

Kukla Değişkenlerle Bağlanım

Nitel Değişkenlerle Bağlanım



Ekonometri 1 – Konu 28
Sürüm 2,0 (Ekim 2011)




UADMK Açık Lisans Bilgisi

İşbu belge, “Creative Commons Attribution-Non-Commercial ShareAlike 3.0 Unported” (CC BY-NC-SA 3.0) lisansı altında bir açık ders malzemesi olarak genel kullanıma sunulmuştur. Eserin ilk sahibinin belirtilmesi ve geçerli lisansın korunması koşulu ile özgürce kullanılabilir, çoğaltılabilir ve değiştirilebilir. Creative Commons örgütü ve “CC-BY-NC-SA” lisansı ile ilgili ayrıntılı bilgi “<http://creativecommons.org>” adresinde bulunmaktadır. Bu ekonometri ders notları setinin tamamına “<http://www.acikders.org.tr>” adresinden ulaşılabilir.

A. Talha Yalta

TOBB Ekonomi ve Teknoloji Üniversitesi

Ekim 2011 

Ders Planı

- 1 Nitel Deęişkenlerle Baęlanım
 - VARÇÖZ Modelleri
 - KOVÇÖZ Modelleri

Kukla Değişkenler

- Bağlanım çözümlerinde bağımlı değişken, sayısal büyüklükler yanında nitel değişkenlerden de etkilenebilir.

Nicel Değişkenler

Gelir, üretim, fiyat, maliyet, enflasyon, işsizlik oranı, yaş, boy, çocuk sayısı, ...

Nitel Değişkenler

Cinsiyet, ırk, din, savaş, coğrafi bölge, hükümet politikalarında değişme, grev, ...

- Görgül çalışmalarda karşılaşılan pek çok sorunu çözmeye nitelik bildiren **“kukla”** (dummy) değişkenlerden yararlanır.
- Bu değişkenler aynı zamanda **“nitel”** (qualitative) değişken, **“ulamsal”** (categorical) değişken veya **“gösterge”** (indicator) değişkeni olarak da adlandırılmaktadırlar.

Kukla Değişkenler

- Kukla değişkenler bir veri sınıflandırma aracıdır.
- Nitel özellikleri nicel olarak gösterebilmek için, niteliğin varlık ya da yokluğunu gösteren 1 ve 0 değerlerini alırlar.
- Örnek olarak bir kimsenin üniversite mezunu olduğu 1 ile, olmadığı ise 0 ile gösterilebilir.
- Kuklaların mutlaka 0 ya da 1 değerleri almaları gerekmez. Doğrusal ilişkili herhangi bir sayı çifti kullanılabilir.
- Diğer yandan, yorumlamada sağladığı kolaylıktan dolayı uygulamada $\{0, 1\}$ çifti yeğlenmektedir.
- Bağlanım modellerinde kukla değişkenlerin kullanılması ek bir zorluk getirmemektedir.
- Kukla içeren bağlanımların hesaplanması bilindik şekilde olur. Katsayıların anlamlılığı da t istatistiği ile sınanabilir.

VARÇÖZ Modelleri

- Bir bağlanım modelindeki tüm açıklayıcı değişkenlerin birer kukla değişken olması mümkündür.
- Böyle modellere “**VARÇÖZ**” (ANOVA) modelleri de denir.
- VARÇÖZ modelleri özellikle toplumbilim, psikoloji, eğitim, pazar araştırması gibi alanlarda yaygındır.

VARÇÖZ Modelleri

Örnek olarak, Şubat-Mart 2011 döneminde Ankara Çankaya'da satılığa çıkartılan ikinci el daire fiyatlarının üç ayrı semtte nasıl farklılık gösterdiğini inceleyen aşağıdaki modeli ele alalım.

$$Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + u_i$$

- Y burada satış ilanı verilen apartman dairesinin 1000 TL olarak fiyatıdır.
- Kukla değişkenler ise D ile gösterilmektedir.
- $D_{2i} = 1$, daire Dikmen'de ise; $D_{2i} = 0$, eğer değilse.
 $D_{3i} = 1$, daire Kavaklıdere'de ise; $D_{3i} = 0$, eğer değilse.
- **Not:** Eğer $D_{2i} = 0$ ve $D_{3i} = 0$ olursa daire Kavaklıdere ya da Dikmen'de değil, üçüncü seçenek olan Cebeci semtinde bulunuyor demektir.

VARÇÖZ Modelleri

- Bu modelin bize gösterdiği şey şudur:

Dikmen'de ortalama fiyat: $E(Y_i | D_{2i} = 1, D_{3i} = 0) = \beta_1 + \beta_2$

Kavaklıdere'de ortalama fiyat: $E(Y_i | D_{2i} = 0, D_{3i} = 1) = \beta_1 + \beta_3$

Cebeci'de ortalama fiyat: $E(Y_i | D_{2i} = 0, D_{3i} = 0) = \beta_1$

- Bağlanıma ilişkin tahmin sonuçları aşağıdaki gibidir:

$$\hat{Y}_i = 91,4615 + 47,2746 D_{2i} + 94,1218 D_{3i}$$

öh	(11,6983)	(13,6480)	(16,8850)	
t	(7,8184)	(3,4638)	(5,5743)	$R^2 = 0,3491$

- D_2 ve D_3 anlamlı olduğu için, Ankara'nın bu üç semtinde daire fiyatlarının farklı olduğunu söyleyebiliriz.
- Buna göre Cebeci'deki ortalama bir dairenin fiyatı yaklaşık 91 bin TL iken, Dikmen ve Kavaklıdere'deki ortalama fiyat ise sırası ile $91 + 47 \approx 140$ ve $91 + 94 \approx 185$ bin liradır.

