

PAPER DETAILS

TITLE: Sosyal Bilimler Arastirmalarinda Tasnif Edilmis Bagimli Degiskenler

AUTHORS: Kultuk Sümer

PAGES: 75-92

ORIGINAL PDF URL: <https://dergipark.org.tr/tr/download/article-file/100888>

SOSYAL BİLİMLER ARAŞTIRMALARINDA TASNİF EDİLMİŞ BAĞIMLI DEĞİŞKENLER (Discrete Dependent Variable)

Dr. Kutluk Kağan SÜMER*

ÖZET

Bu çalışma iki veya daha fazla nitel seçimin bağımlı değişken olduğu modeller üzerine bina edilmiştir. Bu modeller anket verilerinin analizi değerlendirilmesinde de kullanılmaktadır. Birçok anket davranışsal cevaplara dayalıdır.

Özellikle discrete choice modelleri ve bunların sosyal bilimler anketlerindeki modellerindeki spesifikasyon ve tahminleri üzerinde durmaya çalışacağız. Bu çalışmada ele alınacak kesikli bağımlı değişken modelleri şunlardır: Binary Logit, Binary Probit, Ordered Logit, Ordered Probit ve Multinomial Logit.

Anahtar Kelimeler: Kesikli Değişken, Kesikli Bağımlı Değişken, Regresyon, Logit, Probit, Ordered Logit, Ordered Probit, Multinomial Logit, Anket, Anket Değerlendirme

ABSTRACT

In this paper we construct models in which the dependent variable involves two or more qualitative choices. These models are valuable in the analysis of survey data. In most surveys the behavioral response are qualitative.

We discuss initially the specification and estimation of discrete choice models and discrete choice models applications in Social Science Surveys. The type of the discrete dependent models are Binary Logit, Binary Probit, Ordered Logit, Ordered Probit and Multinomial Logit.

Key Words: Discrete variable, Discrete dependent variable, regression, Logit, Probit, Ordered Logit, Ordered Probit, Multinomial Logit, Survey, Pool.

* İ.Ü. İktisat Fakültesi, Ekonometri Bölümü.

Sosyal bilimler araştırmalarının en yaygın uygulama biçimi bilindiği üzere anketlerdir. Anket çalışmalarında en yaygın olarak frekans tabloları ve çapraz tablolar kullanılır. Anketlere verilen cevaplar genellikle kesikli (discrete) değişkenler olduğu için alternatif istatistiki metotlar çok fazla kullanılamaz. Örneğin iki veya daha çok anket sorusu için regresyon kurulabilmesi pek mümkün değildir. İki cevap şıklı soruların bağımsız değişken (sebebe) olduğu yapılar için dummy (gölge) değişkenler kullanılabilir. Daha çok cevap şıklarının bağımsız değişken (sebebe) olduğu yapılar içinde dummy (gölge) değişkenlerin çeşitli formlarını kullanmak mümkündür.

Cevap şıklarının bağımlı değişken (sonuç) olması durumu ise şimdilik sosyal bilimler araştırmaları için çok yaygın kullanımları olmayan yeni teknikler olarak kalmaktadırlar.

Bu çalışmada kesikli (discrete) bağımlı değişken olarak kullanılacak veri yapılarını ve bunlar için kullanılacak alternatif modelleme ve çözüm tekniklerini incelemeye çalışacağız.

Aşağıdaki tablonun (Tablo-1) birinci sütununda kesikli (discrete) değişken tiplerini, ikinci sütununda ise çözüm metodu olarak kullanılacak regresyon modeli tiplerini tanımlamaya çalışacağız.¹

Tablo 1. Kesikli Değişken tipleri ve hesaplama için uygulanan regresyon metodu

Bağımlı Değişkenin Tipi	Regresyon Modelinin Tipi
Binary	Logistic Regresyon Probit Regresyon
Ordinal	Ordered Logistic Regresyon Ordered Probit Regresyon
Nominal	Multinomial Logistic Regresyon
Count	Poisson Regresyon Negatif Binomial Regresyon

Tablo-1 de sunulan birinci sütun kesikli sonuç değişkenleridir. Bu değişken tipleri mutlak suretle (-1,0,1,2) gibi tam sayı değerler alırlar. Küsuratlı değerler alması durumunda bu tür değişken yapılarına uymazlar.

