8. KUKLA DEĞİŞKENLER

Popülasyonda ya da örneklemde yer alan denek değeri veri olarak tanımlanır. Veriler sürekli veya kesikli olabileceği gibi nicelik veya nitelik de belirtebilir. Herhangi bir iktisadi olayı açıklamada kullanacağımız değişkenlerden bazıları nicelik özelliklerden ziyade nitelik özelliklere sahip olabilir.

Kukla değişkenler uygulamalarda dummy değişkenler, yapay değişkenler, gruplandırılmış değişkenler, niteliksel değişkenler ve gölge değişkenler olarak da anılmaktadır (Akkaya, 1991). Ekonometrik uygulamalarda sık sık kullanılan değişkenler olarak bilinir. Modelde hangi niteliksel faktörün etkisini temsil edecek olan bağımsız değişkenler ya da açıklayıcı değişkenler, niteliksel değişkenin özelliğine göre iki ya da daha çok değer alarak dahil edilir. Genel itibariyle kukla ya da dummy değişkenler bazı özellikleri sağlaması halinde bir, sağlamaması halinde ise sıfır değeri almaktadır. Kukla, dummy ya da gölge değişkenler regresyon modellerine dahil edilen, niceliksel olmayan ancak niteliksel değişkenler olarak tanımlanabilir.

Kukla değişkenler genel itibariyle şu amaçlar için modele dahil edilmektedir:

a)Savaş ve barış yılları,

b)Ekonomik kriz dönemleri,

c)Farklı sosyal ve ekonomik politikaların uygulandığı dönemler,

d)Herhangi bir iktisadi, sosyal ve politik bir olayda bölgesel farklılıkların ortaya konulması,

e)Cinsiyet, medeni hal, meslek grupları, eğitim düzeyleri, sosyal statü vb. gibi faktörlerin etkisinin modelde incelenmesi.

8.1. Tek Kukla Bağımsız Değişkeni Olan Modeller

Basit doğrusal regresyon modelimiz,

$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ daha önce bu şekilde verilmişti. Kukla değişkenler, basit doğrusal regresyon modelinde yer alan bağımsız değişkenler arasında olabilir. Bir tane bağımsız değişkenli basit doğrusal regresyon modelindeki bağımsız değişken kukla değişken olabilir. Bu durumda yukarıda verilen basit doğrusal regresyon modelinde yer alan bağımsız değişken X_i değerlerinin bir kısmı sıfır (0) değerini $X_i=0$ alırken bir kısmı bir (1) değerini $X_i=1$ alabilmektedir. X_i'nin sıfır (0) ve bir (1) değerlerine göre Y_i beklenen değeri,

$$E(Y_i / X_i = 0) = \beta_0$$

$$E(Y_i / X_i = 1) = \beta_0 + \beta_1$$ şeklinde olacaktır.

Şöyle bir tanımlama yapalım. X_i =0 olan gözlemlere ilişkin Y_i'nin ortalaması $\overline{Y}_0$, X_i =1 olan gözlemlere ilişkin Y_i'nin ortalaması $\overline{Y}_1$ olsun. Bu durumda,

$\overline{Y}_0 = \beta_0$ (X_i =0 olan gözlemler için)

$\overline{Y}_1 = \beta_0 + \beta_1$ (X_i =1 olan gözlemler için) olur. Bu iki eşitlikten,

$$\overline{Y}_1 - \overline{Y}_0 = (\beta_0 + \beta_1) - \beta_0$$

$\overline{Y}_1 - \overline{Y}_0 = \beta_1$ elde edilir.

Bu durumda X_i'nin sıfır (0) ve bir (1) değeri alması durumunda ortalama β_1 kadar bir değişiklik söz konusu olacaktır (Güriş ve Çağlayan, 2010).

Örnek 8.1. Tarım sektöründe faaliyet gösteren bir işletmede çalışan toplam 10 adet daimi işçi ve geçici işçinin aylık net ücretleri aşağıda verilmiştir. Burada, Ortalama aylık net ücret = f(daimi işçi ücretleri, geçici işçi ücretleri) şeklinde bir fonksiyonel ilişki mevcuttur. Buna göre kukla değişkenli modeli tahmin ediniz.

Çizelge 18. Daimi ve Geçici İşçilerin Ortalama Aylık Net Ücretleri (000 TL)

Statü	Ortalama aylık net ücret (Y) (000 TL)
Daimi işçi	1,0
Daimi işçi	1,2
Geçici işçi	0,8
Geçici işçi	0,9
Geçici işçi	0,9
Daimi işçi	1,3
Geçici işçi	0,8
Daimi işçi	1,1
Geçici işçi	0,8
Daimi işçi	1,2

Çözüm:

Kukla değişkenler aşağıdaki gibi tanımlanır:

$$D_i = \begin{cases} Daimi\ \text{İşçi}\ ise & 1 \\ Degilse & 0 \end{cases}$$

Kukla değişkenlere göre Çizelge yeniden düzenlenirse, Çizelge 19. Daimi ve Geçici İşçilerin Ortalama Aylık Net Ücretleri (000 TL) ve Kukla Değişken

