Çoklu Bağlanım Çözümlemesi

Çıkarsama Sorunu

Yrd. Doç. Dr. A. Talha YALTA
Ekonometri 1 Ders Notları
Sürüm 2,0 (Ekim 2011)

TÜRKİYE BİLİMLER AKADEMİSİ

ULUSAL AÇIK DERS MALZEMELERİ KONSORSİYUMU

Açık Lisans Bilgisi

A. Talha Yalta
TOBB Ekonomi ve Teknoloji Üniversitesi
Ekim 2011

Ders Planı

1. *T* Sınamaları
   - Çoklu Bağlanımda Önsav Sınaması
   - Tek Bir Katsayının Sınanması
   - İki Katsayının Eşitliğinin Sınanması
2. *F* Sınamaları
   - Bağlanımın Bütününün Anlamlılık Sınaması
   - Bir Açıklayıcı Değişkenin Marjinal Katkısı
   - Sınırlamalı Enküçük Kareler Yöntemi
3. Diğer Sınama ve Konular
   - Chow Sınaması
   - MWD Sınaması
   - Diğer Bazı Sınama ve Konular

Ders Planı

Çoklu Bağlanımda Önsav Sınaması

- Bu bölümde daha önce iki değişkenli bağlanım modelleri için ele almış olduğumuz aralık tahmini ve önsav sınaması kavramlarını çok değişkenli modellere genişleteceğiz.
- Bilindiği gibi amacımız yalnızca bağlanım katsayılarını tahmin etmek değil, aynı zamanda bu katsayılara ilişkin çeşitli çıkarsamalar ve önsav sınamaları da yapmaktır.
- Bu doğrultuda u_i hatalarının sıfır ortalama ve σ^2 sabit varyanslı normal dağılıma uydukları varsayımını çoklu bağlanım modelleri için de sürdüreceğiz.

Çoklu Bağlanımda Önsav Sınaması

İkili bağlanım modelinin basit dünyasından dışarı çıkıldığında önsav sınaması aşağıdaki gibi farklı şekiller almaktadır:

1. Tek bir kısmi bağlanım katsayısına ilişkin önsav sınaması,
2. Tahmin edilen bağlanım modelinin bütününün sınanması,
3. İki ya da daha çok katsayının eşitliğinin sınanması,
4. Katsayıların belli sınırlamalara uygunluğunun sınanması,
5. Modelin farklı veri setlerindeki kararlılığının sınanması,
6. Bağlanım modellerinin işlev biçimlerinin sınanması.

İzleyen bölümde bu sınama çeşitleri ayrı ayrı ele alınacaktır.

Çoklu Bağlanımda Önsav Sınaması

- Farklı önsav sınama biçimlerini göstermek için, Türkiye'de 81 ile ait ve 2000 yılı verileri kullanılarak tahmin edilmiş şu modeli ele alalım:

$$\begin{array}{llllll} \hat{Y}_i = & 7{,}3778 & +1{,}4718\,X_{2i} & -0{,}2014\,X_{3i} & & \\ \text{öh} & (1{,}0689) & (0{,}3850) & (0{,}0717) & R^2 = 0{,}3139 \\ t & (6{,}9021) & (3{,}8223) & (-2{,}8078) & \bar{R}^2 = 0{,}2963 \end{array}$$

- Burada

  Y ilin aldığı göçün toplam il nüfusuna oranını (%),

  X_2 cari fiyatlarla kişi başına düşen GSYH'yi (1000 TL),

  X_3 erkek nüfustaki işsizlik oranını (%)

  göstermektedir.
- Sonuçlara göre, milli gelirdeki 1000 liralık artış ilin göç alma yüzdesini yaklaşık 1,5 puan yükseltirken işsizlikteki benzer bir artış ise % 0,2'lik eksi yönlü bir etkiye neden olmaktadır.
- Katsayılar anlamlıdır ve önsel beklentilerle de uyumludur.

Tek Bir Katsayının Sınanması

- $u_i \sim N(0, \sigma^2)$ varsayımı altında, herhangi bir tekil bağlanım katsayısına ilişkin önsavlar için t sınamasını kullanabiliriz.
- Örnek olarak, erkek işsizlik oranının göç alma üzerinde bir doğrusal etkisi olmadığı varsayımını şöyle sınarız:

$$H_0 : \beta_3 = 0, \qquad H_1 : \beta_3 \neq 0$$

$$t = \frac{\hat{\beta}_3 - \beta_3^*}{\text{öh}(\hat{\beta}_3)} = \frac{-0{,}2014 - 0}{0{,}0717} = -2{,}8089$$

- $\alpha = 0{,}05$ seçilirse, 78 (81-3) sd ile kritik $t_{\alpha/2} = 1{,}9908$ olur.
- Hesaplanan t değeri kritik t değerini aştığı için, istatistiksel olarak β_3'ün anlamlı olduğunu ya da diğer bir deyişle sıfırdan anlamlı ölçüde uzak olduğunu söyleyebiliriz.

