Uludağ Üniversitesi
İktisadi ve İdari Bilimler Fakültesi Dergisi
Cilt XI, Sayı 1-2, Mart-Kasım 1990

# REGRESYON MODELLERİNDE KUKLA DEĞİŞKENLERİN ROLÜ

Mustafa SEVÜKTEKİN*

## ÖZET

*Bu makalenin amacı regresyon analizinde nitel açıklayıcı değişkenlerin rolünü incelemektir. Nitel değişkenlerin sık sık kukla değişkenler olarak takdim edildiğini görmekteyiz. Kukla değişkenleri kullanma tekniği geniş bir uyarlama içerisinde yapılabilir ve sonuçta regresyon modellerindeki nitel değişkenlerin dahil edilmesiyle tahmin yapabilme (düz bir doğru uydurabilme) gerçekleştirilebilir. Nitel değişkenlerin gösterimi için kullanılan teknikte eğer sözkonusu nitel değişkenin bağımlı değişken üzerinde bir etkisi olduğu düşünülüyorsa 1 değeri, herhangi bir etkisi yoksa 0 değeri verilir. Burada incelenen "varyans analizi" ve "kovaryans analizi" nitel açıklayıcı değişken içeren regresyon modelleri iyi bilinen istatistiki tekniklerdendir.*

## SUMMARY

*The purpose of this paper is to consider the role of qualitative explanatory variables in regression analysis. I will be shown that the introduction of qualitative variables, often called dummy variables. The technique of using dummy variables has een widely adopted, and the result is the straighforward inclusion of qualitative variables in regression models. This technique uses the value 1 for the presence of the qualitative attribute that is assumed to have an impact on the dependent variable and 0 for the absence of the given attribute. Included here are the well-know statistical techniques of "analysis of variance" and "analysis of covariance" both of which can be regarded as regression models with qualitative explanatory variables.*

---

* *Yard. Doç. Dr.; U.Ü. İ.İ.B.F. Ekonometri Bölümü, Ekonometri Anabilim Dalı Öğretim Üyesi.*

# 1. GİRİŞ

Ekonometrik araştırmalarda kullanılan değişkenlerin büyük çoğunluğu ölçülebilen, sayılabilen nicel vasıftaki değişkenlerden oluşmaktadır. Örneğin, tüketim harcamaları, kullanılabilir gelir, yatırım harcamaları ve fiyatlar parasal birimlerle; faiz oranları ve işsizlik oranları yüzde kavramlarla; herhangi bir malın talep edilen miktarı, arz edilen miktarı ve üretim miktarı ağırlık veya adet ölçü birimleriyle ifade edilmektedir. Ancak bazı ekonometrik araştırmalarda yer verilmesi düşünülen öyle değişkenler vardır ki, herhangi bir ölçü birimiyle ölçmemiz ve saymamız mümkün olmayabilir. Ekonometrik araştırmalarda nitel vasıftaki gerçeklerin (fenomenlerin) temsil edildiği değişkenlerin birer etkileyici faktör olarak modellerde sık sık yer aldığı gözlenmektedir. Ölçülemiyen veya sayılamıyan fakat etkisi olduğuna inanılan bu unsurun regresyon modellerinde yer alması düşünülüyorsa, bu nitel vasıftaki etkileyici faktörler, "kukla değişkenler" yardımıyla yaklaşık olarak ölçülmeye çalışılır[1].

Meslek, din, ırk, eğitim düzeyleri, cinsiyet gibi bazı faktörler belirli malların tüketimini etkileyen önemli unsurlardan bazılarıdır. Bu nitel unsurların (gerçeklerin) etkilerini, herhangi bir tüketim fonksiyonunda görmek istiyorsak, kukla değişkenler yardımıyla bunları kolaylıkla yaklaşık olarak ele alabiliriz. Kukla değişkenlerin hazırlanması ve kullanımı oldukça keyfi olmasına karşın, ilgili faktörün etkisini yansıtmasından ötürü geçerlidirler[2]. Kukla değişkenlerin yapısını incelemeye geçmeden önce, aşağıda verilen örneklere kısaca gözatmamızda yarar vardır[3]:

i. Yatay-kesit verilerinin kullanıldığı hanehalklarını ilgilendiren bir tüketim fonksiyonu analizinde, kentsel ve kırsal alanlarda ikamet edenlerin herhangi bir veri gelir düzeyinde tüketim kalıplarının farklı olacağı açıktır. Aynı zamanda yalnızca kentsel alanları dikkate aldığımızda, kentin belirli bölgeleri (veya semtleri) arasında da tüketim davranışının farklı olduğu söylenebilir. Hanehalkları açısından farklı bir tüketim davranışı, bireylerin evli veya bekâr olması halinde de beklenebilir.

ii. Zaman-dizisi verilerinin kullanıldığı bir başka tüketim fonksiyonu analizinde, tüketim yalnızca kullanılabilir gelire bağlı olmayabilir. Tüketim ayrıca kişilerin yapacakları tüketim harcamalarına da bağlıdır. Analizde dikkate alınan mevcut veri dönemi boyunca ülke bir savaş hali yaşamış olabilir. Savaş halinin sürdüğü dönemlerde ülkede muhtemelen bazı tüketim malları üzerine bir takım

---

1 Bakınız Sevüktekin (1988).

2 Daha fazla bilgi için bakınız Koutsoyiannis (1977:18).

3 Bu örnekler kukla değişkenlerin yapısını daha yakından tanımamıza neden olacaktır, benzer diğer örnekler için bakınız Stewart ve Wallis (1981:171-2); Kmenta (1971:409-10) ve Dutta (1975: 159-60).

kısıtlamalar getirilmiş olabilir. (Örneğin karne ile satış veya sınırlı miktarlarda tüketim gibi). Aynı gözlem döneminde hem savaş halinin, hem de barış halinin birlikte bulunması analizde bazı sorunlar yaratabilir. Örneğin, barış halinin sürdüğü dönemlerde, savaş halinin sürdüğü dönemlerden daha düşük bir kullanılabilir gelirle tüketim harcamalarında bulunulabilir.

iii. Diğer bir yatay-kesit analizinde bireylerin kazançları arasındaki farklılıkları incelemekte olduğumuzu düşünelim. İş-gücü piyasasında bireylerin eğitim düzeyleri, vasıfları, işgücüne katılma yaşları, cinsiyetleri gibi faktörler kazanılan gelirin farklılaşmasına neden olan faktörlerin başında gelirler.

iv. Bir zaman-dizisi çalışmasında siyasi olayların, ulusal ekonomi ile ilgili olarak bütüncül veriler üzerine etkileri araştırılmak istenebilir. Bir ulusal hükümetin izlediği politikalar partiler ve iktidarlar açısından farklılıklar gösterebilir ve dolayısıyla uygulanan farklı politikaların sonuçları da muhtemelen farklı olacaktır. Ulusal sınırlar içerisinde bazı doğal felaketler, politik krizler ve savaş ve barış hali gibi durumlar mutlaka ekonomik ve sosyal politikalar üzerine etkileri sözkonusu olacaktır.

Yukarıda verilen örneklerin sayısını arttırmak mümkündür. Örneklerde de açıkça görüldüğü gibi, ölçülebilen, sayılabilen faktörlerin yanında, etkilerinin olduğu düşünülen fakat ölçülemiyen, sayılamıyan faktörler mevcuttur. Din, eğitim, aile reisinin cinsiyeti, ırk, yerleşim yerleri, aile reisinin evlilik durumu, ailenin büyüklüğü, etnik bilgiler (gelenekler, örfler ve töreler gibi), politik üyelikler, sendika veya dernek üyeliği, savaş hali, barış hali, ulusal felaketler (seller, depremler, kuraklık gibi) ve daha sayabileceğimiz birçok faktör ekonometrik araştırma yapanlar tarafından modellere ilave edip incelemek isteyebilirler. Bu amaçla bu faktörleri temsilen kukla değişken kullanarak modellere katabilirler.

## 2. KUKLA DEĞİŞKENLERİN YAPISI

Kukla değişkenler genellikle iki mümkün durumdan birisini değer olarak alırlar. Değerlerden birisi, bir nitel durumun mevcut olduğunu belirtir, diğeri ise nitel bir durumun olmadığı hali gösterir. Geleneksel olarak kukla değişkenler "sıfır" ve "bir" değerlerini alırlar. "Bir" değeri bir nitel durumun vukuu bulduğunu, gerçekleştiğini gösterir, ya da o durumun bir karakteristik görünüşü yansıtır. "Sıfır" değeri ise, bir nitel durumun vukuu bulmadığını, gerçekleşmediğini gösterir, ya da karakteristik bir ifade edişin olmadığını vurgular[4].