Statü	Ortalama aylık net ücret (Y) (000 TL)	Di
Daimi işçi	1,0	1
Daimi işçi	1,2	1
Geçici işçi	0,8	0
Geçici işçi	0,9	0
Geçici işçi	0,9	0
Daimi işçi	1,3	1
Geçici işçi	0,8	0
Daimi işçi	1,1	1
Geçici işçi	0,8	0
Daimi işçi	1,2	1

Regresyon modellerinin tahmini sırasında kukla değişkenlerin bir eksiği modele dahil edilir. Yani bir kantitatif değişkenin m tane sınıfı varsa, m-1 tane yardımcı değişken kullanılmaktadır. Örneğin; cinsiyet özelliğinin erkek ve kadın olarak m=2 tane sınıfı olduğundan m-1=1 tane yardımcı değişken kullanılması gerekir.

$$X'X = \begin{bmatrix} 10 & 5 \\ 5 & 5 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 10 \\ 5{,}80 \end{bmatrix}$$

$$(X'X)^{-1} = \begin{bmatrix} 0{,}2 & -0{,}2 \\ -0{,}2 & 0{,}4 \end{bmatrix}$$

$$\hat{\beta}=\begin{bmatrix} 0{,}2 & -0{,}2 \\ -0{,}2 & 0{,}4 \end{bmatrix}\begin{bmatrix} 10 \\ 5.80 \end{bmatrix}$$

$$\hat{\beta}=\begin{bmatrix} 0{,}84 \\ 0{,}32 \end{bmatrix}$$

Sonuç vektörüne göre basit doğrusal regresyon denklemini $\hat{Y}$=0,84+0,32D_i şeklinde yazılacaktır.

D_i=0 için $\hat{Y}$=0,84+0,32D= 0,84

D_i=1 için $\hat{Y}$=0,84+0,32(1)=1,16

Regresyon denkleminde 0.84 katsayısı geçici işçilerin ortalama aylık net ücretini ifade ederken, 0.84+0.32=1.16 ise daimi işçilerin ortalama aylık net ücretini göstermektedir. Sabit terim D_i=0 olma özelliğine göre yorumlanır. $\hat{\beta}_1$'in değeri pozitif ise daimi işiçilerin aylık net ücretinin, geçici işçilere göre $\hat{\beta}_1$ kadar fazla olduğunu gösterir.

8.2. İki Kukla Bağımsız Değişkeni Olan Modeller

Tek bağımsız değişkenli basit doğrusal regresyon modelinde yer alan bağımsız değişken X_i değerlerinin bir kısmı sıfır (0) değerini $X_i = 0$ alırken bir kısmı bir (1) değerini $X_i = 1$ almaktadır. Regresyon modellerinde ikiden fazla kukla değişken olabilir. Burada, modelde iki kukla değişken olması halinde çift kukla değişkenli modeller sözkonusu olacaktır. Regresyon modeli,

$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i$ şeklinde olsun. Bu modelde yer alan X_{i1} ve X_{i2} sıfır (0) ve bir (1) değerlerini alan kukla değişkenler ise Y_i beklenen değeri,

$$E(Y_i / X_{i1} = 0, X_{i2} = 0) = \beta_0$$

$$E(Y_i / X_{i1} = 1, X_{i2} = 0) = \beta_0 + \beta_1$$

$E(Y_i / X_{i1} = 0, X_{i2} = 1) = \beta_0 + \beta_2$ şeklinde elde edilir.

Kukla değişkenler tanım nedeniyle aynı anda bir (1) değerini almamaktadır. Buna göre,

$$\overline{Y}_1 = \beta_0$$

$$\overline{Y}_2 = \beta_0 + \beta_1$$

$\overline{Y}_3 = \beta_0 + \beta_2$ olacaktır (Güriş ve Çağlayan, 2010).

Örnek 8.2. Tarım sektöründe faaliyet gösteren bir işletmede çalışan toplam 10 adet daimi işçi ve geçici işçinin aylık net ücretleri aşağıda verilmiştir. Aile işçisi aylık ortalama net ücretlerini de ilave edelim. Bu durumda kukla değişkenli modeli tahmin ediniz.

Çizelge 20. Daimi ve Geçici İşçilerin Ortalama Aylık Net Ücretleri (000 TL) ve Bunlara İlişkin Kukla Değişkenler

Statü	Ortalama aylık net ücret (Y) (000 TL)	D_1 (Daimi işçi)	D_2 (Geçici işçi)
Daimi işçi	3	1	0
Aile işçisi	4	0	0
Geçici işçi	5	0	1
Aile işçisi	3	0	0
Geçici işçi	6	0	1
Daimi işçi	2	1	0
Aile işçisi	7	0	0
Daimi işçi	5	1	0
Aile işçisi	4	0	0
Geçici işçi	3	0	1

Tarım sektöründe faaliyet gösteren bir işletmede çalışan daimi işçi, geçici işçi ve aile işçisi aylık ortalama net ücretleri verildi. Bu duruma göre üç farklı durum ortaya çıkmış olup üç değer alan tek kukla değişken kullanılabilir.

