
Samoa	2,22	6,1		40,8	4,5
Tayland	7,76	9,6	85	25,6	5,8
Sudi Arabistan	6,46	0	29	39,1	2,7
Kazakistan	1,57	8,6	82	26	7,5
Jameika	22,84	13,6	86	30,8	7,1
Lübnan	11,71	2,3	39	29,6	6,2
Fiji	6,1	6	48	32,7	3,7
Ermenistan	1,58	4,6	88	21,4	9,2
Filipinler	4,4	17,2	61	36,6	3,7
Maldivler	5,34	6	80	43,1	3,2
Peru	9,35	18,3	44	33,6	5
Türkmenistan	0,17	26	82	34,6	4,5

Saint Vincent	5,98	22,7		31,1	6,7
Türkiye	7,28	4,4	62	30,7	5,7
Paraguay	1,73	8,8	44	38,8	3,6
Jordan	5,77	7,9	36	38	3
Azerbaycan	3,69	10,5	76	30,1	6,1
Tunus	5,17	11,5	48	28,5	5,9
Grenada	14,15	28,6			
China	4,6	20,2	86	23,7	7,1
Dominica	16,03	18,8			
Srilanka	1,06	4,4	56	25	6,9
Georgia	1,49	7,2	78	19,2	13,8
Dominican Republic	3,64	15,4	48	32,5	4,6
Belize	10,89	9,3	32	37,9	4,4
Ekvator	4,16	16	40	33,1	5
İran	4,85	4,1	38	32,6	4,5
Occupied Palestinian Territories	3,04		14	46,1	3,4
El Salvador	4,65	10,7	56	35,1	5,3
Guyana	14,22	22	50	30	5,1
Cape Verde	3,64	11,1	53	39,9	4,4
Syrian	1,29	12	38	38,3	3
Özbekistan	1,09	7,2	85	34,5	4,9
Algeria	1,6		41	33,5	4,2
Equatorial Guinea	0,36	5	52	43,6	3,9
Kırgızistan	2,98	6,7	85	32,6	6,3
Endonezya	3,77	8	68	29,9	5,1
Vietnam	1,85	27,3	91	31,7	5,4
Moldova	3,41	12,9	84	21,2	9,7
Bolivya	3,24	17,8	58	39	4,4
Honduras	2,52	5,5	48	40,7	3,6
Tacikistan	0,05	12,4	81	37,4	4,8
Mongolia	2,06	10,5	88	33,2	3,8
Nicaragua	1,68	20,7	57	41,9	3,1
Kuzey Afrika	6,82	27,9	59	33,2	3,9
Mısır	2,82	3,6	45	35,2	4,6
Guatemala	3,33	8,2	43	43	3,6
Gabon	1,92	11	77	41	4,6
Sao Tome and Principe	7,28	9,1		40,2	4,5
Solomon adaları	0,5	0	92	42,9	2,6
Morocco	2,36		53	31,8	4,4
Namibya	2,67	21,4	67	43,2	3,7
İndia	1,59	9,3	50	33,3	5,1
Botswana	2,97	17	77	39,8	2,7
Vanuatu	3,46	1,9		40,6	3,4
Kamboçya	0,22	10,9	97	41,9	2,9
Gana	0,78	9	98	40,1	3,3
Myanmar	0,05	0	75	32,3	4,6
Papua Yeni Gine	1,37	0,9	79	41,2	2,4
Bhutan	1,45	9,3	65	41,8	4,3

Lao People's Dem. Rep.	0,27	22,9	85	42	3,5
Comoros	0,42	0	73	42,3	2,4
Swaziland	1,94	16,8	52	43,7	3,3
Bangladeş	0,15	2	76	38,3	3,2
Sudan	0,26	9,7	41	39,7	3,5
Nepal	0,34		67	40,2	3,7
Kamerun	0,38	8,9	59	42,4	3,7
Pakistan	1,03	20,8	44	41,5	3,7
Togo	4,1	7,4	62	43,9	3,1
Kongo	0,15	10,6	71	46,8	2,9
Lesotho	0,97	17	56	39,9	4,7
Uganda	0,4	24,7	88	50,1	2,6
Zimbabve	4,3	10	78	43,1	3,4
Kenya	1,25	7,1	85	42,1	2,9
Yemen	0,51	0,3	37	48,7	2,3
Madagaskar	0,35	6,4	78	44,6	3
Nijerya	0,35	5,8	56	44,6	3,1
Mauritanya	0,37	4,4	74	43,2	3,3
Haiti	0,96	9,1	70	39,1	3,9
Djibouti	0,69	10,8		43	3,2
Gambia	1,88	13,2	78	40,9	3,5
Eritrea	0,23	22	87	45,5	2,1
Senegal	1,04	19,2	72	43,5	2,4
Timur Leste		26,1	86	39,3	2,8
Rwanda	0,31	45	88	45,2	2,5
Guinea	0,46	19,3	89	44	2,9
Benin	0,74	7,2	90	45,6	2,7
Tanzanya	0,23	21,4	93	45,3	2,3
Cote d'Ivoire	0,55	8,5	51	41,8	3,2
Zambiya	0,48	12	74	46,5	3
Malawi	0,26	9,3	90	46,2	3,5
Angola	0,29	15,5	82	47,5	2,7
Chad	0,19	5,8	77	46	3,1
Congo Dem. Rep.	0,09	7,4	72	46,8	2,6
Central Africa	0,13	0	78	43,1	4
Etyopya	0,07	7,8	67	45,7	2,9
Mozambik	0,27	30	92	44	3,2
Guinea-Bissau	0,4	0	63	47,1	3,1
Burundi	0,12	18,5	89	46,9	2,9
Mali	0,24	10,2	79	49,2	2,4
Burkina Faso	0,21	11,7	85	48,9	2,7
Niger	0,13	1,2	75	50	2
Sierra Leone	0,16	14,5	54	44,2	2,9