Tek Bir Katsayının Sınanması

- Bilindiği gibi önsav sınamasına diğer bir yaklaşım da güven aralığı yöntemidir.
- Örnek olarak β_2'nin yüzde 95 güven aralığı şöyledir:

$$\begin{aligned} \hat{\beta}_2 - t_{\alpha/2}\text{öh}(\hat{\beta}_2) &\leq \beta_2 \leq \hat{\beta}_2 + t_{\alpha/2}\text{öh}(\hat{\beta}_2) \\ 1{,}4718 - 1{,}9908(0{,}3850) &\leq \beta_2 \leq 1{,}4718 + 1{,}9908(0{,}3850) \\ 0{,}7053 &\leq \beta_2 \leq 2{,}2383 \end{aligned}$$

- 81 gözlemli 100 farklı örneklem seçilir ve $\hat{\beta}_2 \pm t_{\alpha/2}\text{öh}(\hat{\beta}_2)$ gibi böyle 100 güven aralığı bulunursa, bunlardan 95'inin anakütledeki gerçek β_2'yi içermesi beklenir.

İki Katsayının Eşitliğinin Sınanması

- Şimdi de β_2 ve β_3 eğim katsayılarının birbirine eşit olup olmadığını sınamak istediğimizi varsayalım.
- Bunun için sıfır ve almaşık önsavları iki şekilde yazabiliriz:

$$H_0 : \beta_2 = \beta_3$$
$$H_1 : \beta_2 \neq \beta_3$$

$$H_0 : (\beta_2 - \beta_3) = 0$$
$$H_1 : (\beta_2 - \beta_3) \neq 0$$

- Milli gelir ve işsizlik oranı katsayılarının eşit olmasını doğal olarak beklemiyoruz. Dolayısıyla örneğimizde bu sınama iktisat bağlamında gereksizdir.
- Diğer yandan, uygulamada bu tür önsav sınamasına sıkça başvurulur.
- Örnek olarak Y bir mala olan talebi, X_2 ve X_3 de sırasıyla tüketicinin gelir ve servetini gösteriyor olsun.
- Log-doğrusal model için yukarıdaki sıfır önsavları talebin gelir ve servet esnekliklerinin aynı olduğu anlamına gelir.

İki Katsayının Eşitliğinin Sınanması

- İki katsayı tahmininin eşitliğini sınamak için t sınaması yöntemi kullanılabilir:

$$t = \frac{(\hat{\beta}_2 - \hat{\beta}_3) - (\beta_2 - \beta_3)^*}{\text{öh}(\hat{\beta}_2 - \hat{\beta}_3)}$$

- Klasik varsayımlar altında $n - k$ sd (örneğimizde $k = 3$) ile t dağılımına uyan yukarıdaki istatistik şöyle de yazılabilir:

$$t = \frac{\hat{\beta}_2 - \hat{\beta}_3}{\sqrt{\text{var}(\hat{\beta}_2) + \text{var}(\hat{\beta}_3) - 2\text{cov}(\hat{\beta}_2, \hat{\beta}_3)}}$$

- Yukarıda, $\text{öh}(\hat{\beta}_2 - \hat{\beta}_3) = \sqrt{\text{var}(\hat{\beta}_2) + \text{var}(\hat{\beta}_3) - 2\text{cov}(\hat{\beta}_2, \hat{\beta}_3)}$ ve H_0'a göre $\beta_2 - \beta_3 = 0$ özdeşliklerinden yararlanılmıştır.

İki Katsayının Eşitliğinin Sınanması

- Üçlü bağlanım örneğimiz için $\text{cov}(\hat{\beta}_2, \hat{\beta}_3) = 0{,}0104$'tür.
- $\text{var}(\hat{\beta}_2)$ ve $\text{var}(\hat{\beta}_3)$ değerleri ise ölçünlü hataların karesi alınarak 0,1482 ve 0,0051 olarak bulunur.
- Buna göre $\beta_2 - \beta_3 = 0$ sınamasını şöyle yaparız:

$$t = \frac{1{,}4718 + 0{,}2014}{\sqrt{0{,}1482 + 0{,}0051 - 0{,}0208}} = 4{,}5966$$

- Eldeki değer 78 sd ve çift kuyruklu sınama için hesaplanan $t = 1{,}9908$ kritik t değerini aştığı için, X^2 ve X^3'e ait katsayı değerlerinin aynı olduğu sıfır önsavı reddedilir.