Ancak bu olayı iki değer alan kukla değişkenler ile açıklanabileceğini düşünelim. İki kukla değişken ile bu olayı açıklamak mümkün olabilecektir.

Çizelge 21. Daimi ve Geçici İşçilerin Ortalama Aylık Net Ücretleri (000 TL) ve İstatistiki Değerleri

Değişken	Katsayı	Std. Hata	t-İstatistik
Sabit	4,500000	0,809174	5,561229
D1	-1,166667	1,236033	-0,943880
D2	0,166667	1,236033	0,134840
R-kare	0,151235		

Bu durumda regresyon modeli

$Y_i = 4{,}50 - 1{,}17_1 D_1 + 0{,}17 D_2$ olur.

Kukla değişkenler tanım nedeniyle aynı anda bir (1) değerini almamaktadır. Buna göre,

$\overline{Y_1}$ =4,500

$\overline{Y_2}$ = 4,500+(1,167)=3,33

$\overline{Y_3}$ =4,500+0,167=4,67

Ortalama aylık net ücret aile işçileri için 4,50 TL, daimi işçiler için 3,33 TL ve geçici işçiler için ise 4,67 TL'dir.

8.3. Hem Kukla Hem Bağımsız Değişkenli Modeller

Modelde bir veya daha fazla bağımsız değişken olabilir. Modelde yer alan bağımsız değişken ya da değişkenlerin yanına ilaveten bağımsız bir kukla değişken ilavesi yapılabilir. Bu durumda, eklenen bu bağımsız kukla değişkenin hangi katsayıyı etkileyeceği modelin ilk oluşturulma aşamasında belirlenmesi gerekir. Bu bağımsız kukla değişken, sadece sabit katsayıyı, sadece bağımsız değişken katsayılarını ve son olarak her iki katsayıyı yani sabit ve bağımsız değişken katsayılarını birlikte etkileyebilir. Bu durumların hepsinin ayrı ayrı irdelenmesi gerekmektedir.

Örnek 8.3. Tarım sektöründe faaliyet gösteren bir işletmede çalışan toplam 11 adet daimi işçinin aylık net ücretleri ve çalışma süreleri aşağıda verilmiştir. Burada, Ortalama aylık net ücret = f(Ortalama çalışma süresi, Cinsiyet) şeklinde bir fonksiyonel ilişki mevcuttur. Buna göre kukla değişkenli modeli tahmin ediniz.

Çizelge 22. Daimi İşçi ve Cinsiyete Ait Ortalama Aylık Net Ücretleri (000 TL) ve Ortalama Çalışma Süreleri

Cinsiyet	Ortalama aylık net ücret (Y) (000 TL)	Ortalama çalışma süresi (X_1)	D_i
Erkek	1,0	8	1
Kadın	1,2	7	0
Erkek	0,8	9	1
Kadın	0,9	8	0
Erkek	0,9	9	1
Kadın	1,3	7	0
Erkek	0,8	10	1
Kadın	1,1	8	0
Erkek	0,8	9	1
Kadın	1,2	8	0

Çözüm:

Kukla değişkenli çoklu regresyon denkleminde,

$$D_1 = \begin{cases} Erkek & 1 \\ Degil \;\; ise & 0 \end{cases}$$

şeklinde tanımlanır. Modeli sadece sabit katsayıyı etkileyecek şekilde $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 D_i + \varepsilon_i$ oluşturalım.

Buna göre çoklu regresyon modeli tahmininde kukla değişkenlerin bir eksiği modele katılarak sadece D_1 değişkeni dikkate alınarak tahmin edilen model aşağıda verilmiştir.

$$X'X = \begin{bmatrix} 10 & 83 & 5 \\ 83 & 697 & 45 \\ 5 & 45 & 697 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 10 \\ 81.6 \\ 4.3 \end{bmatrix}$$

$\hat{\beta}=(X'X)^{-1}X'Y$ formülü uygulandığında tahmin edilen kukla değişkenli regresyon denklemini aşağıdaki şekli ile elde edebiliriz:

$\hat{Y}$=2,14-0,13X_1-0,096D_i olarak tahmin edilir.

D_i = 1 için $\hat{Y}$=2,14+(-0,096)-0,13X_1

D_i = 0 için $\hat{Y}$=2,14-0,13X_1 olur. Görüleceği gibi sabit terimlerde farklılık ortaya çıkacaktır.

Sabit terim kadınlar için β_0, erkekler için $\beta_0 + \beta_2$ olacaktır. β_0 =ortalama çalışma süresi sıfır olduğunda kadınlar için ortalama aylık net ücrettir. β_1 ortalama aylık net ücretle ortalama çalışma süresi arasında ilişki kuran katsayıdır. Marjinal ortalama aylık net ücreti göstermekte olup, kadın ve erkekler için aynıdır. β_2 kadınlar ile erkekler arasındaki ortalama aylık net ücret farkını gösterir. Buna fark katsayısı denir.