Nitel Tepki Bağlanım Modelleri

Nitel Tepki ve Doğrusal Olasılık Modeli

TÜRKİYE BİLİMLER AKADEMİSİ

UADMK ULUSAL AÇIK DERS MALZEMELERİ KONSORSİYUMU

Ekonometri 2 – Konu 17

Sürüm 2,0 (Ekim 2011)

UADMK Açık Lisans Bilgisi

A. Talha Yalta
TOBB Ekonomi ve Teknoloji Üniversitesi
Ekim 2011

Ders Planı

1. Nitel Tepki ve Doğrusal Olasılık Modeli
   - Nitel Bağımlı Değişkenler
   - Doğrusal Olasılık Modeli
   - DOM Tahminindeki Güçlükler

Nitel Bağımlı Değişkenler

- Daha önceki bölümlerde açıklayıcı değişken olarak nicel ya da nitel değişkenler kullanılabileceğini görmüştük.
- Bağımlı değişken ise bir nicel değişken olmak zorunda idi.
- Bu bölümde, bağımlı değişkeni sınırlı değerler alan, örnek olarak bir kukla değişken olabilen modelleri ele alacağız.
- Bu tür modellere özgü bazı tahmin sorunlarını incelerken, almaşık nitel tepki modellerini de tanıtacağız.

Nitel Bağımlı Değişkenler

Bağımlı değişkenin 0 ve 1 gibi yalnızca iki değer alabileceği modellere ilişkin olarak şu örnek konu başlıkları gösterilebilir:

- Ev sahibi olup olmamayı belirleyen etmenler
- Sendika üyesi olup olmamanın nedenleri
- Bir kredi başvurusunun reddedilip reddedilmeyeceği
- Bir seçimde evet ya da hayır oyunu nelerin belirlediği
- İşgücüne katılıp katılmamanın nelerden etkilendiği
- Kişilerin sigorta yaptırıp yaptırmayacakları
- Şirketlerin hisse senedi çıkartıp çıkartmayacakları
- Şirketlerin ele geçirilmeye hedef olup olmayacakları
- Bir ülkede idam cezasının olup olmayacağı

Doğrusal Olasılık Modeli

- Nitel bağımlı değişkene örnek olarak şu modeli ele alalım.

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

- Burada X hanehalkının gelirini göstermektedir.
- $Y = 1$, aile ev sahibi ise; $Y = 0$, eğer değilse.
- Bir kukla değişken olan Y'yi X açıklayıcı değişken(ler)inin doğrusal işlevi olarak belirten yukarıdaki gibi modellere "doğrusal olasılık modeli" (linear probability model), ya da kısaca "DOM" (LPM) adı verilir.
- Bu modellerde X_i veriliyken Y_i'nin koşullu beklenen değeri, olayın gerçekleşme koşullu olasılığı olarak yorumlanabilir.

Doğrusal Olasılık Modeli

- Eldeki modele doğrusal olasılık denilme nedenini görmek için $E(u_i) = 0$ varsayımını anımsayalım ve şunu yazalım:
$$E(Y_i|X_i) = \beta_1 + \beta_2 X_i$$
- $Y_i = 1$ olduğunda olayın gerçekleştiğini ve bunun olasılık değerinin de P_i olduğunu söyleyelim. Bu durumda, Y'nin olasılık dağılımı şöyledir:

Y	Olasılık
0	$1 - P_i$
1	P_i
Toplam	1

- Beklenen değer tanımından yararlanarak şunu görebiliriz:
$$E(Y_i) = 0(1 - P_i) + 1(P_i) = P_i$$
- P_i bir olasılık olduğu için burada $0 \leq E(Y_i|X_i) \leq 1$ şeklinde bir sınırlama olduğu unutulmamalıdır.

DOM Tahminindeki Güçlükler

Yukarıda gördüklerimiz, SEK yönteminin kolaylıkla nitel bağımlı değişkenler için de kullanılabileceği kanısını uyandırıyorsa da durum gerçekte böyle değildir.

DOM tahmini, aşağıda gösterilen dört sorunu da beraberinde getirmektedir.