¹ Michael R. Frone; (1997): Regression models for discrete and limited dependent variables; Pace University Research Methods Forum

Dört değişik kesikli değişken tipi vardır. Bunları şu şekilde özetlemek mümkündür.

1. Binary veya dichotomous sonuçlar sadece 2 değer alırlar (örneğin çalışıyor=0 çalışmıyor=1) Binary sonuçlar probit veya logit regresyonuyla analiz edilebilirler.

2. Ordinal sonuçlar 3 veya daha fazla değerın sıralanmasıdır. Ayrıca, sıra gösteren sonuçların mutlaka sonuç seçeneklerinin arasında eşit farkların olması zorunludur. (Örneğin 1= önemli değil, 2=biraz önemli, 3=çok önemli.) Ordinal sonuçlar ordered probit regresyonu veya ordered logistic regresyon ile hesaplanabilir.

3. Nominal sonuçlar Üç veya daha fazla sıra düzenli olmayan değerlerdir. (Örneğin 1= işçi, 2= işsiz, 3= part time işçi 4= memur ve 5=mevsimlik işçi gibi.) Nomial sonuçlar multinomial lojistik regresyonla analiz edilir.

4. Count sonuçlar: Tipik olarak sayma şeklindeki sonuçlardır. (Örneğin çocuk sayısı: 1,2,3 gibi) Poisson regresyonu veya negatif binomal regresyon metodlarıyla analiz edilebilirler.

Biz bu çalışmamızda ilk üç kesikli veri tipinin bağımlı değişken olarak kullanıldığı yapılardan ve çözüm tiplerinden bahsedeceğiz. Çözüm teknikleri için Minitab² istatistiki paket programının kullanacağız ve bu paket program yardımıyla çözüm metodlarını göstereceğiz. Çalışmamız için bir örnek anket hazırlayalım.

1. Cinsiyetiniz?

0- Erkek

1-Bayan

2. Yaşınız?

3. Medeni durumunuz?

1-Bekar

2- Evli

3- Dul

4- Boşanmış

² Minitab Statistical Software; Release 13.2 ; MINITAB INC

4. Tahsiliniz?

- 1-Tahsili yok
- 2- İlkokul
- 3- Ortaokul
- 4- Lise
- 5- Üniversite
- 6- Lisans Üstü

5. Günde Kaç Saat çalışıyorsunuz?

6. İş pozisyonunuz nedir?

- 1- İşsiz
- 2- Ücretli
- 3- İş veren
- 4- Yönetici
- 5- Kendi Hesabına Çalışan

7. Hane Geliriniz sizce nasıl?

- 1- Kıt kanaat geçiniyorum
- 2- Sadece Geçinmeye yetiyor
- 3- Rahat Yaşamaya yetiyor.
- 4- Rahat yaşamaya ve isteklerimi yerine getirmeye yetiyor
- 5- Tasarruf edebilecek kadar iyi
- 6- Çok iyi

8. İşinizi Seviyormusunuz?

- 1- Hiç sevmiyorum
- 2- Az seviyorum
- 3- Orta derecede seviyorum
- 4- Çok seviyorum

9. Sigara Kullanıyormusunuz?

0- Hayır

1- Evet

Anketimizdeki soruların değişken tipleri aşağıdaki tabloda verilmiştir. (Tablo-2)