4 Kukla değişkenlerin sayısallaştırılmasına dair verilen 0 ve 1 değerleri tamamen keyfi bir tanımlamadır, bu konuda daha fazla bilgi için bakınız Sevüktekin (1988); Intriligator (1978:58); Common (1976:371) ve Ferber ve Verdoorn (1962:369).

Herhangi bir sorunun nitel vasfını 1 ve 0 gibi iki değer ile nitelendirmek, yalnızca verilen niteliğin varlığını ve yokluğunu belirtmek için yararlandığımız ve seçtiğimiz herhangi iki değerden başka birşey değildir. Dolayısıyle ikili bir değere sahip olan böyle bir nitel yapıdaki değişkene "ikili-binary" veya "kukla" değişken denilmektedir[5]. Nitel türdeki değişkenlere ayrıca "gösterge", "kategorik", "gölge" ve "dikotomi" değişken adları da verilmektedir[6]. Bu çalışmamızda daha ziyade kukla değişken kavramı ağırlıkla kullanılacaktır.

Nitel yapıdaki kukla değişkenin yapısını basit birkaç örnek vasıtasıyla daha iyi tanımaya çalışalım. Örneğin bir birey eğer erkek ise 1, kadın ise 0; yüksek okul mezunu ise 1, değil ise 0; evli ise 1, değil ise 0; müslüman ise 1, değil ise 0; zenci ise 1, değil ise 0; kentsel bir alanda ikâmet ediyorsa 1, kırsal alanda ise 0 gibi değerlendirmeler yapılabilir. Görüldüğü gibi kukla değişkeni tanımaya çalışırken verdiğimiz bu örneklerde değerlendirmeler tamamıyla farazi olarak yapılmıştır. Dolayısıyle değerlendirmede daha çok modelde hangi nitel vasfın etkili olduğuna bakılarak yapılır. Örneğin cinsiyet açısından etkin veya etkileyici vasıf "erkek" ise 1 değeri erkek vasfına verilir, "kadın" ise kadın vasfına verilir. Yüksek okul mezunu olma vasfı modelde etkileyici bir unsur ise 1 değeri bu vasfa verilir. Benzer şekilde diğer bütün nitel faktörlerin değerlendirilmesi yapılabilir.

## 3. VARYANS ANALİZİ MODELLERİ (AOV)

Basit bir regresyon modelinde açıklayıcı değişken olarak kukla değişkenlere yer vermek mümkündür. Kukla değişkenlerin bu kullanımı nicel değişkenlerde olduğu gibi kolaylıkla regresyon denklemlerine adapte edilebilir. Regresyon denklemlerinde açıklayıcı değişken olarak yalnızca bir veya daha fazla kukla değişken yer alabilir. Yalnızca kukla değişkenlerin yer aldığı regresyon modellerine "Varyans Analizi Modelleri" (AOV) denilmektedir[7].

Aşağıda varyans analizi modellerine dair bir örnek verilmektedir:

$$Y_i = \alpha_0 + \alpha_1 D_i + u_i \qquad (3.1)$$

burada

$Y_i$ = Özel bir eğitim kurumunda çalışan bir öğretmenin yıllık geliri,

$D_i$ = 1 eğer öğretmen bir doktora derecesine sahip ise,

= 0 diğer durumlar (yani öğretmenin bir doktora derecesi yoktur),

5 Bakınız Kmenta (1971:409).

6 Birçok isimlendirme değişik çalışmalarda karşımıza çıkmaktadır. Ancak burada verilenler kukla değişkenler olarak kısaca adlandıracağımız nitel türdeki değişkenler için en çok kullanılanlarıdır. Bakınız Gujarati (1979:287) ve Common (1976:371).

7 Varyans modelleri için bakınız Gujarati (1979:288-9); Kmenta (1971:410) ve Dutta (1975:159-60)

bu modelde dikkat edilirse nicel türde bir açıklayıcı değişken yerine bir kukla değişken kullanılmıştır. Modelde her ne kadar özel bir eğitim kurumunda çalışan öğretmenlerin maaşlarındaki farklılığı yalnızca eğitim düzeyinin yansıtması karşısında, bu farklılığı verecek, örneğin öğretmenlerin bulundukları kadro dereceleri, branşları ve deneyimli olup olmadıkları gibi faktörlerin sabit tutulduğu gözden kaçmamalıdır. Modelde yer alan kalıntı terim $u_i$'nin klasik doğrusal regresyon modeli için ileri sürülen varsayımları sağladığını kabul ettiğimizde, Model (3.1)'deki $Y_i$'nin beklenen değerleri $D_i$'nin iki farklı değerine karşılık gelecektir[8]. Bu durumlar aşağıda verilmektedir.

$$E(Y_i | D_i = 0) = \alpha_0 \tag{3.2}$$

$$E(Y_i | D_i = 1) = \alpha_0 + \alpha_1 \tag{3.3}$$

Denklem (3.2) özel eğitim kurumunda çalışan bir doktorasız öğretmenin, Denklem (3.3) ise bir doktoralı öğretmenin ortalama yıllık gelirlerini (maaşlarını) vermektedir. Bu sonuçlar ayrıca modelde yer alan parametrelerin değerlerinden de çıkarılabilir. Kesme terimi $\alpha_0$ özel eğitim kurumunda çalışan doktorasız bir öğretmenin ortalama yıllık gelirini, $\alpha_0 + \alpha_1$ ise doktoralı bir öğretmenin ortalama yıllık gelirini vermektedir. $\alpha_1$ bir doktoralı öğretmenin, bir doktorasız öğretmenin yıllık ortalama gelirinden ne kadar farklı olduğunu belirtir. Bir başka ifadeyle doktoralı öğretmenlerin $\sigma^2$ varyansla ve $\mu_1$ ortalama ile doktorasız öğretmenlerin de yine $\sigma^2$ varyans ve $\mu_0$ ortalama ile normal dağıldığını kabul ettiğimizde kısaca şu sonuca ulaşabiliriz[9].

$$\alpha_0 = \mu_0 \tag{3.4}$$

$$\alpha_0 + \alpha_1 = \mu_1 \tag{3.5}$$

$$\alpha_1 = \mu_1 - \mu_0 \tag{3.6}$$

Bu ortalamalar anakütle regresyon doğrusu (3.1)'in kesmesini ölçer, eğim ise doktoralı öğretmen ile doktorasız öğretmenin yıllık ortalama gelir farklılığını verir. Kurulacak bir hipotez testinde $\alpha_1$'nin sıfıra eşit bulunması halinde doktoralı ve doktorasız öğretmenin yıllık ortalama gelirleri arasında bir fark olmadığı imajı ortaya çıkar.

Denklem (3.1)'in katsayıları en küçük kareler yöntemi ile tahmin edilebilir. Klasik doğrusal regresyon varsayımları altında elde edilen tahminler arzu edilen özellikleri sağlayacaktır. Bunun için en küçük kareler tahmincilerini şöyle verebiliriz[10].

