

Kukla Değişkenlerle Bağlanım

Kukla Değişken Kullanım Şekilleri




Ekonometri 1 – Konu 29
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 Kukla Deęişken Kullanım Şekilleri
 - Chow Sınamasının Kukla Almaşığı
 - Karşılıklı Etkileşim
 - Parça-Yollu Doğrusal Bağlanım

Chow Sınamasının Kukla Değişken Almaşığı

- Önceki örnekte, nitel değişkenlerin sabit terimi etkilediği ama eğim katsayısını etkilemediği varsayılmıştı.
- Diğer yandan, eğer farklı ulamların eğim katsayısı da farklı ise sabit terim farklarını sınamanın pek anlamı yoktur.
- Birden fazla bağlanımın aynı olup olmadığını sınamak için çok adımlı Chow sınamasının kullanılabildiğini biliyoruz.
- Farklı bağlanımları sabit terimler, eğimler ya da her ikisi yönünden ayırt edebilen daha genel bir sınama yöntemi kukla değişkenler ile olanaklıdır.

Chow Sınamasının Kukla Değişken Almaşığı

- Türkiye için tüketim harcamaları ve milli gelir verilerimizi anımsayalım:

Çizelge: Türkiye’de Tüketim ve GSYH (1987–2006)

Yıl	C	Y	Yıl	C	Y
1987	51.019	74.416	1997	77.620	112.892
1988	51.638	76.143	1998	78.113	116.541
1989	51.105	76.364	1999	76.077	111.083
1990	57.803	83.371	2000	80.774	119.147
1991	59.366	84.271	2001	73.356	110.267
1992	61.282	88.893	2002	74.894	118.923
1993	66.545	96.391	2003	79.862	125.778
1994	62.962	91.600	2004	87.897	137.110
1995	66.011	97.729	2005	95.594	147.200
1996	71.614	104.940	2006	100.584	156.249

Chow Sınamasının Kukla Değişken Almaşığı

Türkiye'deki 1994 krizini anımsayalım. Verileri 1994 öncesi ve sonrası olarak ikiye ayıralım ve şu iki modeli inceleyelim:

$$1987-1993 \text{ dönemi: } Y_t = \lambda_1 + \lambda_2 X_t + u_{1t}, \quad n_1 = 7$$

$$1994-2006 \text{ dönemi: } Y_t = \gamma_1 + \gamma_2 X_t + u_{2t}, \quad n_2 = 13$$

Yukarıdaki iki model dört farklı olasılık sunmaktadır:

- 1 Eğer $\lambda_1 = \gamma_1$ ve $\lambda_2 = \gamma_2$ ise, iki bağlanım sabit terim ve eğim olarak aynıdır: “Çakışan” (coincident) bağlanımlar.
- 2 Eğer $\lambda_1 \neq \gamma_1$ ve $\lambda_2 = \gamma_2$ ise, iki bağlanım yalnızca sabit terimler yönünden farklıdır: “Koşut” (parallel) bağlanımlar.
- 3 Eğer $\lambda_1 = \gamma_1$ ve $\lambda_2 \neq \gamma_2$ ise, iki bağlanım aynı sabit terimli ama farklı eğimlidir: “Uyumlu” (concurrent) bağlanımlar.
- 4 Eğer $\lambda_1 \neq \gamma_1$ ve $\lambda_2 \neq \gamma_2$ ise, iki bağlanım bütünüyle farklıdır: “Benzemez” (dissimilar) bağlanımlar.

Chow Sınamasının Kukla Değişken Almaşığı

- Elimizdeki iki modeli karşılaştırabilmek için tüm n_1 ve n_2 gözlemlerini toplayıp aşağıdaki bağlanımı tahmin edelim:

$$Y_t = \alpha_1 + \alpha_2 D_t + \beta_1 X_t + \beta_2 (D_t X_t) + u_t$$

- $E(u_t) = 0$ varsayımı ile şu iki bağlanımı buluruz:

$$E(Y_t | D_t = 0, X_t) = \alpha_1 + \beta_1 X_t$$

$$E(Y_t | D_t = 1, X_t) = (\alpha_1 + \alpha_2) + (\beta_1 + \beta_2) X_t$$

- Y_t ve X_t farklı yıllar için tüketim ve geliri göstermektedir.
- $D_t = 0$ 1994 öncesi, $D_t = 1$ ise 1994 ve sonrası dönemdir.
- α_2 sabit terim farkıdır.
- β_2 ise eğim katsayısı farkı olup, ikinci dönem işlevinin eğim katsayısının ilk ya da temel döneme ait eğim katsayısından ne kadar farklı olduğunu gösterir.