1. Bozukluk terimi u'nun normal-dışılığı
2. Bozukluklarda farklıserpilimsellik görülmesi
3. R^2'nin yakışma ölçütü olarak kuşkulu değeri
4. $0 \leq E(Y_i|X_i) \leq 1$ koşulunun sağlanamaması

Bozukluk Terimi *u*'nun Normal-dışılığı

- DOM tahminininde u_i'lerin normal dağılması olanaksızdır.
- Aşağıda da görüldüğü gibi, Y_i'ler yalnızca 2 değer aldıkları için, 2 farklı u_i kümesi ortaya çıkar.

$$u_i = Y_i - \beta_1 - \beta_2 X_i$$

$Y_i = 1$	ise	$u_i = 1 - \beta_1 - \beta_2 X_i$
$Y_i = 0$	ise	$u_i = -\beta_1 - \beta_2 X_i$

- Bu durumda u_i'ler normal dağılımı değil, kesikli Bernouilli dağılımını izlerler.
- Diğer taraftan, nokta tahminleri yansız olmayı sürdürürler.
- Ayrıca merkezi limit kanıtsavına göre örneklem büyüklüğü artarken kalıntıların normale yaklaşacağı unutulmamalıdır.
- Dolayısıyla, büyük örneklemlerde *u*'nun normal-dışılığı tahmin ve çıkarsama açısından bir sorun yaratmayabilir.

Bozukluklarda Farklıserpilimsellik

- Bozuklukların Bernouilli dağılımına uyduğu bilindiğine göre, DOM tahmininde u_i'lerin aynıserpilimsel olduğu varsayımını korumak da olanaksızdır.
- Anımsayacak olursak, ikiterimli dağılımın genel biçimi olan Bernoulli dağılımının ortalaması p, varyansı $p(1-p)$'dir.
- Buna göre, doğrusal olasılık modelinin varyansı da şu olur.

$$\mathrm{var}(u_i) = P_i(1 - P_i)$$

- $P_i = E(Y_i|X_i) = \beta_1 + \beta_2 X_i$ olduğuna göre, u_i sonuçta X_i değerlerine bağlıdır ve bu nedenle aynıserpilimsel olamaz.
- Farklıserpilimsellik altında SEK tahminlerinin yansız olmayı sürdürürken enaz varyanslı olamadıklarını anımsayalım.
- Büyük örneklemlerde bu da DOM için sorun olmayabilir. Farklıserpilimsellik ağırlıklı en küçük kareler ya da White ölçünlü hataları kullanılarak aşılmaya çalışılabilir.

ULUSAL AÇIK DERS MALZEMELERİ KONSORSİYUMU

R^2'nin Yakışma Ölçütü Olarak Kuşkulu Değeri

- Nitel bağımlı değişkenler, tanım gereği, gözlenen Y_i'lerin yalnızca kesikli 0 ve 1 değerlerini alabilmeleri demektir.
- DOM tahmininden elde edilen $\hat{Y}_i$'ler ise yakıştırılan doğru üzerinde farklı ve sürekli değerler alabilirler.
- DOM'ların böyle bir serpilime iyi yakışması beklenemez.
- Bu nedenle, DOM tahmininde R^2 genellikle düşük çıkar. Uygulamada genellikle 0.2 ile 0.6 arası değerler beklenir.
- Genel olarak, her türden nitel bağımlı değişken modelinde belirleme katsayısı R^2'yi bir özet istatistik olarak kullanmak sakıncalı kabul edilmektedir.

$0 \leq E(Y_i|X_i) \leq 1$ Koşulunun Sağlanamaması

- DOM tahminine ilişkin asıl ciddi bir sorun, $0 \leq E(Y_i|X_i) \leq 1$ koşulunun sağlanamamasıdır.
- Bu modeller X veriliyken Y olayının gerçekleşme koşullu olasılığını ölçtüğü için, $E(Y_i|X_i)$ değerinin 0 ile 1 arasında yer alması önemlidir.
- SEK yöntemi böyle bir matematiksel sınırlama içermediği için, DOM tahmini sonrasında yakıştırılan değerlerin eksi değerli ya da 1'den büyük çıkmasına sıkça rastlanır.
- Böyle durumlarda eksi değerli $\hat{Y}_i$'leri sıfır, 1'den büyük $\hat{Y}_i$'leri ise 1 varsaymak yoluna gidilebilir.
- Ancak böyle sakıncalı ek varsayımlara gerek yoktur çünkü tahmin edilen olasılıkların 0 ve 1 arasında olmasını güven altına alan almaşık yöntemler bulunmaktadır.