---

8 Bakınız Gujarati (1979:288-9) ve Kmenta (1971:410).
9 Kmenta (1971:410-1).
10 Bakınız İşyar (1989:156-8) ve Kılıçbay (1980:266-7).

$$\hat{\alpha}_1 = \frac{\Sigma (D_i - \overline{D})(Y_i - \overline{Y})}{\Sigma (D_i - \overline{D})^2} \tag{3.7}$$

$$\hat{\alpha}_0 = \overline{Y} - \hat{\alpha}_1 \overline{D} \tag{3.8}$$

Modelimizi bir adım daha ileriye götürerek bir genelleştirme yapalım. n sayıda bir örneklem alarak aşağıdaki sonuçları çıkarmaya çalışalım:

$n_1$: Örneklemdeki doktoralı öğretmenlerin sayısı

$n_0$: Örneklemdeki doktorasız öğretmenlerin sayısı

$\overline{Y}_1$: Doktoralı öğretmenlerin örneklem ortalama yıllık geliri

$\overline{Y}_0$: Doktorasız öğretmenlerin örneklem ortalama yıllık geliri

Buna göre,

$$\sum_{i=1}^{n} D_i = n_1$$

$$\sum_{i=1}^{n} D_i^2 = n_1$$

$$\sum_{i=1}^{n} Y_i = n_1\overline{Y}_1 + n_0\overline{Y}_0$$

$$\sum_{i=1}^{n} D_iY_i = n_1\overline{Y}_1$$

elde edilebilir. Bu sonuçları tahmin denklemleri için kullanmak istersek:

$$\sum_{i=1}^{n} (D_i - \overline{D})(Y_i - \overline{Y}) = \Sigma D_iY_i - \frac{1}{n}(\Sigma D_i)(\Sigma Y_i)$$

$$= n_1\overline{Y}_1 - \frac{n_1}{n}(n_1\overline{Y}_1 + n_0\overline{Y}_0) = \frac{n_0 n_1}{n}(\overline{Y}_1 - \overline{Y}_0)$$

ve

$$\sum_{i=1}^{n} (D_i - \overline{D})^2 = \Sigma D^2_i - \frac{1}{n}(\Sigma D_i)^2 = n_1 - \frac{n^2_1}{n} = \frac{n_0 n_1}{n}$$

sonuçlarını elde ederiz. Bu sonuçlar veri iken tahminciler şu şekilde olacaktır:

$$\hat{\alpha}_1 = \frac{(n_0n_1/n)\,(\overline{Y}_1 - \overline{Y}_0)}{n_0n_1/n} = \overline{Y}_1 - \overline{Y}_0 \qquad (3.9)$$

$$\hat{\alpha}_0 = \frac{1}{n}(n_1\overline{Y}_1 + n_0\overline{Y}_0) - (\overline{Y}_1 - \overline{Y}_0)\frac{n_1}{n} = \overline{Y}_0 \qquad (3.10)$$

Görüldüğü gibi regresyon denkleminin en küçük kareler eğim tahmincisi doktoralı bir öğretmenin örneklem ortalaması ile doktorasız bir öğretmenin örneklem ortalaması arasındaki farka eşdeğerdir. Regresyon denkleminin en küçük kareler kesme tahmincisi doktorasız bir öğretmenin örneklem ortalaması yıllık gelirine (maaşına) denktir.

Yukarıda açıklanmaya çalışılan regresyon modelinde, açıklayıcı değişken olarak yer alan kukla değişken ikili bir yapı özelliği göstermektedir, yani iki mümkün durumu ancak incelemektedir. Fakat buradaki kukla değişkenin yapısı gereği ikili değil de (yani yalnızca doktoralı ve doktorasız) daha karmaşık olduğunu düşünelim; örneğin lisans derecesi, yüksek lisans derecesi ve doktora derecesi gibi üçlü bir durumu ele alalım. Bu durumda bir önceki denklemi kullanmamız mümkün olamıyacaktır. Varsayalım ki, öğretmenlerin yıllık gelirleri (maaşları) onların lisans, yüksek lisans ve doktora derecelerine sahip olmasına bakılmaksızın $\sigma^2$ varyansla ve $\mu$ ortalama ile normal bir dağılım göstersin. Lisans mezunu bir öğretmen için başlangıç maaş ortalaması $\mu_A$'ya, yüksek lisans dereceli bir öğretmenin başlangıç maaş ortalaması $\mu_B$'ye ve doktora dereceli bir öğretmenin başlangıç maaş ortalaması da $\mu_C$'ye eşit olsun. Bu durumu açıklayan en uygun regresyon denklemi şöyle verilebilir[11]:

$$Y_i = \alpha_0 + \alpha_1 D_{i2} + \alpha_2 D_{i3} + u_i \qquad (3.11)$$

burada

$Y_i$ = i'ninci öğretmenin bir yıllık geliri

$D_{i2}$ = 1 öğretmenin en yüksek derecesi doktora ise

= 0 Diğer durumlar

$D_{i3}$ = 1 Öğretmenin en yüksek derecesi yüksek lisans ise

= 0 Diğer durumlar.

Dikkat edilirse $D_{i2} = 1$ olduğu zaman, $D_{i3}$ sıfıra eşit olmaktadır ve tersi durum da varittir. $Y_i$'nin beklenen değerleri bu kez $D_i$'nin üç farklı değerine karşılık gelmektedir, bunlar:

---

11 Bakınız Kmenta (1971:412).

$$E(Y_i | D_{i2} = 1, D_{i3} = 0) = \alpha_0 + \alpha_1 \quad (3.12)$$

$$E(Y_i | D_{i2} = 0, D_{i3} = 1) = \alpha_0 + \alpha_2 \quad (3.13)$$

$$E(Y_i | D_{i2} = 0, D_{i3} = 0) = \alpha_0 \quad (3.14)$$

Aynı zamanda şu sonuçları da buradan kolaylıkla çıkarabiliriz:

$$\alpha_0 = \mu_A \quad (3.15)$$

$$\alpha_1 = \mu_C - \mu_A \quad (3.16)$$

$$\alpha_2 = \mu_B - \mu_A \quad (3.17)$$

Bu sonuçlara dikkat edilirse (3.4) ile (3.5) ve (3.6)'da çıkardığımız sonuçlara benzemektedir.

Şu ana kadar yukarıda sunulan aşamalarda kukla değişkenler ikili yapıda, yani 0 ve 1 değerlerini aldığı varsayımı altında incelendi. Bu açıdan modelin üç farklı ayırıma tabi tutulması gereği ortaya çıktı. Oysa şimdiye değin vurgulamadığımız bir yapıda açıklayıcı değişken olarak modelde yer verdiğimiz kukla değişkeni tanımlamamız sözkonusu olabilir. Öğretmenlerin derecelerini yansıtan kukla değişkeni üçlü bir yapıda tanımlayabiliriz. Örneğin lisans derecesine sahip olana 0, yüksek lisans derecesine sahip olana 1 ve doktora derecesine sahip olana da 2 değerini verebiliriz. Regresyon modelini bu durumda şöyle tanımlayabiliriz[12]:

$$Y_i = \alpha_0 + \alpha_1 W_i + u_i \quad (3.18)$$

burada $W_i$, 0, 1 ve 2 değerlerini alan üçlü yapıdaki bir kukla değişkendir. $Y_i$'nin beklenen değeri $W_i$'nin üç farklı değerine karşılık gelecektir:

$$E(Y_i | W_i = 0) = \alpha_0 \quad (3.19)$$

$$E(Y_i | W_i = 1) = \alpha_0 + \alpha_1 \quad (3.20)$$

$$E(Y_i | W_i = 2) = \alpha_0 + 2\alpha_1 \quad (3.21)$$

buradan şu sonuçları kolaylıkla çıkarabiliriz. Bir lisans derecesine sahip öğretmen ile bir yüksek lisans derecesine sahip olan öğretmenin ortalama maaşları arasındaki fark

$$(\alpha_0 + \alpha_1) - \alpha_0 = \alpha_1 \quad (3.22)$$

ile verilmektedir. Bir doktora derecesine sahip olan öğretmen ile bir yüksek lisans derecesine sahip olan öğretmenin ortalama maaşları arasındaki fark ise

---

12 Bakınız Kmenta (1971:412-3).

$$(\alpha_0 + 2\alpha_1) - (\alpha_0 + \alpha_1) = \alpha_1 \tag{3.23}$$

ile verilmekte ve nihayet bir doktora dercesine sahip olan öğretmen ile bir lisans derecesine sahip olan öğretmenin ortalama maaşları arasındaki fark

$$(\alpha_0 + 2\alpha_1) - \alpha_0 = 2\alpha_1 \tag{3.24}$$

ile verilmektedir. Yani 0, 1 ve 2 değierlerini alan bir değişkenin kullanılmasıyla gerçekte doktoralı bir öğretmen ile yüksek lisanslı bir öğretmenin maaşı arasındaki fark bir yüksek lisanslı öğretmen ile lisanslı bir öğretmenin maaşı arasındaki farka eşittir. Ancak bizim burada önsel bir bilgimiz olmadıkça böyle bir varsayımın doğruluğunu kanıtlayamayız.

Dikkat edilirse iki tane ikili değişken yerine (regresyon denkleminde sabit terimi düşürmeksizin) üçlü yapıdaki bir değişkene yer veremeyiz. Bunun için regresyon modelini aşağıdaki gibi düzenlemeliyiz:

$$Y_i = \alpha_0 + \alpha_1 D_{i2} + \alpha_2 D_{i3} + \alpha_3 D_{i4} + u_i \tag{3.25}$$

burada $D_{i2}$ ve $D_{i3}$ Denklem (3.11) de tanımlandığı gibidir.