Tablo 2. Kesikli Değişken tipleri ve hesaplama için uygulanan regresyon metodu

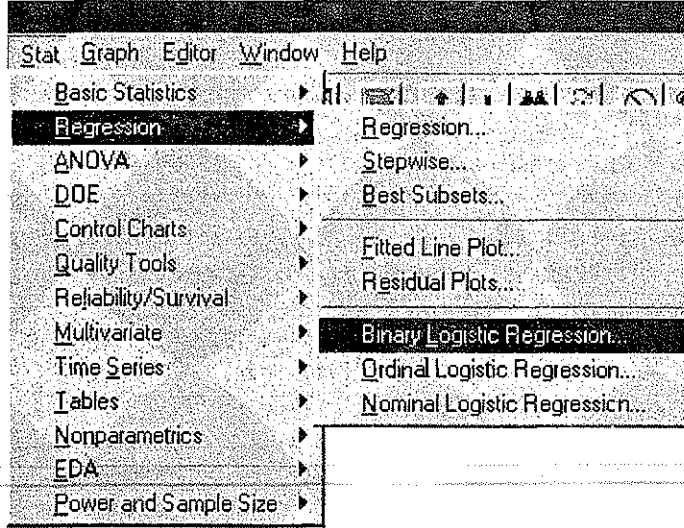
SORU	DEĞİŞKEN TİPİ
1. Cinsiyetiniz?	Binary(ikili)
2. Yaşınız?	Count(Sayaç)
3. Medeni durumunuz?	Nomial
4. Tahsiliniz?	Ordered (sıralı)
5. Günde Kaç Saat çalışıyorsunuz?	Ordered(Sıralı)
6. İş pozisyonunuz nedir?	Nomial
7. Hane Geliriniz sizce nasıl?	Ordered(Sıralı)
8. İşinizi Seviyormusunuz?	Ordered(Sıralı)
9. Sigara Kuilamyormusunuz?	Binary(ikili)

Söz konusu anketi 40 deneye uygulayalım. Aldığımız sonuçlara göre aşağıdaki tabloda üç regresyon metodunun kullanılacağı bağımlı ve bağımsız değişkenler görülmektedir. (Tablo-3)

Tablo 3. Kullanılacak regresyon metodu, bağımlı ve bağımsız değişkenler

Regresyon Metodu	Bağımlı Değişken	Bağımsız Değişkenler
Logit ve Probit	Sigara alışkanlığı	Cinsiyet Yaş Medeni Durum Tahsil İşteki durum Gelir
Ordered Logit ve Ordered Probit	Gelir	Cinsiyet Yaş Medeni Durum Tahsil Çalışılan saat İşteki durum
Multinomial Logit	İşteki Durum	Cinsiyet Tahsil Gelir

Daha öncede belirttiğimiz gibi söz konusu anketi değerlendirmek için Minitab paket programını kullanacağız. İlk evvela Binary Logistic Regression (Logit) e bakalım.



Şekil 1

Minitab programında “Stat” menüsü altındaki “Regression” altında sırasıyla “Binary Logistic Regression” Logit ve Probit, “Ordinal Logistic Regression” Ordered Logistic ve Ordered Probit, “Nominal Logistic Regression” ise Multinomial Logistic çözümlerini yapar. (Şekil-1)

Bunlardan “Binary Logistic Regression” u seçecek olursak karşımıza değişken tanımlama ekranı (Şekil-2) çıkacaktır. Bunlardan “Response” a bağımlı değişkenimiz olan sigara alışkanlığı ve “Model” e sırasıyla bağımsız değişkenlerimiz olan cinsiyet, yaş medeni durum, tahsil, işteki pozisyon ve geliri koyup “OK” e bastığımızda hesaplamalar yapılacaktır.

(Tablo-4) de hesaplanan değerleri görmekteyiz. Bunlardan bizim için önemli olan “Logistic Regression Table” altında tanımlanan değişkenlere ait parametre değerleri ve bunların test istatistikleridir. “Coef” sütunu parametreleri (katsayıları) “SE Coef” sütunu parametrelerin standart hataları, “Z” sütunu hesaplanan test istatistiğini, “P” sütunu ise hesaplanan test istatistiğinin olasılık değerini vermektedir.

Binary Logistic Regression

☒ **Response:** sigara **Frequency:** (optional)

☐ **Success:** **Trial:**

☐ **Success:** **Failure:**

☐ **Failure:** **Trial:**

Model:

cinsiyet yas medeni tahsil pozisyon gelir

Factors (optional):

