

Ekonometrik Modelleme

Belirtim Hatalarının Niteliği




Ekonometri 2 – Konu 14
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 Belirtim Hatalarının Niteliđi
 - Belirtim Hatası Trleri ve Bunların Sonuđları

Model Belirtim Konusu

KDBM'nin 9. varsayımı, kullanılan modelin “dođru” belirtilmiř olduđudur. Bu varsayım altında řu ana kadar katsayı tahmini ve buna iliřkin sınamalar üzerine odaklanılmıřtı.

Ancak, eđer model dođru belirtilmediyse “**model belirtim hatası**” (model specification error) ya da “**model belirtim yanlılıđı**” (model specification bias) sorunu ortaya çıkar.

Bu bölümde řu sorulara yanıt arayacađız:

- 1 Uygulamada karřılařılan belirtim hataları nelerdir?
- 2 Bu hatalar hangi sonuçları dođurur?
- 3 Belirtim hataları nasıl saptanabilir?
- 4 Düzeltmek için ne gibi önlemler alınabilir?
- 5 Almařık modeller arasında nasıl seçim yapılır?

Model Belirtim Konusu

Model belirtimi konusu, uzmanlar arasında zaman zaman bakış ayrılıkları da olabilen geniş bir alandır. Ancak yaygın görüşe göre çözümlenmede kullanılacak model şu özellikleri taşımalıdır.

- 1 **Onanırılık:** Model çıkarımlarının kabul edilebilir olması.
- 2 **Kuram ile uyumluluk:** Modelin iktisat düşüncesi açısından anlamlı olması.
- 3 **Açıklayıcı değişken dıştüreliđi:** Bağlayanların hata terimi ile ilintisiz olması.
- 4 **Deđiştirge deđişmezliđi:** Deđiştirge tahminlerinin farklı örneklerde deđişmemesi.
- 5 **Veriler ile bağdaşma:** Kalıntıların tümüyle rastsallık, diđer bir deyişle “beyaz gürültü” (white noise) özelliđi göstermesi.
- 6 **Kapsayıcılık:** Modelin açıklama gücü bakımından alması modeller içinde en iyisi olması.

Belirtim Hatası Türleri

Bir modelin yukarıda sözü edilen özellikleri kaybetmesine yol açabilecek dört önemli hata türü şunlardır:

- “Atlanan deđişken hatası” (omitted variable error)
- “İlgisiz deđişken hatası” (irrelevant variable error)
- “Yanlış işlev biçimi” (wrong functional form)
- “Ölçüm hataları yanlılığı” (measurement errors bias)

Şimdi bu sorunları ve neden oldukları olumsuz sonuçları kısaca ele alalım.

Modeli Eksik Belirtme

- Atlanan değişken hatasını açıklamak için, aşağıdaki üç değişkenli modelin “doğru” olduğunu varsayalım:

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

- Bunun yerine ise aşağıdaki “**eksik belirtilmiş model**” (under specified model) kullanılsın:

$$Y_i = \alpha_1 + \alpha_2 X_{2i} + v_i$$

- X_3 'ün X_2 'ye göre ikili bağlanımındaki eğim katsayısı b_{32} olsun. Bu durumda şu eşitliğin geçerli olduğu gösterilebilir:

$$E(\hat{\alpha}_2) = \beta_2 + \beta_3 b_{32}$$

- Eşitlik gösteriyor ki α_2 , β_2 'nin yanlış bir tahmincisidir.
- Örnek olarak, X_3 'ün Y üzerindeki etkisi (β_3) ile X_3 'ün X_2 üzerindeki etkisi (b_{32}) aynı anda artı değerli ise, $\hat{\alpha}_2$ yukarı doğru yanlış olacak ve gerçek β_2 'den hep yüksek çıkacaktır.