$D_{i4}$ = 1 Öğretmenin en yüksek derecesi lisans ise,

= 0 diğer durumlar

ve $\alpha_0$, $\alpha_1$, $\alpha_2$ ve $\alpha_3$ için çözüm belirsiz olacaktır. Zira sonuç

$$D_{i4} = 1 - D_{i2} - D_{i3} \tag{3.26}$$

bu durumda en küçük kareler normal denklemleri tanımsızdır, yani (X'X) matrisi tekil bir matristir, katsayılar için bir çözüm yapılalmaz[13].

## 4. KOVARYANS ANALİZİ MODELLERİ (ACOV)

Yukarıda açıklamaya çalıştığımız varyans analizi modelleri (AOV), sosyal bilimlerin daha çok, sosyoloji, psikoloji, eğitim ve piyasa araştırmaları gibi değişik alanlarda başvurulan modellerdir, ekonomide bu tür modellerin kullanımı pek yaygın değildir. Ekonomide daha ziyade nicel türdeki değişkenlerin yanında açıklayıcı bir değişken olarak nitel (kukla) değişkenlere yer veren regresyon modelleri kullanılmaktadır. Nicel ve nitel (kukla) değişkenlerin birlikte yer aldığı modellere "Kovaryans Analizi Modelleri" (ACOV) denilmektedir[14].

---

13 Böyle bir sorundan kurtulmak için başvurulacak değişik yöntemler sözkonusu olabilir. Bu konuda daha detaylı bilgi aşağıda beşinci başlıkta "kukla değişken tuzağı" sorunu adı altında incelenecektir.

14 Kovaryans analizi modelleri için bakınız Gujarati (1979:289); Kmenta (1971:419-20); Kelejian ve Oates (1981:169-70); Common (1976:372-3); Maddala (1979:135-6) ve Draper ve Smith (1966: 136-41).

Yukarıda ele aldığımız (3.1)'deki modele yıllık eğitim harcamalarını yansıtan bir nicel değişken ilave ederek bir kovaryans analizi modeli (ACOV) geliştirebiliriz:

$$Y_i = \alpha_0 + \alpha_i D_i + \beta X_i + u_i \tag{4.1}$$

burada

$Y_i$ = Özel eğitim kurumunda çalışan bir öğretmenin yıllık maaşı,

$X_i$ = Yıllık eğitim harcamaları,

$D_i$ = 1 öğretmen eğer doktora derecesine sahip ise,

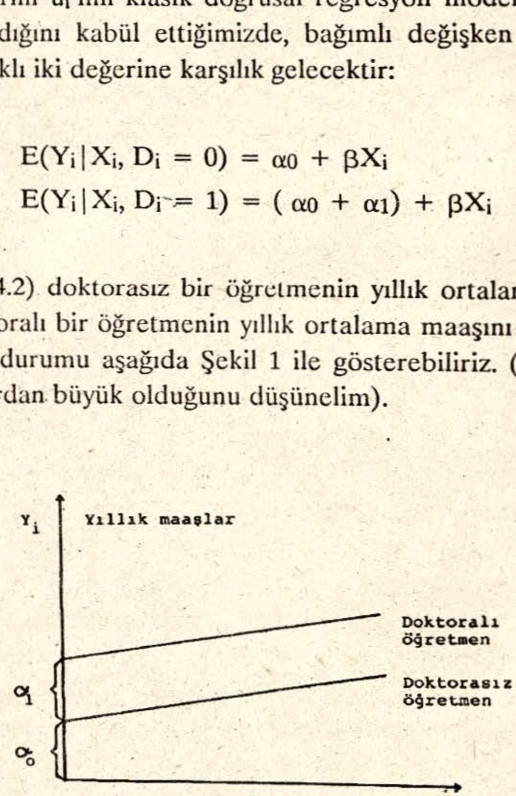
= 0 diğer durumlar.

Görüldüğü gibi model bir nicel değişken $X_i$ (yıllık eğitim harcamaları) ile bir nitel değişken $D_i$ (öğretmenlerin eğitim düzeyleri) ihtiva etmektedir. Modelde yer alan kalıntı terim $u_i$'nin klasik doğrusal regresyon modeli için ileri sürülen varsayımları sağladığını kabül ettiğimizde, bağımlı değişken $Y_i$ için beklenen değerler $D_i$'nin farklı iki değerine karşılık gelecektir:

$$E(Y_i | X_i, D_i = 0) = \alpha_0 + \beta X_i \tag{4.2}$$

$$E(Y_i | X_i, D_i = 1) = (\alpha_0 + \alpha_1) + \beta X_i \tag{4.3}$$

Denklem (4.2) doktorasız bir öğretmenin yıllık ortalama maaşını, Denklem (4.3) ise doktoralı bir öğretmenin yıllık ortalama maaşını vermektedir. Geometrik olarak bu durumu aşağıda Şekil 1 ile gösterebiliriz. (Farazi olarak izah amacıyla $\alpha_1$'in sıfırdan büyük olduğunu düşünelim).




*Şekil: 1*
*Öğretmenlerin yıllık maaşları ile yıllık eğitim harcamaları arasındaki farazi ilişki*

Model (4.1)'deki doktoralı ve doktorasız öğretmenlerin yıllık maaş fonksiyonlarını (4.2) ve (4.3)'deki gibi ayrı ayrı ele alıp incelediğimizde her iki fonksiyonun aynı eğim katsayısına ($\beta$'ya) sahip olduklarını fakat kesme terimlerinin farklı olduklarını görürüz. Bir başka ifadeyle doktoralı öğretmenlerin ortalama maaş düzeyleri, doktorasız öğretmenlerin ortalama maaşlarından $\alpha_1$ kadar farklıdır, fakat yıllık eğitim harcamaları ile ortalama yıllık maaşlardaki değişme oranı her iki durum için de aynıdır.

Yukarıda vurgulanan "eğim katsayısının eşit olduğu" varsayımı geçerli olursa ve eğitim düzeyleri arasında bir farklılık olmaması halinde (4.2 ve 4.3'ün kesme terimleri aynı ise) (4.1)'deki genel regresyon için bir hipotez testini kolaylıkla uygulamamız mümkün olabilir ve geleneksel t-testine başvurarak $\alpha_1$'in istatistiki olarak anlamlı olup olmadığını belirleyebiliriz[15].

## 5. KUKLA DEĞİŞKEN TUZAĞI VE BAZI PROBLEMLER

Regresyon denklemlerine kukla değişkenlerin katılmasıyla geliştirilen modelleri daha ileri aşamalara götürmeden, oldukça önemli bazı problemler üzerinde durmamız gereklidir. Yukarıda ele aldığımız örneğe tekrar geri dönerek model üzerinde biraz daha değişik düşünmeye çalışalım: Örneğimizde nitel bir değişken olarak eğitim düzeylerini ele almış ve bunu iki kategori içerisinde (doktoralı öğretmen ve doktorasız öğretmen) incelemiştik. Burada eğitim düzeylerini bir nitel vasıf olarak doktoralı ve doktorasız gibi ayırıma tabi tutmak ile ancak bir kukla değişkene $D_i$'ye sahip olabiliriz. Eğer kukla değişken $D_i = 1$ değerini alıyorsa bu sürekli olarak doktora derecesine sahip olan bir öğretmeni, $D_i = 0$ değerini alıyorsa bu da sürekli olarak doktora derecesine sahip olmayan bir öğretmeni temsil edecektir. Çünkü ancak iki durum mevcuttur. Dolayısıyla bir kukla değişken iki kategori ayırımına tabi tutulması yeterli olacaktır. Kovaryans analizi modelini şimdi aşağıdaki gibi yazdığımızı düşünelim:

$$Y_i = \alpha_0 + \alpha_1 D_{1i} + \alpha_2 D_{2i} + \beta X_i + u_i \qquad (5.1)$$

burada $Y_i$ ve $X_i$ daha önce tanımlandığı gibidir.

$D_{1i}$ = 1 eğer öğretmen doktoralı ise,
= 0 diğer durumlar

$D_{2i}$ = 1 eğer öğretmen doktorasız ise,
= 0 diğer durumlar.