Şekil 2

Tablo 4

Binary Logistic Regression: sigara versus cinsiyet; yas; ...							
Link Function: Logit							
Response Information							
Variable	Value	Count					
sigara	1	23 (Event)					
	0	17					
	Total	40					
Logistic Regression Table							
Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI Lower Upper	
Constant	1,351	2,379	0,57	0,570			
cinsiyet	0,0283	0,8222	0,03	0,973	1,03	0,21	5,15
yas	-0,07515	0,03963	-1,90	0,058	0,93	0,86	1,00
medeni	0,8681	0,6281	1,38	0,167	2,38	0,70	8,16
tahsil	-0,2478	0,2665	-0,93	0,052	0,78	0,46	1,32
pozisyon	-0,0086	0,3178	-0,03	0,978	0,99	0,53	1,85
gelir	0,3119	0,3264	0,96	0,339	1,37	0,72	2,59
Log-Likelihood = -23,517							
Test that all slopes are zero: G = 7,514; DF = 6; P-Value = 0,027							
Goodness-of-Fit Tests							
Method	Chi-Square		DF	P			
Pearson	42,203		30	0,049			
Deviance	47,035		30	0,025			
Hosmer-Lemeshow	18,380		8	0,019			

Klasik Regresyon teorisinde hipotezler:

$$H_0 : \beta=0$$

$$H_1 : \beta \neq 0$$

Şeklinde kurulmaktadır. P-değerinin 0,05 den küçük değerleri için H_0 hipotezi reddedilecektir ki buda bize değişkenimizin modelde yeri olduğunu gösterir. Modele aldığımız değişkenlerden bazılarının “P” (olasılık) değerleri 0,05’in üzerindedir. Bu durumda bu değişkenler geçerli olmayan (açıklayıcı olmayan-sıfır etkili) değişkenlerdir. Bunları modelden çıkartmak gerekir.

Tablo 5

Binary Logistic Regression: sigara versus yas, tahsil							
Link Function: Logit							
Response Information							
Variable	Value	Count					
sigara	1	23 (Event)					
	0	17					
	Total	40					
Logistic Regression Table							
Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI	
						Lower	Upper
Constant	3.578	1.671	2.14	0.032			
yas	-0.04884	0.03078	-1.59	0.049	0.95	0.90	1.01
tahsil	-0.3350	0.2403	-1.39	0.043	0.72	0.45	1.15
Log-Likelihood = -24,842							
Test that all slopes are zero: G = 4.864; DF = 2; P-Value = 0.048							
Goodness-of-Fit Tests							
Method	Chi-Square		DF	P			
Pearson	30.489		29	0,039			
Deviance	34.637		29	0,021			
Hosmer-Lemeshow	3.970		7	0,037			

(Tablo-5)’de açıklayıcı olmayan değişkenler model dışında bırakıldığında yaş ve tahsilin geçerli değişkenler olduğunu görmekteyiz. Model:

$$\text{Sigara Kullanımı} = 3.578 - 0.048(\text{Yaş}) - 0.335(\text{Tahsil})$$

Şeklinde olacaktır. Burada $Z = 3.578 - 0.048(\text{Yaş}) - 0.335(\text{Tahsil})$ şeklindeki hesaplamadan hareketle verilere göre sigara alışkanlığı ile yaş ve tahsil arasındaki bağlantı (Denklemler-1) şeklindedir.

$$P(Y) = \frac{1}{1 + e^{-z}} \quad (\text{Denklem 1})^3$$

Bu durumda örnek vermek gerekirse yaşı 35 olan ve üniversite mezunu (5) bir kişinin sigara içme ihtimali $Z = 3.578 - 0.048(35) - 0.335(5) = 0.223$ den hareketle:

$$P(Y) = 1 / (1 + \text{Exp}(-0.223)) = 0.5553 \quad (\text{Denklem 2})$$

olacaktır. Bu değer $P(Y) < 0.50$ durumunda $Y = 0$ yani sigara içmeyen aksi durumda $Y = 1$ olarak değerlendirilecektir ki örnek olaydaki denek Sigara içiyor çıkmalıdır.

Bakmamız gereken diğer testler ise "Test that all slopes are zero:" satırının sonundaki "P-Value" olmalıdır. Yine "P-Value" 0.05 den küçük değer almamalıdır. "Goodness-of-Fit Test"leri sırasıyla Pearson Chi-Square, Deviance Chi-Square ve Hosmer-Lemeshow Chi-Square istatistikleridir ki bunlarında P-değerleri "P" sütununda görülmektedir. Uygulamada en önemlisi Pearson Chi-Square dir. Yine 0.05 den küçük değer almaları tercih edilir.