Ders Planı

Bağlanımın Bütününün Anlamlılık Sınaması

- Türkiye örneğimize dönelim ve β_2 ve β_3'ün aynı anda sıfır olduğunu öneren $H_0 : \beta_2 = \beta_3 = 0$ önsavını ele alalım.
- Bu sıfır önsavının sınanmasına, bağlanıma ilişkin "bütünün anlamlılığı" (overall significance) sınaması adı verilir.
- Bu sınama tekil anlamlılık sınamalarından farklıdır.
- Bunun nedeni şudur: $\hat{\beta}_2$ ve $\hat{\beta}_3$ gibi farklı katsayılar için tekil anlamlılık sınaması yaparken, her bir sınamanın farklı ve bağımsız bir örnekleme dayandığı varsayılır.
- Diğer yandan, verili bir örneklemde $\text{cov}(\hat{\beta}_2, \hat{\beta}_3) = 0$ geçerli olmayabilir. Diğer bir deyişle, $\hat{\beta}_2$ ile $\hat{\beta}_3$ ilişkili olabilirler.
- Bu durumda, $\hat{\beta}_2$ ile $\hat{\beta}_3$'nın aynı anda $[\hat{\beta}_2 \pm\ t_{\alpha/2}\text{öh}(\hat{\beta}_2)]$ ve $[\hat{\beta}_3 \pm\ t_{\alpha/2}\text{öh}(\hat{\beta}_3)]$ aralıklarında bulunma olasılığı $(1-\alpha)^2$ değildir.

Bağlanımın Bütününün Anlamlılık Sınaması

- Anakütle kısmi bağlanım katsayılarının aynı anda sıfır olduğu yönündeki ortak önsavı sınamak için varyans çözümlemesi yöntemi kullanılabilir:

$$\begin{aligned} \sum y_i^2 &= \hat{\beta}_2 \sum y_i x_{2i} + \hat{\beta}_3 \sum y_i x_{3i} && + \sum \hat{u}_i^2 \\ \text{TKT} &= \text{BKT} && + \text{KKT} \end{aligned}$$

- Buna göre aşağıdaki VARÇÖZ çizelgesini düzenleyebiliriz:

Değişimin Kaynağı	KT	sd	OKT (KT/sd)
Bağlanımdan (BKT)	$\hat{\beta}_2 \sum y_i x_{2i} + \hat{\beta}_3 \sum y_i x_{3i}$	$k-1$	$\frac{\hat{\beta}_2 \sum y_i x_{2i} + \hat{\beta}_3 \sum y_i x_{3i}}{k-1}$
Kalıntılardan (KKT)	$\sum \hat{u}_i^2$	$n-k$	$\frac{\sum \hat{u}_i^2}{n-k} = \hat{\sigma}^2$
Toplamlarından (TKT)	$\sum y_i^2$	$n-1$	

- Burada k, sabit terim ile birlikte tahmin edilen toplam anakütle katsayılarının sayısıdır.

Bağlanımın Bütününün Anlamlılık Sınaması

- Üçlü model için, hata teriminin normal dağıldığı varsayımı ve $\beta_2 = \beta_3 = 0$ sıfır önsavı altında şu istatistik hesaplanır:

$$F = \frac{(\hat{\beta}_2 \sum y_i x_{2i} + \hat{\beta}_3 \sum y_i x_{3i})/(k-1)}{\sum \hat{u}_i{}^2/(n-k)} = \frac{\text{BKT}/sd}{\text{KKT}/sd}$$

- Yukarıda verilen değişkenin $(k-1)$ ve $(n-k)$ sd ile F dağılımına uyduğu gösterilebilir.
- Buna göre, hesaplanan F istatistiğinin p değeri yeterince küçükse H_0 reddedilir.

Bütünün Anlamlılık Sınaması Açıklayıcı Örnek

Bağlanımın bütününün anlamlılığının sınanmasına örnek olarak Türkiye için gelir, işsizlik, ve göç alma modelimize dönelim ve aşağıdaki VARÇÖZ çizelgesini oluşturalım:

Değişimin Kaynağı	KT	sd	OKT
Bağlanım	205,022	2	102,511
Kalıntılar	448,080	78	5,74461
Toplam	653,102	80	

- F değeri çizelgeden aşağıdaki gibi hesaplanır:

$$F = \frac{102{,}511}{5{,}74461} = 17{,}8448$$

- Yüzde 5 anlamlılık düzeyinde ve 2 ile 78 sd için kritik değer $F_{0{,}05}(2,78) = 3{,}1138$'dir.
- Hesaplanan F değeri anlamlı olduğu için H_0 reddedilir.