DOM Açıklayıcı Örnek

- Açıklayıcı örnek olarak, çoğu ilde tek bir "ticaret ve sanayi odası" varken, bazı büyük illerde ise sanayi odasının ayrı bir kuruluş olarak hizmet verdiğini göz önüne alalım.
- Bir ilde sanayi odası olup olmayacağını sanayi sektöründe faaliyet gösteren firma sayısı ile açıklamak istiyor olalım:

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

- X_i burada toplam firma sayısını (100 birim) göstermektedir.
- $Y_i = 1$, ilde sanayi odası var ise; $Y_i = 0$, eğer yok ise.
- Önsel beklentimiz, $\hat{\beta}_2$'nın artı değerli ve (0,1) aralığında tahmin edileceği yönündedir.

DOM Açıklayıcı Örnek

- TOBB Sanayi Veri Tabanı verilerine dayalı DOM bağlanım sonuçları aşağıdaki gibidir.

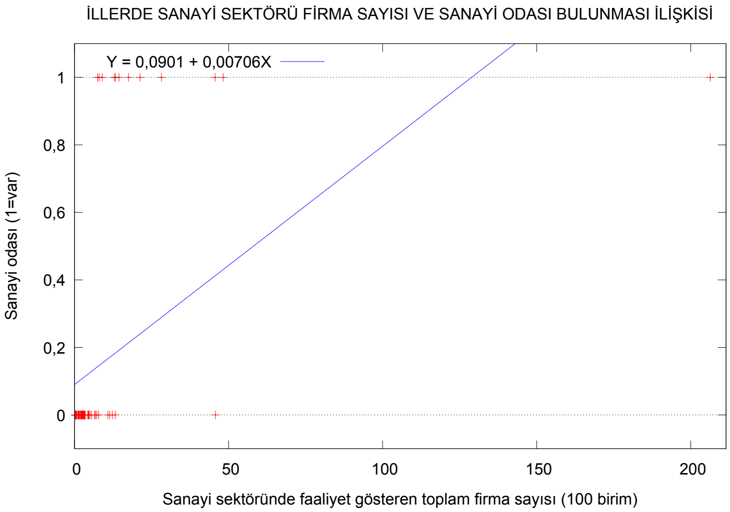
$$\begin{array}{llll} \hat{Y}_i = & 0{,}0901 & +\,0{,}0070\ X_i & \\ \text{öh} & (0{,}0376) & (0{,}0015) & \\ t & (2{,}3933) & (4{,}8092) & r^2 = 0{,}2287 \end{array}$$

- Sonuçlar, ilde açılacak her 100 yeni firmanın sanayi odası kurulma olasılığını yüzde 0,7 artıracağını göstermektedir.
- İlişki doğrusal tahmin edildiği için, ilde firma sayısı 1000 de olsa 5000 de olsa olasılığın aynı kaldığı varsayılmaktadır.
- Kukla bağımlı değişkene bir doğru yakıştırmak güç olduğu için, r^2 değeri beklenildiği gibi düşük bulunmuştur.
- Ayrıca, toplam 20601 sanayi firması bulunan İstanbul için tahmin edilen olasılık 1'den büyük çıkmaktadır:

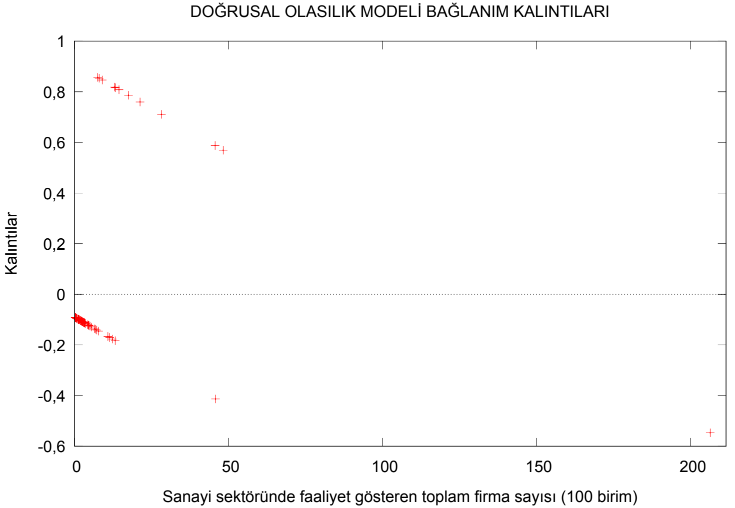
$$0{,}0901 + (0{,}0070 \times 206{,}01) = 1{,}547$$

- Y_i'ye bağlı iki ayrı kalıntı kümesi olduğundan, hataların normal dağılması da söz konusu değildir.

DOM Yakıştırması



DOM Kalıntıları



Önümüzdeki Dersin Konusu

Önümüzdeki ders

Doğrusal-dışı yaklaşım ve olabirim modeli