Buraya kadar Binary Logit hesaplamasını anlattık Binary Probit analizi yapabilmemiz için değişken tanımlama ekranımızda (Şekil-2) küçük bir değişiklik yapmamız gerekmektedir. Bu ekranın sağ alt köşesinde "Options.." şeklinde bir düğme vardır. Bu düğmeye dokunduğumuzda karşımıza Binary Logit analizine ait seçenekler gelecektir (Şekil-3). Burada (Şekil-3)de de görebileceğimiz şekilde Fonksiyonu "Logit" den "Normit/Probit" e geçirmemiz ve "OK"e basmamız gerekmektedir.

$$P(Y) = F(Z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^Z e^{-s^2/2} ds \quad (\text{Denklem 3})^4$$

Bu daha karmaşık formül olasılığın hesaplanmasını nispeten zorlaştırmaktadır. (Tablo-6) da Z değerlerine göre F(Z) yani P(Y) değerlerinin hesaplanmasını kolaylaştırmak için farklı Z değerlerinin tekabül ettiği F(Z) ler verilmiştir.

³ Prof. Dr. Kazım Özdamar (1999); Paket Programlar ile istatistiksel veri analizi-1; S:487; Kaan Kitabevi

⁴ R. Pindyck, D. Rubinfeld (1998); Econometric Models and Econometric Forecast; S.305; McGraw-Hill; 4th Ed.

Tablo 6⁵. Kesikli Değişken tipleri ve hesaplama için uygulanan regresyon metodu

Z	F(Z)	Z	F(Z)
-3.0	0.001	0.5	0.691
-2.5	0.006	1.0	0.841
-2.0	0.023	1.5	0.933
-1.5	0.067	2.0	0.977
-1.0	0.159	2.5	0.994
-0.5	0.309	3.0	0.999
0.0	0.500	3.5	0.999

Şekil 3

Burada biraz önceki örnekteki gibi Probit (Tablo-7) den hareket edecek olursak $Z = 2.1189 - 0.02930(35) - 0.1968(5) = 0.1094$ değerinin $P(Y)$ değeri yaklaşık olarak 0.5 ile 0.691 arasında değer almaktadır ki yine şahıs sigara içiyor çıkmalıdır.

⁵ R. Pindyck, D. Rubinfeld (1998): Econometric Models and Econometric Forecast; S.305; McGraw-Hill; 4th Ed.

Tablo 7

Binary Logistic Regression: sigara versus yas, tahsil				
Link Function: Normit				
Response Information				
Variable	Value	Count		
sigara	1	23	(Event)	
	0	17		
	Total	40		
Logistic Regression Table				
Predictor	Coef	SE Coef	Z	P
Constant	2.1189	0.9648	2.20	0.028
yas	-0.02930	0.01825	-1.61	0.049
tahsil	-0.19.68	0.1430	-1.38	0.043
Log-Likelihood = -24,900				
Test that all slopes are zero: G = 4.749; DF = 2; P-Value = 0.043				
Goodness-of-Fit Tests				
Method	Chi-Square	DF	P	
Pearson	30.489	29	0,039	
Deviance	34.751	29	0.021	
Hosmer-Lemeshow	3.967	7	0.078	

Şimdi Ordered Logit ve Ordered Probit modellerine değinelim. Daha önce de belirtildiği üzere Ordered (Sıralı) Bağımlı değişkende seçenekler bir sıra takip etmek zorundadır. Burada Gelir ile medeni hal, tahsil ve işteki pozisyon arasında bir regresyon kurulmaya çalışılacaktır. (Şekil-1) de görüldüğü üzere "Stat/Regression/Ordinal Logistic Regression" menüsü altından hesaplama yapılacaktır. Karşımıza gelecek değişken tanımlama ekranı (Şekil-2) ile aynıdır. Burada bağımlı değişkenimiz gelir bağımsız değişkenlerimiz ise medeni hal, tahsil ve işteki pozisyonudur.