15 Daha fazla bilgi için bakınız Gujarati (1979:289).

Modelde yer alan parametrelerin OEKK tahminlerini veren $(X'X)^{-1} X'y$ denkleminde (X'X) matrisinin elemanlarının tanımlanmasında dikkat edilmesi gerekli önemli birkaç husus vardır. Zira Denklem (5.1)'de $D_{1i}$ ve $D_{2i}$ arasında tam bir doğrusal bağımlılık sözkonusudur[16].

Farazi bir örnek üzerinde durarak bu sorunu açıklamaya çalışalım. Farzedelim ki, üç doktoralı ve iki doktorasız öğretmenin yer aldığı bir küçük örnekleme sahibiz. Veri matrisinin şu şekilde olduğunu düşünelim:

| Öğretmenler | ($\alpha_0$ için çarpan) | $D_{1i}$ | $D_{2i}$ | $X_i$ |
|---|---|---|---|---|
| Doktoralı | 1 | 1 | 0 | $X_1$ |
| Doktorasız | 1 | 0 | 1 | $X_2$ |
| Doktorasız | 1 | 0 | 1 | $X_3$ |
| Doktoralı | 1 | 1 | 0 | $X_4$ |
| Doktoralı | 1 | 1 | 0 | $X_5$ |

Dikkat edilirse yukarıda verilen veri matrisinin sağ tarafındaki ilk sütun genel kesme teriminin (zımni olarak yer aldığı düşünülen değişkenin) bir çarpan değerini vermektedir. Bu tablodan kolaylıkla $D_{1i} = 1 - D_{2i}$ veya $D_{2i} = 1 - D_{1i}$ sonucunu çıkarabiliriz, yani $D_{1i}$ ve $D_{2i}$ tam olarak doğrusal bağımlıdır. Tahminci denkleminde yer alan (X'X) matrisini aşağıdaki gibi daha açık yazmaya çalışalım: $\alpha_0$'ın önündeki değişkeni $X_0$ olarak tanımlayalım,

$$(X'X) = \begin{bmatrix} \Sigma X_0^2 & \Sigma X_0 D_1 & \Sigma X_0 D_2 & \Sigma X_0 X \\ \Sigma X_0 D_1 & \Sigma D_1^2 & \Sigma D_1 D_2 & \Sigma D_1 X \\ \Sigma X_0 D_2 & \Sigma D_1 D_2 & \Sigma D_2^2 & \Sigma D_2 X \\ \Sigma X_0 X & \Sigma D_1 X & \Sigma D_2 X & \Sigma X^2 \end{bmatrix}$$

Yukarıda veri matrisinde yer alan değerleri (X'X) matrisinde kullanırsak ortaya çıkabilecek sorunları daha açık bir şekilde görebiliriz:

16 Kukla değişken tuzağı hakkında daha fazla bilgi için bakınız Dutta (1975:160-1); Kmenta (1971: 413); Maddala (1979:134); Pindyck ve Rubinfeld (1981:113); Gujarati (1979: 291); İşyar (1989: 151); Jonston (1981:222-3) ve Daniel (1957).

$$(X'X) = \begin{bmatrix} 5 & 3 & 2 & \sum_1^5 X \\ 3 & 3 & 0 & \sum_1^3 X \\ 2 & 0 & 2 & \sum_1^2 X \\ \sum_1^5 X & \sum_1^3 X & \sum_1^2 X & \sum_1^5 X^2 \end{bmatrix}$$

Dikkat edilirse, (X'X) matrisinin 1'inci sütunu, 2'inci ve 3'üncü sütunların toplamıyla bulunmaktadır. Eğer matrisin birinci sütunu diğer iki sütunun toplamına eşitse, tekil olma koşulu (matris kuralı) doğmasından dolayı bu matrisin tersi alınamaz. OEKK tahminci denkleminde $(X'X)^{-1}$ X'y 'den (X'X) matrisinin tersi alınamadığı durumlarda, yani $(X'X)^{-1}$ 'in hesaplanamadığı hallerde yukarıdaki tahminci denklemimiz için de bir çözüm bulunamayacaktır. Dolayısıyla parametreleri hesaplamamız mümkün olamıyacaktır. Parametrelerin tahminlerinin hesaplanamadığı böyle bir duruma "kukla değişken tuzağı" denilmektedir.

Kukla değişken tuzağı sorunu sürekli karşımızda olduğu müddetçe biz herhangi bir tahminde bulunmamız sözkonusu olamaz. Bu açıdan kukla değişken tuzağından kurtulmak gereklidir. Bu amaçla yararlanabileceğimiz bazı yöntemler mevcuttur. Bunlardan en basit ve en çok başvurulanı modelde yer alması düşünülen nitel değişkenin sahip olduğu vasıf (kategori, sınıf gibi) sayısına göre hareket etmektir. Burada uygulanan kural şudur: *Eğer bir nitel değişken m sayıda vasfa (kategoriye, sınıfa) sahipse, modele yalnızca m–1 sayıda kukla değişken katılır.* Yukarıdaki durumda eğitim düzeyi iki kategorili olduğundan 2 – 1 = 1, yani tek bir kukla değişken ile temsil edilebilir. Dolayısıyla yukarıda hazırlanan veri matrisinde $D_1$ veya $D_2$'den herhangi birini ihmal edersek çoklu doğrusal bağlantı sorunu ortadan kalkmış olacaktır, aksi takdirde kukla değişken tuzağı bir sorun olarak kalmaya devam edecektir.

Doktoralı ve doktorasız gibi iki kategoriye 1 ve 0 değerlerinin atanması genel anlamda tamamen keyfi olarak verilmiştir. Değer atamasını tersi bir durum için de yapabiliriz, doktorasız öğretmene 1, doktoralı öğretmene 0 değerleri verilebilir. Bu durumda doktorasız bir öğretmen için $Y_i$'nin beklenen değeri

$$E(Y_i | X_i, D_i = 1) = (\alpha_0 + \alpha_1) + \beta X_i \qquad (5.2)$$

doktoralı bir öğretmen için beklenen değer ise

$$E(Y_i | X_i, D = 0) = \alpha_0 + \beta X_i \qquad (5.3)$$

olacaktır.

Diğer bir sorun ise, nitel değişkenin taşıdığı vasıflardan hangisinin temel-baz kabül edileceğidir. Yukarıdaki örneğimizde temel kategori için doktorasız öğretmen kabul edilmişti. Bir başka ifadeyle 0 değerinin atandığı kategori temel kategoriyi oluşturmaktadır. Dikkat edilirse kesme terimi $\alpha_0$ temel kategori için kesme terimi olacaktır. Eğer biz sürekli olarak doktorasız öğretmenleri dikkate alan regresyon modeli ile çalışırsak kesme terimi $\alpha_0$ olacaktır.

Burada kukla değişkene bağlı $\alpha_1$ katsayısı diferansiyel (fark) kesme terimi olarak bilinmektedir ve bu kesme terimi temel-bas kategori kesme teriminden farklıdır[17].

## 6. KOVARYANS ANALİZİ MODELLERİNDE (ACOV) İKİ VEYA DAHA FAZLA KATEGORİLİ BİR NİTEL (KUKLA) DEĞİŞKENİN YER ALMASI

Yukarıdaki örneğimizde kovaryans analizi modelinde yer alan kukla değişkenin, ya da tanımlanan eğitim düzeyi iki kategorili olarak verilmiş ve bu şekliyle Denklem (4.1) hazırlanmıştı. Bu kez aynı örneğimizde eğitim düzeyini yüksek lisans derecesinden az (lisans veya dengi bir düzey), yüksek lisans ve doktora dereceleri gibi bir ayırıma tabi tutabiliriz. Yukarıda vurgulanan kural ile modele katacağımız kukla değişken sayısı belirlenmeye çalışılır. Bu amaçla kukla değişken sayısı nitel değişkenin sahip olduğu kategori sayısından bir eksiği olacağından, modele en fazla iki kukla değişken (3-1-2) katabiliriz. Dolayısıyla önerilen model:

$$Y_i = \alpha_0 + \alpha_1 D_{1i} + \alpha_2 D_{2i} + \beta X_i + u_i \qquad (6.1)$$

burada $Y_i$ ve $X_i$ daha önce tanımlandığı gibidir,

$D_{1i}$ = 1 yüksek lisans dereceli ise

= 0 diğer durumlar

$D_{2i}$ = 1 doktora dereceli ise,

= 0 diğer durumlar.