Chow Sınamasının Kukla Değişken Almaşığı

- Model tahmini şu sonuçları vermektedir:

$$\hat{Y}_t = -4,7884 + 16,2163 D_t + 0,7455 X_t - 0,1796 D_t X_t$$

öh	(6,9547)	(7,5961)	(0,0836)	(0,0874)
t	(-0,6885)	(2,1348)	(8,9146)	(-2,0556)
R^2	= 0,9887			

- Buna göre 1987-94 dönemi tasarruf-gelir bağlanımı şudur:

$$\hat{Y}_t = -4,7884 + 0,7455 X_t$$

- 1994-2006 dönemi tasarruf-gelir bağlanımı ise şöyledir:

$$\hat{Y}_t = (-4,7884 + 16,2163) + (0,7455 - 0,1796) X_t$$

$$= 11,4279 + 0,5659 X_t$$

- Sabit terim farkı ve eğim farkının her ikisinin de istatistiksel olarak anlamlı bulunması, bu iki bağlanımın “benzemez” olduğunu göstermektedir.

Chow Sınamasının Kukla Deęişken Almaşıđı

Kukla deęişken yönteminin Chow sınamasına üstünlükleri şunlardır:

- 1 Kukla deęişken yaklaşımı, tek bir bağlanım tahmini içerdiği için uygulama yönünden basittir.
- 2 Kukla deęişkenler, iki bağlanımın farklı olup olmadığının yanı sıra farkın sabit terimden mi yoksa eğimden mi kaynaklandığını da göstermektedir.
- 3 Tek bağlanım olması önsav sınamalarında kolaylık sağlar.
- 4 Verilerin bir arada kullanılması serbestlik derecesini arttırır.
Dikkat: Modele eklenen her kukla deęişkenin serbestlik derecesini bir azalttığı unutulmamalıdır.

Karşılıklı Etkileşim

Kukla değişkenlerin bir diğer kullanım alanı da açıklayıcı değişkenler arası karşılıklı etkileşimi incelemektir.

Ankara örneğimize dönelim ve şimdi de şu modeli ele alalım:

$$Y_i = \alpha_1 + \alpha_2 D_{2i} + \alpha_3 D_{3i} + \beta X_i + u_i$$

- Y_i burada evin fiyatını, X_i ise m^2 alanını göstermektedir.
- $D_{2i} = 1$, kot daire ise; $D_{2i} = 0$, eğer değilse.
- $D_{3i} = 1$, su deposu bulunuyorsa; $D_{3i} = 0$, eğer yoksa.
- Model tahmin sonuçları aşağıdaki gibidir.

$$\hat{Y}_i = 1,2103 - 46,2989 D_{2i} + 20,4479 D_{3i} + 1,2023 X_i$$

öh	(19,0367)	(12,8606)	(9,7909)	(0,1633)
t	(0,0636)	(-3,6001)	(2,0885)	(7,3639)
R^2	= 0,5942			

- Bulgular kot dairelerin yaklaşık 46 bin lira ucuz olduğunu, apartmanda su deposu bulunmasının ise ortalama daire fiyatını yaklaşık 20 bin TL yükselttiğini göstermektedir.

Karşılıklı Etkileşim

- Tahmin etmiş olduğumuz modeldeki üstü kapalı varsayım, D_2 ve D_3 'ün fark etkilerinin birbirinden bağımsız olduğudur.
- Diğer bir deyişle, su deposu olsa da olmasa da kot dairenin fark etkisinin aynı olduğu kabul edilmektedir.
- Belli bir uygulamada bu varsayım savunulamayabilir.
- D_2 ve D_3 gibi iki ayrı nitel değişken arasında var olabilecek karşılıklı etkileşim şu şekilde ele alınır:

$$\hat{Y}_i = \alpha_1 + \alpha_2 D_{2i} + \alpha_3 D_{3i} + \alpha_4 (D_{2i} D_{3i}) + \beta X_i$$

- Burada
 - α_2 kot dairenin fark etkisini,
 - α_3 su deposu bulunmasının fark etkisini,
 - α_4 kot daire ve su deposu olmasının birlikte fark etkisini göstermektedir.