Gelirin aldığı değerler (1,2,3,4,5,6) olduğuna göre son değer (6) hariç beş paralel doğru varsayımına göre⁶ Const(1,2,3,4,5) şeklinde beş sabit katsayı hesaplanmıştır. Bu katsayıları $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$ şeklinde gösterir ve olasılıkları hesaplayacağımız formüller aşağıda verilmiştir (Denklem-4). Burada $P(Y=6)$ değeri direk orijinden geçmektedir. Yani sabit parametre olmayan değer Hane gelirinin çok iyi olması durumunu ifade etmektedir.

⁶ Prof. Dr. Kazım Özdamar (1999); Paket Programlar ile istatistiksel veri analizi-1; S:503; Kaan Kitabevi

$$P(Y=1) = \frac{1}{1 + e^{-(\beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3)}}$$

$$P(Y=2) = \frac{1}{1 + e^{-(\alpha_1 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3)}} \quad (\text{Denklem 4})$$

$$\Downarrow$$

$$P(Y=6) = \frac{1}{1 + e^{-(\alpha_5 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3)}}$$

Tablo 8

Ordinal Logistic Regression : gelir versus medeni; tahsil; pozisyon							
Link Function: Logit							
Response Information							
Variable	Value	Count					
gelir	1	3					
	2	6					
	3	6					
	4	7					
	5	12					
	6	6					
	Total	40					
Logistic Regression Table							
Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI Lower Upper	
Const (1)	-5.568	1.578	-3.53	0.000			
Const (2)	-4.032	1.445	-2.79	0.005			
Const (3)	-2.883	1.372	-2.10	0.036			
Const (4)	-1.687	0.581	-2.92	0.020			
Const (5)	0.552	0.277	1.99	0.048			
medeni	0.9173	0.4012	2.29	0.022	2.50	1.14	5.49
tahsil	0.7047	0.2304	3.06	0.002	2.02	1.29	3.18
pozisyon	-0.8739	0.2551	-3.43	0.001	0.42	0.25	0.69
Log-Likelihood = -56.064							
Test that all slopes are zero: G = 25.007; DF = 3; P-Value = 0.000							
Goodness-of-Fit Tests							
Method	Chi-Square		DF	P			
Pearson	154.336		127	0.050			
Deviance	87.854		127	0.997			

Tablo-8 de ordered (sıralı) Logistic Regresyon denklemini görmekteyiz.

Yine bir örnek üzerinde çalışacak olursak medeni hali bekar olan(1); tahsili olmayan (1) bir iş veren (3) e ait hesap yapacak olursak:

$$Z1 = -5,568 + 0,9173 (1) + 0,7047 (1) - 0,8739 (3) = -6,5677 \rightarrow P(Y=1) = 0,001403$$

$$Z2 = -4,032 + 0,9173 (1) + 0,7047 (1) - 0,8739 (3) = -5,0317 \rightarrow P(Y=2) = 0,0064$$

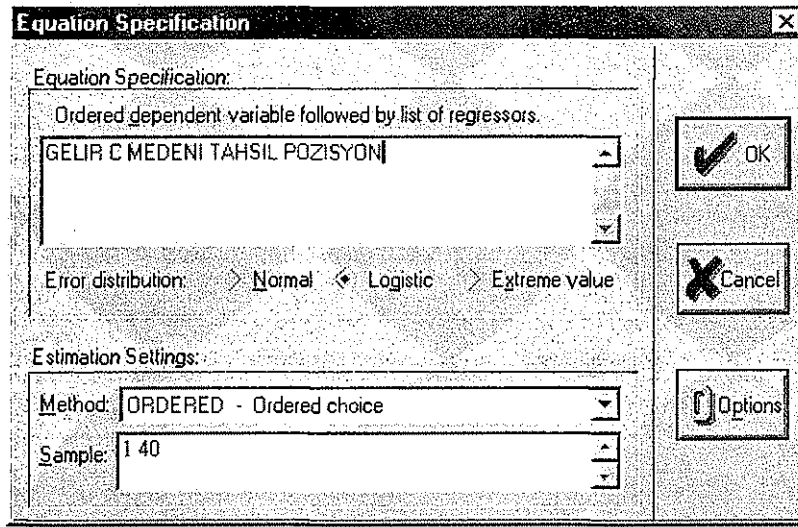
$$Z3 = -2,883 + 0,9173 (1) + 0,7047 (1) - 0,8739 (3) = -3,8827 \rightarrow P(Y=3) = 0,020$$