Burada temel kategori olarak "yüksek lisans derecesinden az" olan kategori keyfi olarak seçilmiştir. Dolayısıyla kesme terimi $\alpha_0$ bu kategoriyi yansıtacaktır. Diferansiyel (fark) kesme terimleri $\alpha_1$ ve $\alpha_2$ temel kategoriden farklı olarak diğer kategorilerin kesme terimlerini yansıtacaktır. Kalıntı terimi $u_i$'nin klasik doğrusal regresyon modeli için ileri sürülen varsayımları sağladığını düşündüğümüzde, $Y_i$ için beklenen değerler aşağıdaki gibi verilebilir:

$$E(Y_i | D_1 = 0, D_2 = 0, X_i) = \alpha_0 + \beta X_i \qquad (6.2)$$

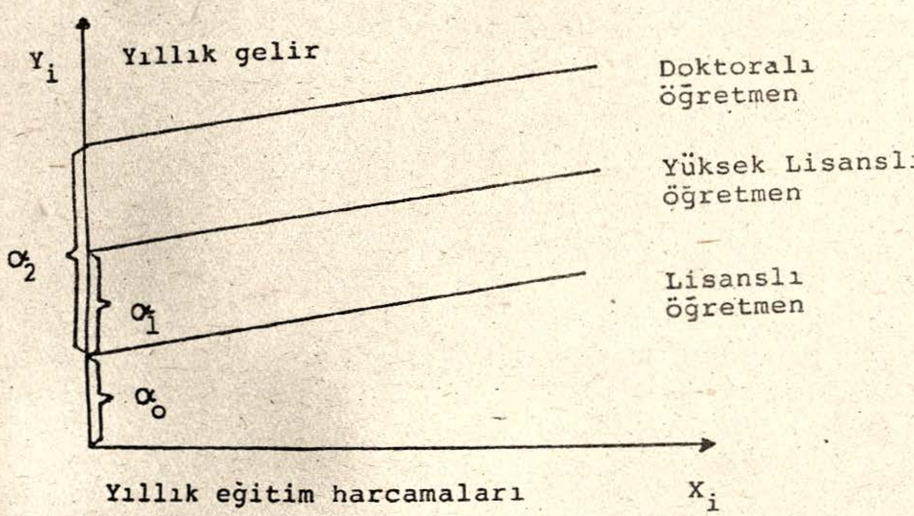
---

17 Gujarati (1979:291).

$$E(Y_i | D_1 = 1, D_2 = 0, X_i) = (\alpha_0 + \alpha_1) + \beta X_i \tag{6.3}$$

$$E(Y_i | D_1 = 0, D_2 = 1, X_i) = (\alpha_0 + \alpha_2) + \beta X_i \tag{6.4}$$

Bu üç denklem sırasıyla yüksek lisans derecesinden az, yüksek lisans dereceli ve doktora dereceli öğretmenlerin ortalama yıllık gelir ile ilgili fonksiyonlarını vermektedir. Geometrik olarak bu durumu aşağıdaki gibi gösterebiliriz (izah amacıyla $\alpha_2 > \alpha_1$ olduğunu varsayalım):




*Şekil: 2*

*Farklı üç eğitim düzeyi için yıllık gelir ve eğitim harcamaları ilişkisi*

Yukarıdaki (6.1) ile verilen kovaryans analizi modelinde, biz kukla değişkenler için farklı bir değer atama şeması uyguladığımızda temel kategori de değişecektir. Örneğin şemayı $D_{1i} = 1$ olarak yüksek lisans derecesinden az; $D_{2i} = 1$ olarak yüksek lisans derecesi ataması yaptığımızda, temel kategori doktora derecesi olacaktır[18].

## 7. KOVARYANS ANALİZİ MODELLERİNDE (ACOV) FARKLI İKİ VEYA DAHA FAZLA KUKLA DEĞİŞKENİN YER ALMASI

Kovaryans analizi modellerinde nicel değişken yanında birbirinden farklı iki veya daha fazla kukla değişkene yer vermek mümkündür. Yukarıdaki örneğimizi burada da kullanabiliriz. Denklem (4.1) ile verdiğimiz kovaryans analizi modelinde eğitim düzeyleri ile ilgili bir nitel değişkene yer vermiş ve nitel değişke-

18 Bakınız Gujarati (1979:292-3) ve Dutta (1975:161-5).

nin iki kategorili olmasından ötürü de tek bir kukla değişken ile de temsil etmiştik. Şimdi aynı modele bir ikinci nitel değişken kattığımızı düşünelim. Öğretmenlerin yıllık kazançlarına, eğitim harcamaları ve eğitim düzeyleri yanında bir de onların yabancı bir dili bilip bilmedikleri etkili olsun, yani etkileyici bir faktör olarak modele yabancı bir dil bilme unsurunu katalım. Bu amaçla modeli şu şekilde yeniden yazabiliriz[19].

$$Y_i = \alpha_0 + \alpha_1 D_{1i} + \alpha_2 D_{2i} + \beta X_i + u_i \tag{7.1}$$

burada $Y_i$ ve $X_i$ daha önce tanımlandığı gibidir.

$D_{1i}$ = 1 eğer öğretmen doktora derecesine sahip ise
= 0 diğer durumlar

$D_{2i}$ = 1 eğer öğretmen bir yabancı dil biliyorsa
= 0 diğer durumlar.

Herbir nitel değişken için ayrı ayrı kukla değişkenler kullanılmıştır. Taban kategori olarak "doktorasız ve yabancı dil bilmeyen" kategorisini dikkate alacağız. Kalıntı terim $u_i$'nin daha önce olduğu gibi klasik doğrusal regresyon varsayımlarını sağladığını düşünmekteyiz. Dolayısıyla $Y_i$ için beklenen değerler aşağıdaki gibi olacaktır:

Doktorasız ve yabancı dil bilmeyen bir öğretmen için ortalama yıllık kazanç

$$E(Y_i | D_{1i} = 0, D_{2i} = 0, X_i) = \alpha_0 + \beta X_i \tag{7.2}$$

denklemi ile, doktoralı ve yabancı dil bilmeyen bir öğretmen için ortalama yıllık kazanç

$$E(Y_i | D_{1i} = 1, D_{2i} = 0, X_i) = (\alpha_0 + \alpha_1) + \beta X_i \tag{7.3}$$

denklemi ile, doktorasız ve yabancı dil bilen bir öğretmen için ortalama yıllık kazanç

$$E(Y_i | D_{1i} = 0, D_{2i} = 1, X_i) = (\alpha_0 + \alpha_2) + \beta X_i \tag{7.4}$$

denklemi ile ve doktoralı ve yabancı dil bilen bir öğretmen için ortalama yıllık kazanç

$$E(Y_i | D_{1i} = 1, D_{2i} = 1, X_i) = (\alpha_0 + \alpha_1 + \alpha_2) + \beta X_i \tag{7.5}$$

denklemi ile verilmektedir.

---

19 Ekonometride çokça başvurulan bu modeller için bakınız Common (1976:372-5); Dutta (1975:165); Kelejian ve Oates (1981:172-3); Maddala (1979:134-6); Kmenta (1971:419-23) ve Kennedy (1979:142-3).

Yukarıdaki dört denklemin hepsinde dikkat edilirse eğim katsayıları eşittir, kesme terimleri ise birbirlerinden farklıdır. Denklem (7.1)'e OEKK uygulayarak tahmin edilen katsayıların anlamlılıklarını test etmek için bir hipotez testine tabi tutabiliriz. Örneğin, $\hat{\alpha}_2$ istatistiki olarak anlamlı ise bir öğretmenin yıllık kazancı yabancı dil bilme faktöründen etkilenecektir. Benzer biçimde $\hat{\alpha}_1$ istatistiki olarak anlamlı ise, bu kez bir öğretmen yıllık kazancı doktora yapma faktöründen etkilenecektir. Eğer bu her iki diferansiyel kesme terimleri istatistiki olarak birlikte anlamlı iseler doktora faktörü yanında, bir yabancı dil bilme faktörü de bir öğretmenin yıllık kazancını belirleyen iki önemli etken olarak kabul edilir.

## 8. KUKLA DEĞİŞKENLERİN TOPLAMA, ÇARPMA VE BİRARADA KULLANILMA KURALLARI

Regresyon modellerinde kukla değişkenlerin etkilerini varyans analizi ve kovaryans analizi modelleri gibi iki farklı yapıda ele alarak incelemeye çalıştık. Kukla değişkenlerin regresyon modellerinde kullanımına ilişkin diğer bir ayırım da; toplama, çarpma ve birarada (karma) kullanımına göre yapılmaktadır.