Karşılıklı Etkileşim

- Karşılıklı etkileşimi öneren model tahminleri şöyledir:

$$\hat{Y}_i = 1,1340 - 44,7608 D_{2i} + 21,1225 D_{3i} - 4,3179 (D_{2i}D_{3i}) + 1,2014 X_i$$

öh	(19,2074)	(16,1315)	(10,7339)	(26,9206)	(0,1648)
<i>t</i>	(0,0590)	(-2,7747)	(1,9678)	(-0,1604)	(7,2912)
$R^2 =$	0,5944				

- “Etkileşim kuklası” (interaction dummy) α_4 'ün istatistiksel olarak anlamlı olup olmadığı yine *t* sınamasıyla bulunabilir.
- Sonuçlar, bir apartmanda su deposu bulunmasının kot daire fiyatlarını da diğer daireler ile aynı şekilde artırdığını göstermektedir.

Parça-Yollu Doğrusal Bağlanım

- Kukla değişkenlerin bir diğer kullanım alanı da “**parça-yollu bağlanım**” (piecewise regression) modelleridir.
- Bu modellere yönelik olarak, Ankara’daki satılık daireler örneğimizdeki fiyat-metrekare ilişkisini göz önüne alalım.
- Daire fiyatlarının “**eşik**” (threshold) düzeyi denilen bir X^* değeri öncesinde ve sonrasında farklı şekilde değiştiğini varsayalım.
- Örnek olarak, daire fiyatları metrekareye göre doğrusal olarak artsın ancak X^* eşik düzeyinden sonra daha dik bir eğimle artıyor olsun.
- Buna göre, elimizdeki model iki farklı parçadan oluşan bir doğrusal bağlanım modelidir.
- Bu tür modeller daha genel bir tür olan “**kama işlevleri**” (spline functions) yaklaşımına bir örnektir.

Parça-Yollu Doğrusal Bağlanım

- Parça-yollu bağlanımı açıklamak için şu modele bakalım:

$$Y_i = \alpha_1 + \beta_1 X_i + \beta_2 (X_i - X^*) D_i + u_i$$

- Y_i burada dairenin fiyatını, X_i de metrekare genişliğini göstermektedir.
- X^* değeri genişliğin eşik düzeyidir ve önceden bellidir.
- $D_i = 1$, eğer $X_i \geq X^*$ ise;
 $D_i = 0$, eğer $X_i < X^*$ ise.
- $E(u_i) = 0$ varsayımı altında şunu görebiliriz:

Eşiğe kadar: $E(Y_i | D_i = 0, X_i, X^*) = \alpha_1 + \beta_1 X_i$

Eşik sonrası: $E(Y_i | D_i = 1, X_i, X^*) = \alpha_1 - \beta_2 X^* + (\beta_1 + \beta_2) X_i$

- Buna göre β_1 parça-yollu bağlanımın birinci parçasının, $(\beta_1 + \beta_2)$ ise ikinci parçasının eğimini vermektedir.
- Kırılma yoktur diyen önsav için $\hat{\beta}_2$ 'nin p değerine bakılır.

Parça-Yollu Doğrusal Bağlanım

- Verilerden, $X^* = 120m^2$ sonrasında fiyatların değişiyor olabileceğini çıkardığımızı varsayalım.
- Fiyat (Y) ve genişlik (X) verilerini bir parça-yollu doğrusal bağlanım modeline yakıştırırsak şu bulguları elde ederiz:

$$\hat{Y}_i = -0,4698 + 1,1967 X_i + 0,2490 (X_i - X^*)D_i$$

öh	(33,2861)	(0,3203)	(0,5660)	
t	(-0,0141)	(3,7365)	(0,4400)	$R^2 = 0,4822$

- Dairelerin metrekare fiyatı yaklaşık 1200 TL kadardır.
- 120 metrekare üstünde fiyat (1200 + 250) olmakla birlikte, aradaki fark istatistiksel olarak anlamlı değildir.
- Öyleyse $(X - X^*)D$ değişkeni modelden çıkartılabilir.

Önümüzdeki Dersin Konusu

Önümüzdeki ders

Kukla deęişkenlere ilişkin konular