$$Z4 = -1,687 + 0,9173 (1) + 0,7047 (1) - 0,8739 (3) = -2,6867 \rightarrow P(Y=4) = 0,063$$

$$Z5 = -1,687 + 0,9173 (1) + 0,7047 (1) - 0,8739 (3) = -0,4477 \rightarrow P(Y=5) = 0,389$$

$$Z6 = 0,9173 (1) + 0,7047 (1) - 0,8739 (3) = 2,5672 \rightarrow P(Y=6) = 0,9287$$

yani bu durumda deneginizin Gelir beklentisi Çok iyi yani 6 olmalıdır.



Şekil 4

Ordinary Logistic ve Ordinary Probit çözümü için en ideal paket program E-Views 3.1' dir⁷. E-Views 3.1' de çok daha basit bir çalışma mantığı söz konusudur ki modelin kuruluşu (Şekil-4) de verilmiştir. Modelin hesaplamaları (Tablo-9) da görülmektedir. Yapılması gereken sadece $Y = -0.917 (1) - 0.704 (1) + 0.873 (3) = 0.999$ değerini bularak limit noktalarının hangisi arasında olduğuna bakmaktır. Söz konusu işlemin teorik açılımı (Denklemler-5) de verilmiştir. Bu formülde b matrisi parametreleri m ler ise limit noktalarını ifade etmektedir. Sistem 6 değer için 5 limit noktası hesaplamaktadır. Bu limit noktalarının en küçüğünden küçük olan ilk seçeneğe en büyüğünden büyük olan son seçeneğe

⁷ EViews; Version 3.1; © 1994-1998; Quantitative Micro Software

($0 < \mu_1 < \mu_2 < \dots < \mu_{J-1}$) tekabül etmektedir. Örneğimize dönecek olursak Y değerini 0.999 olarak hesaplamıştık (Tablo-9) daki $LIMIT_6:C(8) = 0.551669$ bu değer Y değerimiz olan 0.999 dan küçüktür yani yine gelir beklentisi Çok iyi yani 6 olmalıdır.

$$\begin{aligned} \text{Prob}(y = 1) &= \Phi(-\beta'x) \\ \text{Prob}(y = 2) &= \Phi(\mu_1 - \beta'x) - \Phi(-\beta'x) \\ \text{Prob}(y = 3) &= \Phi(\mu_2 - \beta'x) - \Phi(\mu_1 - \beta'x) \\ &\vdots \\ \text{Prob}(y = J) &= 1 - \Phi(\mu_{J-1} - \beta'x) \end{aligned}$$

(Denklem 5)⁸

Tablo 9

Dependent Variable: GELİR				
Method: ML - Ordered Logit				
Sample: 1 40				
Included observations: 40				
Number of ordered indicator values: 6				
Convergence achieved after 7 iterations				
Covariance matrix computed using second derivatives				
Coefficient	Std. Error	z-Statistic	Prob.	
MEDENİ	-0.917255	0.391552	-2.342614	0.0191
TAHSİL	-0.704746	0.234300	-3.007874	0.0026
POZİSYON	0.873881	0.247798	3.526588	0.0004
Limit Points				
LIMIT_2:C(4)	-5.567646	1.590489	-3.500587	0.0005
LIMIT_3:C(5)	-4.031584	1.493364	-2.699667	0.0069
LIMIT_4:C(6)	-2.883313	1.399314	-2.060519	0.0393
LIMIT_5:C(7)	-1.687418	1.309154	-1.288938	0.0197
LIMIT_6:C(8)	0.551669	1.313663	0.419947	0.047
Akaike info criterion	3.203193	Schwarz criterion	3.540969	
Log likelihood	-56.06386	Hannan-Quinn criter.	3.325322	
Restr. log likelihood	-68.56742	Avg. log likelihood	-1.401596	
LR statistic (3 df)	25.00712	LR index (Pseudo-R2)	0.182354	
Probability(LR stat)	1.54E-05			