Kukla değişkenlerin toplama kuralı: Basit bir kovaryans analizi modelini ele aldığımızı düşünelim ve modelimizde bir kesme (sabit) terimin bulunduğunu kabul edelim. Modelde yer alan kukla değişkenin, örneğin iki kategorisi varsa, herbir kategori için, bağımlı değişkenin beklenen değeri farklı olacağından iki farklı denkleme sahip olacağız. Bu durumu yukarıdaki örneklerde açıkça verilmektedir. Biz burada bu birbirinden farklı beklenen denklemlere alt-grup denklemleri diyelimmm. Dikkat edildi ise, bu alt-grup denklemlerin herbirinin kesme terimleri birbirinden farklıdır fakat eğim sayıları hepsinde eşittir. Kısaca, regresyon modellerinde kukla değişkenlerin etkisi yalnızca kesme terimleri üzerinde görülüyorsa, kukla değişkenlerin toplama kuralı sözkonusu olacaktır[20].

Kukla değişkenlerin çarpım kuralı: Yine basit bir kovaryans analizi modelini ele alalım fakat bu kez modelimizde yer alan kukla değişkenin etkisi kesme terimleri üzerinde değil de, eğim katsayısı üzerinde görülüyorsa, kukla değişkenin çarpım kuralı geçerli olacaktır. Kukla değişkenin iki kategorili olduğunu düşündüğümüzde, bağımlı değişkenin beklenen değeri için iki farklı alt-grup denklemi elde edeceğiz, ancak bu denklemlerin kesme terimleri yukarıda birinci durumdakinin aksine birbirine eşit olacak, eğim katsayıları birbirinden farklı olacaktır. Yalnız bu durumda kukla değişken modelde kullanılan nicel değişken ile çarpım halindedir. Dolayısıyla kukla değişkenin etkisi nicel değişkenle birlikte görülecektir[21].

---

20 Bakınız Stewart ve Wallis (1981:172-3); Kelejian ve Oates (1981:170-1) ve Stewart (1979:111-3).
21 Bakınız Stewart ve Wallis (1981:174-5); Stewart (1979:113) ve Kmenta (1971:420-1).

Kukla değişkenlerin birlikte kullanım kuralı: Regresyon modellerinde kukla değişkenlerin bu üçüncü kullanım kuralında, yukarıdaki her iki kuralın aynı regresyon denkleminde birarada kullanılması ve dolayısıyla hem kesme, hem de eğim katsayılarında kukla değişkenlerin etkilerinin görülmesine yol açacaktır. Diğer bir ifadeyle, bu kullanımda modeldeki kukla değişken hem toplama kuralına göre alt-grup denklemlerin hepsinde kesme terimlerinin birbirinden farklı olmasına neden olacak, hem de çarpma kuralına göre kukla değişken nicel değişkenle beraber modelde yer alarak eğim katsayısının alt-grup denklemlerinde farklı olmasına neden olacaktır[22].

Geometrik olarak bu üç durumu düşünecek olursak şu farklılıklar karşımıza çıkacaktır: Toplama kuralında regresyon doğrusu eğimi değişmeden aşağı veya yukarıya doğru kaymaktadır değişen yalnızca kesme teriminin büyüklüğüdür. Çarpma kuralına göre kukla değişkenler regresyon doğrusunun kesme terimi üzerinde herhangi bir etkisi olmadığından yalnızca doğrunun bu kez eğimi değişmektedir. Üçüncü durumda, yani kukla değişkenlerin hem toplama, hem de çarpma kuralına göre regresyon denkleminde birarada kullanılması sırasında regresyon doğrusunun hem kesme terimleri, hem de eğimi değişmektedir. Bu durumu aşağıdaki örnek yardımıyla incelemeye çalışalım[23]. Yukarıda sıkça üzerinde durduğumuz öğretmenlerin yıllık kazançları ile ilgili regresyona tekrar dönüp, doktoralı ve doktorasız öğretmenlerin maaş regresyonlarını ayrı ayrı ele almaya çalışalım.

Doktorasız bir öğretmenin maaş regresyonu

$$Y_i = \phi_0 + \phi_1 X_i + u_i \quad i = 1, 2, 3, ...., N_1 \tag{8.1}$$

ile ve doktoralı bir öğretmenin maaş regresyonunu ise

$$Y_i = \theta_0 + \theta_1 X_i + u_i \quad i = 1, 2, 3, ...., N_2 \tag{8.2}$$

ile verilsin. Bu her iki grup regresyonda $N_1$ ve $N_2$ gözlem sayıları eşit olmayabilir. Şimdi bu iki regresyonu şu dört varsayım altında incelemeye çalışalım:

1. $\phi_0 = \theta_0$ ve $\phi_1 = \theta_1$' dir, yani her iki regresyonun hem kesme terimleri, hem de eğim katsayıları birbirine eşittir.

2. $\phi_0 \neq \theta_0$ ve $\phi_1 = \theta_1$'dir, yani her iki regresyonun kesme terimleri birbirinden farklıdır fakat eğim katsayıları birbirine eşittir.

3. $\phi_0 = \theta_0$ ve $\phi_1 \neq \theta_1$'dir, yani her iki regresyonun kesme terimleri birbirine eşittir fakat eğim katsayıları birbirinden farklıdır.

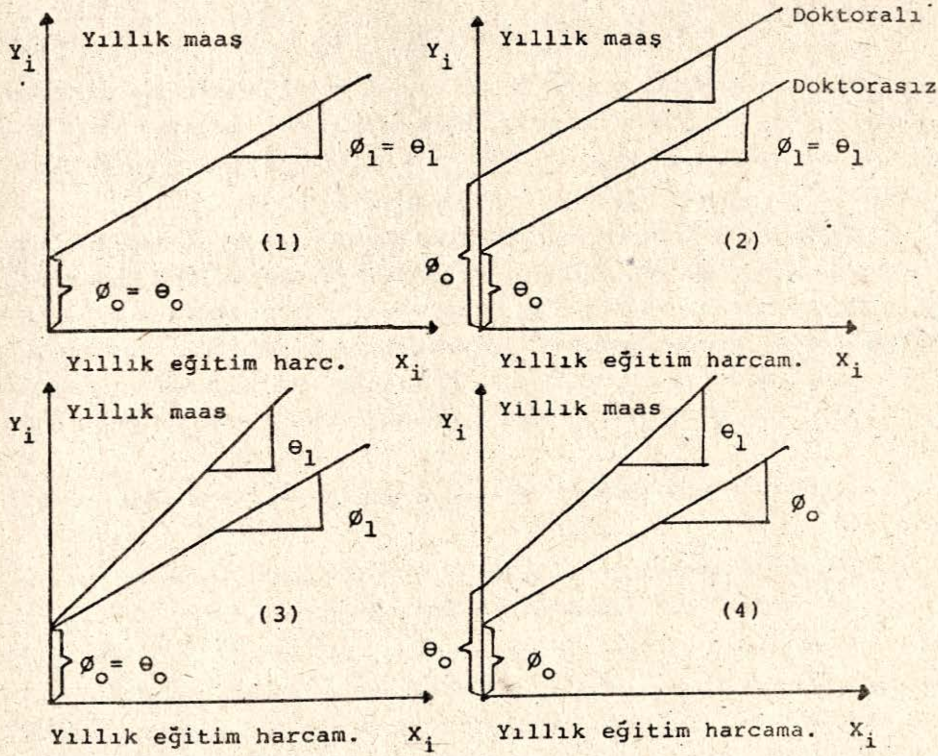
---

22 Bakınız Stewart ve Wallis (1981:175-8).

23 Daha fazla bilgi için bakınız Gujarati (1979:295-7); Kmenta (1971); Stewart ve Wallis (1981:172-8) ve Maddala (1979:136).

4. $\phi_0 \neq \theta_0$ ve $\phi_1 \neq \theta_1$'dir, yani her iki regresyonun hem kesme, hem de eğim katsayıları birbirinden farklıdır.

Yukarıda verilen bu dört olasılığı aşağıdaki gibi farazi dört çizim ile kolaylıkla gösterebiliriz. Şekil 3 doktoralı ve doktorasız bir öğretmenin yıllık maaşları ile eğitim harcamaları arasındaki fonksiyonel ilişkilerini göstermektedir. Dikkat edilirse yukarıda kısaca birer kural olarak açıklamaya çalıştığımız toplama kuralını şekil 3'ün ikinci diyagramında, çarpma kuralını üçüncü diyagramında ve birarada kukla değişkenlerin kullanımı kuralını da dördüncü diyagramında açıkça görmekteyiz.