Modeli Eksik Belirtme

- Şimdi de $\hat{\alpha}_2$ 'nin ve $\hat{\beta}_2$ 'nin varyanslarını karşılaştıralım:

$$\text{var}(\hat{\alpha}_2) = \frac{\sigma^2}{\sum x_{2i}^2} \quad \text{var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_{2i}^2(1 - r_{23}^2)}$$

- $\hat{\beta}_2$, her ne kadar yansız olsa da, daha büyük varyanslıdır. X_2 ve X_3 arasındaki eşdoğrusallığın göstergesi olan ilinti katsayısının karesi arttıkça, aradaki fark da artmaktadır.
- Anlaşıyor ki yanlılık ve varyans arasında bir “ödünleşim” (trade off) bulunmaktadır.
- Bu durumda yüksek eşdoğrusallık altında X_3 'ü modelden çıkartıp, yanlı olsa da, $\hat{\beta}_2$ yerine $\hat{\alpha}_2$ kullanmak yeğlenebilir.
- Diğer yandan, iktisat kuramına dayanarak oluşturulan bir modelden değişken çıkartmanın zorunlu kalmadıkça asla önerilmediği unutulmamalıdır.

Modeli Eksik Belirtme

Özetle, modelde bulunması gereken X_3 deđişkenini atlamak şu sonuçları doğurmaktadır:

- 1 Hatalı modeldeki sabit terim mutlaka yanlıdır ve tutarsızdır. Diđer bir deyişle, örneklem büyüdükçe yanlılık yok olmaz.
- 2 Hatalı modeldeki diđer $\alpha_2, \alpha_3, \dots$ katsayıları da yanlıdır.
- 3 Eđer X_3 ile atlanılmayan bir deđişken arasındaki eğim sıfır ise (örneđimizdeki $b_{32} = 0$ durumu), o zaman katsayı yanlı olmaz. Ancak uygulamada bu durum neredeyse hiç yoktur.
- 4 Yanlı katsayı tahminlerinden dolayı alışıldık güven aralıkları ve önsav sınav sonuçları yanıltıcı olabilir.
- 5 Hatalı modelde varyanslar genellikle daha küçüktür. Ancak yanlılık sorunu olduđu için, hatalı modelin yeđlenebilmesi eşdođrusallığın çok yüksek olduđu durumlar ile sınırlıdır.

Modeli Aşırı Belirtme

- İlgisiz değişken hatasınının sonuçlarını gösterebilmek için, şimdi de doğru modelin şu olduğunu varsayalım:

$$Y_i = \beta_1 + \beta_2 X_{2i} + u_i$$

- Araştırmacı ise aşağıdaki “**aşırı belirtimli**” (over specified) modeli kullanmakta diretiyor olsun:

$$Y_i = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + v_i$$

İlgisiz değişken eklemenin katsayılar üzerindeki etkisi şöyledir:

- Hatalı modeldeki tüm katsayı tahminleri yansız ve tutarlıdır.
- Bu nedenle alışıldık güven aralıkları ve önsav sınamaları geçerlidir.
- Diğer taraftan katsayılar etkin değildir. Diğer bir deyişle, varyanslar doğru modeldekilerden daha büyüktür.

Modeli Aşırı Belirtme

- Aşırı belirtilmiş modeldeki $\hat{\alpha}_2$ tahmininin etkin olmadığını, varyansları karşılaştırarak görebiliriz:

$$\text{var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_{2i}^2} \quad \text{var}(\hat{\alpha}_2) = \frac{\sigma^2}{\sum x_{2i}^2 (1 - r_{23}^2)}$$

- Doğru modelde X_3 olmadığı için paydada $(1 - r_{23}^2)$ teriminin yer almadığına dikkat ediniz.
- Hatalı modelde ise $\hat{\alpha}_2$ varyansı görece yüksek çıkacaktır.
- Aradaki fark X_3 ile X_2 arasındaki ilinti katsayısının karesi ile doğru orantılıdır.
- Demek ki ilgisiz bir değişken eklemek tahmin sonuçlarının kesinliğini azaltmak gibi ciddi bir sonuca yol açabilmektedir.
- Ayrıca, bilimde en az karmaşık açıklama yeğlendiği için, Model belirtiminde **“tutumluluk ilkesi”** (parsimony principle) her zaman özen gösterilmesi gereken önemli bir konudur.