Birkaç Önemli Nokta

Kukla değişken kullanımı ile ilgili dikkat edilmesi gereken bazı noktalar şunlardır:

- 1 Bir nitel değişkende m sayıda sınıf ya da “ulam” (category) varsa, $(m - 1)$ kukla değişken kullanılmalıdır. Aksi halde “kukla değişken tuzağı” (dummy variable trap) denilen “tam eşdoğrusallık” (exact collinearity) oluşur. Ancak, sıfır noktasından geçen bağlanımlarda ulam sayısı kadar kukla değişken koymak mümkündür.
- 2 “Dikmen” ve “Dikmen değil” gibi seçeneklere hangi değer atanacağı isteğe bağlıdır. Eğer Dikmen değil = 1 olursa β_2 katsayısı da eksi çıkar. Demek ki kukla değişken içeren modelleri yorumlarken, 1 ve 0 değerlerinin nasıl verildiğini bilmek önemlidir.

(... devam)

Birkaç Önemli Nokta

- 3 0 değeri verilen ulam, “**yazın**” (literature) içerisinde farklı adlarla karşımıza çıkabilmektedir:

Türkçe	İngilizce
“Taban ulam”	(Base category)
“Kıyas ulamı”	(Benchmark category)
“Karşılaştırma ulamı”	(Comparison category)
“Denetim ulamı”	(Control category)
“Gönderi ulamı”	(Reference category)
“Atlanan ulam”	(Omitted category)

Örneğimizde taban ulam Cebeci semtidir.

Taban ulam, diğerlerini karşılaştırmada kullanılan sınıftır.

Taban ulamı gösteren ya da ölçen terim de β_1 sabit terimidir.

D_2 ve D_3 kukla değişkenlerine gelen β_2 ve β_3 katsayılarına ise “**sabit terim farkı**” (constant term difference) adı verilir.

KOVÇÖZ Modelleri

- Birçok iktisadi araştırmada, yalnızca kukla değişkenlerin kullanıldığı VARÇÖZ modellerine çok sık rastlanmaz.
- Bunun yerine nitel ve nicel değişkenlerin birlikte olduğu “kovaryans çözümlemesi” (analysis of covariance) ya da “KOVÇÖZ” (ANCOVA) modelleri yeğlenir.
- Bu modellerde nicel değişkenlere “denetim değişkeni” (control variable) de denir.

KOVÇÖZ Modelleri

KOVÇÖZ modellerine bir örnek olarak Ankara'daki daire fiyatları örneğimizi geliştirelim.

$$Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 X_i + u_i$$

- Y_i burada satılığa çıkartılan apartman dairesinin fiyatıdır.
- X_i ise dairenin metrekare cinsinden büyüklüğüdür.
- $D_{2i} = 1$, daire Dikmen'de ise; $D_{2i} = 0$, eğer değilse.
 $D_{3i} = 1$, daire Kavaklıdere'de ise; $D_{3i} = 0$, eğer değilse.
- **Not:** Cebeci'yi taban olarak ulamlandırmayı sürdürüyoruz.

KOVÇÖZ Modelleri

- Bağlanım tahminleri aşağıdaki gibidir:

$$\hat{Y}_i = -27,7639 + 15,1179 D_{2i} + 72,4160 D_{3i} + 1,24994 X_i$$

öh	(15,6787)	(9,7259)	(11,4113)	(0,1431)
t	(-1,7708)	(1,5544)	(6,3460)	(8,7354)

$$R^2 = 0,7217$$

- Yeni modeldeki sabit terim fark katsayısının Kavaklıdere için anlamlıyken Dikmen için anlamlı olmadığını görüyoruz.
- Diğer bir deyişle, metrekare sabitken, Dikmen'deki daire fiyatlarının Cebeci ile aynı olduğu reddedilmemekte, Kavaklıdere için ise 72 bin liralık bir fark gözlenmektedir.
- Bulguların Dikmen için farklı çıkmasının nedeni, ilk baştaki modelde X 'i hesaba katmamış olmamızdır.
- Sonuçlar Ankara'nın bu üç semtindeki dairelerin metrekare fiyatının yaklaşık 1250 lira olduğunu göstermektedir.
- Burada sabit terimleri farklı olan ancak aynı β_4 eğimini paylaşan 3 farklı bağlanımı ele aldığımızı dikkat ediniz.

Önümüzdeki Dersin Konusu

Önümüzdeki ders

Kukla değişken kullanım şekilleri