*Şekil: 3*
*Doktoralı ve doktorasız öğretmenlerin farazi maaş fonksiyonları*

Elimizdeki mevcut verileri Denklem (8.1) ve (8.2) için kullanarak uygun istatistiki testleri yapabiliriz. Bununla birlikte burada yapabileceğimiz yukarıda ileri sürülen dört olasılık durumu için $N_1$ ve $N_2$ gözlemlerini biraraya getirerek aşağıdaki regresyonu test edebiliriz:

$$Y_i = \alpha_0 + \alpha_1 D_i + \beta_1 X_i + \beta_2(D_i X_i) + u_i \qquad (8.3)$$

burada $Y_i$ ve $X_i$ daha önce tanımlandığı gibidir,

$D_i$ = 1 öğretmen eğer doktoralı ise

= 0 diğer durumlar (yani doktorasız ise).

Burada sunulan model daha önce (4.1) ile verilen modelden farklıdır. Model (8.3)'de ilave $(D_i X_i)$ değişkeni mevcuttur. Modelde yer alan kalıntı terimi $u_i$'nin klasik doğrusal regresyon modeli için ileri sürülen varsayımları sağladığını kabul ettiğimizde $Y_i$ için beklenen değerler şöyle olacaktır:

$$E(Y_i | D_i = 0, X_i) = \alpha_0 + \beta_1 X_i \qquad (8.4)$$

$$E(Y_i | D_i = 1, X_i) = (\alpha_0 + \alpha_1) + (\beta_1 + \beta_2) X_i \qquad (8.5)$$

Bu denklemler sırasıyla, doktorasız ve doktoralı öğretmenlerin ortalama yıllık maaş fonksiyonlarını vermektedir. (8.1) ve (8.2) de olduğu gibi $\phi_0 = \alpha_0$, $\phi_1 = \alpha_1$, $\theta_0 = (\alpha_0 + \alpha_1)$ ve $\phi_1 = (\beta_1 + \beta_2)$ eşitlikleri çıkarılabilir. Dolayısıyla tahmin edilen (8.3) iki bireysel maaş fonksiyonu ile özdeştir.

Model (8.3)'deki $\alpha_1$ diferansiyel kesme terimidir. $\beta_2$ ise diferansiyel eğim katsayısıdır. $\beta_2$ aynı zamanda doktoralı bir öğretmenin maaş fonksiyonunun eğim katsayısının, doktorasız bir öğretmenin eğim katsayısından ne kadar farklı olduğunu gösterir. Çarpım formundaki Di'nin sunumu ($X_i$ ile $D_i$'nin çarpımı) iki grubun eğim katsayıları arasındaki farklılığı ortaya koyar. Toplama formundaki kukla değişkenin sunumunda ise iki grubun kesmeleri arasındaki ayırım ortaya konur.

Model (8.3)'ün, Model (8.1) ve (8.2)'ye göre avantajlı birkaç yönü vardır. Bunları kısaca şöyle sıralayabiliriz[24]:

1. Bireysel regresyonlar Denklem (8.4) ve (8.5)'deki gibi sonuçlanabilir, oysa (8.3)'deki gibi tek bir denklem elde etmek mümkündür.

2. (8.3) ile elde edilen tekli regresyonu kolaylıkla hipotez testlerinin uygulanmasında kullanabiliriz. Örneğin diferansiyel kesme katsayısı $\hat{\alpha}_1$ istatistik olarak anlamlı ise, iki regresyonun genel bir kesme terimine veya benzer biçimde diferansiyel eğim katsayısı $\hat{\beta}_2$ istatistiki olarak anlamlı ise, iki regresyonun genel bir eğim katsayısına sahip olduğuna karar verebiliriz.

3. Serbestlik derecesindeki artıştan ötürü, tahmin edilen parametrelerin nispi doğrulukları daha da artabilecektir.

---

24 Bakınız Gujarati (1979:297-8).

## 9. SONUÇ

Kukla değişken tekniği çok yönlü bir alet olmasına rağmen dikkatli bir biçimde ele alınıp kullanılması gerekir. Bu amaçla regresyon modellerinde kukla değişkenlerden yararlanılmak isteniyorsa şu noktalara uymakta yarar vardır: Birincisi, eğer regresyon modeli bir sabit terim içeriyorsa, kukla değişkenlerin sayısı modele katılan herbir nitel değişkenin sahip olduğu kategori (nitelik, sınıf ve vasıf) sayısından bir eksiği olmalıdır. İkincisi, kukla değişkenlere bağlı katsayılar daima temel-bas kategorilere dayanarak yorumlanmalıdır ki bas kategori sıfır değerinin verildiği kategoridir. Üçüncü olarak, eğer model birkaç kategorili ve birkaç nitel değişkeni birlikte içeriyorsa, kukla değişkenlerin sunumunda büyük sayıda serbestlik derecesi kaybı sözkonusu olabilir. Bundan ötürü, herhangi bir model ile ilgili çalışmada eldeki mevcut verilerin toplam sayısına bakarak, modele katılacak kukla değişkenlerin sayısına karar verilmelidir.

Bu çalışmamız daha çok kukla değişkenleri birer açıklayıcı değişken olarak alan varyans ve kovaryans analizi modelleri üzerinde yoğunlaştırdık. Nitel bağımlı değişkenlerin yer aldığı veya kullanıldığı olasılık modelleri üzerinde durmadık. Nitel bağımlı değişkenlerin yer aldığı ve kukla değişkenlerin diğer kullanımları ile ilgili modelleri bir başka çalışmaya bırakarak yalnızca açıklayıcı değişken olarak kukla değişkenlerin kullanımı ile sınırlandırmamıştır.

## KAYNAKLAR


**Common, M.S. (1976),** *Basic Econometrics: An Introductory for Economists:* London: Longman.

**Daniel, B.S. (1957),** "Use of the Dummy Variables in Regression Equations", *Journal of the American Statistical Association,* Dec. 52(280), s. 548-51.

**Draper, N.R. ve H. Smith (1966),** *Applied Regression Analysis,* New York: John Wiley and Sons, Inc.

**Dutta, M. (1975),** *Econometric Methods,* Cincinnati: South-Western Publ. Comp.

**Ferber, R. ve P.J. Verdoorn (1962),** *Research Methods in Economics and Business,* New York: The Macmillian Comp.

**Gujarati, D. (1979),** *Basic Econometrics,* London: International Student Edition.

**İşyar, Y. (1989),** *Ekonometrik Modeller I-II,* Bursa: (Basılmamış Ders Notları).

**Intriligator, M.D. (1978),** *Econometric Models, Techniques and Applications,* Amsterdam: North-Holland Publ. Comp.

**Johnston, J. (1981),** *Ekonometrik Metodlar,* Çev. Y. İşyar ve E. Kip, Erzurum: Atatürk Üniversitesi Yayınları.

**Kelejian, H.H., ve W.E. Oates (1981),** *Introduction to Econometrics,* Second Edit., New York: Harper International Edit.

**Kennedy, P. (1979),** *A Guide to Econometrics,* Oxford: Martin Robertson and Comp. Ltd.

**Kılıçbay, A. (1980),** *Ekonometrinin Temelleri,* İstanbul: İ.Ü. İktisat Fakültesi Yayınları, Yayın No: 454.

**Kmenta, J. (1971),** *Elements of Econometrics,* New York: The Macmillian Comp.

**Koutsoyjannis, A. (1977),** *Theory of Econometrics: An Introductory Exposition of Econometric Methods,* Hong-Kong: The Macmillian Press., Ltd.

**Maddala, G.S. (1979),** *Econometrics,* London: McGraw-hill International Book Comp.

**Pindyck, R.S. ve D.L. Rubinfeld (1981),** *Econometric Models and Economic Forecasts,* London: McGraw-Hill International Book Comp.

**Sevüktekin, M. (1988),** "Ekonometrik Araştırmalarda Verilerin Kullanılması-I", *Uludağ Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi,* Mart-Kasım, IX(1-2), s. 219-28.

**Stewart, J. (1979),** *Understandig Econometrics,* London: Hutchinson and Comp. Ltd.

**Stewart, M.B. ve K.F. Wallis (1981),** *Introductory Econometrics,* Oxford: Billing and Sons Ltd.