Ölçüm Hataları

- Ölçüm hataları bir model belirtim hatası değildir. Ancak doğurabileceđi sonuçlar ekonometrik modellemede ölçüm hatalarını da dikkate almayı gerekli kılar.
- Şimdiye kadar olan varsayımımızın aksine, çözümlemede kullandığımız veriler
 - “kaydedici hatası” (clerical error),
 - “atanan değerler” (assigned values),
 - “yuvarlama” (rounding),
 - “içdeğerleme” (interpolation),
 - “dışdeğerleme” (extrapolation)gibi nedenlerden dolayı kesin doğru olmayabilir.
- İkincil kaynaklar tarafından yayınlanan verilerde yer alan hataları bilmek oldukça güçtür. Çođu çalışma böyle verilere dayandığı için, uygulamada bu hata ile sıkça karşılaşılır.

Bağımlı Deđişkendeki Ölçüm Hataları

- Bağımlı deđişkendeki ve açıklayıcı deđişkenlerdeki ölçüm hatalarının etkileri farklıdır. Öncelikle řu modele bakalım:

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

- Burada Y Friedman tarafında öne sürülen “**kalıcı tüketim**” (permanent consumption) harcamasını, X ise cari geliri göstermektedir.
- Gerçekte kavramsal bir araç olan kalıcı tüketim doğrudan ölçülemediđi için, elimizde, gözlenebilen tüketime dayalı řu deđişken vardır:

$$Y_i^* = Y_i + v_i$$

- Yukarıdaki v_i , Y^* 'daki ölçüm hatalarını gösteren rastlantısal bir terimdir.

Bağımlı Değişkendeki Ölçüm Hataları

- Y yerine Y^* kullanıldığında tahmin edilen model şu olur:

$$Y_i^* = \beta_1 + \beta_2 X_i + \epsilon_i$$

- Görülüyor ki yukarıdaki bağlanımda katsayılar aynı ve doğru şekilde tahmin edilebilmektedir.
- Diğer taraftan, $\epsilon_i = u_i - v_i$ biçimindeki bileşik hata teriminin varyansı daha yüksektir:

$$\text{var}(u_i - v_i) = \text{var}(u_i) + \text{var}(v_i) + 2\text{cov}(u_i, v_i)$$

- Öyleyse bağımlı değişkendeki ölçüm hataları katsayı nokta tahminlerini etkilememekte ancak güven aralıklarının geniş olmasına yol açarak etkinliği azaltmaktadır.

Açıklayıcı Değişkenlerdeki Ölçüm Hataları

- Açıklayıcı değişkende yer alan ölçüm hatalarına yönelik olarak, şimdi de şu modeli ele alalım:

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

- Bu modelde Y cari tüketim, X ise “kalıcı gelir” (permanent income) olarak tanımlanmıştır.
- Kalıcı gelir de doğrudan ölçülemediği için, uygulamada gözlenebilen gelire dayalı bir değişken tanımı kullanılır:

$$X_i^* = X_i + w_i$$

- Burada w_i , X_i^* 'deki ölçüm hatasını göstermektedir.

Açıklayıcı Deđişkendeki Ölçüm Hataları

- X yerine X^* kullanılması aşağıdaki modele yol açar:

$$\begin{aligned} Y_i &= \beta_1 + \beta_2(X_i + w_i) + u_i \\ Y_i &= \beta_1^* + \beta_2^* X_i + z_i \end{aligned}$$

- Buradaki bileşik hata terimi $z_i = u_i + \beta_2 w_i$ biçimindedir.
- İçerdiği β_2 teriminden dolayı, z_i , KDBM'nin hata terimi ve açıklayıcı deđişkenlerin ilişkisiz olduđu varsayımını çığner.
- Öyleyse açıklayıcı deđişkenlerdeki ölçüm hataları ciddi bir sorundur, çünkü yukarıdaki durumda SEK tahminleri hem yanlış hem de tutarsızdır.
- Bu yanlışlık sorununu gidermek zordur. Başvurulabilecek bir yol “**araç deđişkenler**” (instrumental variables) yöntemidir.
- Eğer ölçüm hataları küçükse, ki bunu bilebilmek güçtür, uygulamada sorunu gözardı etmek zorunda kalınabilir.
- En dođru yol hatasız, dođru ölçülmüş verilerle çalışmaktır.

Önümüzdeki Dersin Konusu

Önümüzdeki ders

Belirtim hatalarının sınanması