Maalesef E-Views 3.1 de multinomial logit analizi yapmak mümkün değildir. (Tablo-10) da yine Minitab'a ait Ordered Probit analizine ait çıkış görülmektedir. Ordered Probit analizi yapabilmemiz için değişken tanımlama ekranımızda (Şekil-2) küçük bir değişiklik yapmamız gerekmektedir. Bu ekranın sağ

⁸ Greene William (2000): Econometric Analysis; S:876-877; Prentice Hall International Inc.

alt köşesinde "Options.." şeklinde bir düğme vardır. Bu düğmeye dokunduğumuzda karşımıza Ordered Logit analizine ait seçenekler gelecektir (Şekil-3). Burada (Şekil-3)de de görebileceğimiz şekilde Fonksiyonu Logit den Normit/Probit e geçirmemiz ve "OK"e basmamız gerekmektedir. Bu işlemi daha önce Binary Probit analizi yaparken de kullanmıştık. Ordered Probit analizi daha önce anlattığımız Ordered Logit analiziyle paraleldir. Sadece hesaplanan Z değerlerinin beklentileri olan P ler hesaplanırken daha önce Binary Probitte olduğu gibi (Tablo-6) dan faydalanırız.

Tablo 10

Ordinal Logistic Regression : gelir versus medeni ; tahsil ; pozisyon				
Link Function: Logit				
Response Information				
Variable	Value	Count		
gelir	1	3		
	2	6		
	3	6		
	4	7		
	5	12		
	6	6		
	Total	40		
Logistic Regression Table				
Predictor	Coef	SE Coef	Z	P
Const (1)	-2.8330	0.8464	-3.35	0.001
Const (2)	-1.9776	0.7909	-2.50	0.012
Const (3)	-1.3410	0.7679	-1.175	0.081
Const (4)	-0.6655	0.2662	-2.50	0.037
Const (5)	0.5815	0.2907	2.01	0.044
medeni	0.5077	0.2417	2.23	0.026
tahsil	0.3541	0.1250	2.83	0.005
pozisyon	-0.5248	0.1413	-3.72	0.000
Log-Likelihood = -56.740				
Test that all slopes are zero: G = 23.655; DF = 3; P-Value = 0.000				
Goodness-of-Fit Tests				
Method	Chi-Square	DF	P	
Pearson	151.390	127	0.049	
Deviance	89.206	127	0.0496	
Pairs	Number	Percent	Summary Measures	
Concordant	496	76,9%	Somers' D	0,57
Discordant	126	19,5%	Goodman-Kruskal Gamma	0,59
Ties	23	3,6%	Kendall's Tau-a	0,47
Total	645	100,0%		

SONUÇ

Sosyal Bilimler araştırmalarında kullanılan anket çalışmalarında frekans tabloları ve çapraz tablolar haricinde Discrete (Kesikli) Bağımlı Değişkenler metodu başlığı altında Binary, Ordinary ve Multinomial tipteki kesikli bağımsız değişken çeşitleri ile regresyon analizi yapmak mümkündür. Bu amaçla Binary Logit veya Probit, Ordered Logit veya Ordered Probit ve Multinomial Logit metodları kullanılabilir. Bu çalışma bir anket yardımıyla söz konusu metodların kullanımını sosyal bilimler alanında anket çalışması yapan araştırmacılara tarif etmeyi amaçlamaktadır.

KAYNAKÇA

Michael R. Frone. Regresion models for discrete and Limited dependent variables, Pace University Research Methods Forum 1997.

Minitab Statistical Software; Relase 13.2, MINITAB INC

Prof. Dr. Kazım Özdamar. Paket Programlar ile istatistiksel veri analizi-1, S:487, Kaan Kitabevi 1999.

R. Pindyck, D. Rubinfeld. Econometric Models and Econometric Forecast, S.305, McGraw-Hill; 4th Ed 1998.

EViews, Version 3.1, © 1994-1998, Quantitative Micro Software

Greene William. Econometric Analysis, S:876-877, Prentice Hall International Inc 2000.

Backhaus, Erichson, Plinke, Weiber. Multivariate Analysemethoden, S:113, Springer-Lehrbuch 2000.