Çoklueşdoğrusallığın Etkisi

- Göstermiş olduğumuz yöntemle hesaplanan F istatistiği çoğu zaman yüksek çıkar.
- Tüm bağlanım katsayıları tek tek istatistiksel olarak anlamlı değilken, F değerinin anlamlı çıkması olasıdır.
- Bu durum açıklayıcı değişkenler kendi aralarında yüksek derecede ilinti gösteriyorsa karşımıza çıkabilir.
- Bu sorunu ileride çoklueşdoğrusallık başlığı altında ayrıntılı biçimde ele alacağız.
- Şimdilik F ve t sınama sonuçlarını yorumlarken dikkatli olmak gerektiğini vurgulamakla yetiniyoruz.

R^2 ve F Arasındaki İlişki

- Belirleme katsayısı R^2 ile varyans çözümlemesindeki F değeri arasında yakın bir ilişki vardır.
- k değişkenli durumda ve $H_0 : \beta_2 = \beta_3 = \ldots = \beta_k = 0$ sıfır önsavı altında şu gösterilebilir:

$$
\begin{aligned}
F &= \frac{n-k}{k-1}\frac{\text{BKT}}{\text{KKT}} \\
&= \frac{n-k}{k-1}\frac{\text{BKT}}{\text{TKT}-\text{BKT}} \\
&= \frac{n-k}{k-1}\frac{\text{BKT}/\text{TKT}}{1-(\text{BKT}/\text{TKT})} \\
&= \frac{n-k}{k-1}\frac{R^2}{1-R^2}
\end{aligned}
$$

- Burada $R^2 = \text{BKT}/\text{TKT}$ tanımı kullanılmıştır.
- Eşitliğe göre R^2 ile F aynı yönde değişirler.
- Tahmin edilen bağlanımın bütün olarak anlamlılığının ölçüsü olan F demek ki aynı zamanda $H_0 : R^2 = 0$ sınamasına eşdeğerdir.

R^2 ve F Arasındaki İlişki

- F sınamasının R^2 cinsinden gösterilmesinin üstün yanı hesaplama kolaylığıdır. Tek gereken R^2 değeridir.
- VARÇÖZ çizelgesini R^2 ile aşağıdaki gibi düzenleyebiliriz:

Değişimin Kaynağı	KT	sd	OKT
Bağlanım	$R^2(\sum y_i^2)$	$k-1$	$R^2(\sum y_i^2)/(k-1)$
Kalıntılar	$(1-R^2)(\sum y_i^2)$	$n-k$	$(1-R^2)(\sum y_i^2)/(n-k)$
Toplam	$\sum y_i^2$	$n-1$	

Bir Açıklayıcı Değişkenin Marjinal Katkısı

- Bir açıklayıcı değişkenin marjinal katkısına bakmak için, X_2 ve X_3 gibi iki değişkeni modele sırayla ekleyelim.
- Burada görmek istediğimiz, eklenen değişkenin KKT'yi eskiye oranla ne ölçüde azalttığıdır.
- Çoğu görgül çalışmada, çeşitli olası X değişkenleri içinden KKT'yi çok azaltmayanları modele eklememek yeğlenebilir.
- Aynı şekilde KKT'yi önemli ölçüde azaltan, diğer bir deyişle R^2'yi "anlamlı" biçimde yükselten değişkenler de modelden çıkartılmak istenmez.
- Bu yüzden bir X değişkeninin marjinal katkısı uygulamada önemli bir konudur.

Bir Açıklayıcı Değişkenin Marjinal Katkısı

- Ek bir açıklayıcı değişkenin KKT'yi anlamlı biçimde azaltıp azaltmadığını bulmak için yine varyans çözümlemesinden yararlanılabilir.
- Türkiye'de illerin aldığı göç örneğimize dönelim ve şu ikili bağlanımı tahmin edelim:

$$\begin{array}{llll} \hat{Y}_i = & 4{,}8832 & + 1{,}8800\, X_{2i} & \\ \text{öh} & (0{,}6197) & (0{,}3717) & \\ t & (7{,}8797) & (5{,}0574) & r^2 = 0{,}2446 \end{array}$$

- Bulgular, kişi başına düşen GSYH'yi gösteren X_2'nin Y'yi anlamlı biçimde etkilediğini göstermektedir.

Bir Açıklayıcı Değişkenin Marjinal Katkısı

- İkili bağlanıma ait VARÇÖZ çizelgesi aşağıdaki gibidir:

Değişimin Kaynağı	KT	sd	OKT
Bağlanım	159,732	1	159,732
Kalıntılar	493,370	79	6,24519
Toplam	653,102	80	

- Şimdi, ildeki erkek işsizlik oranını gösteren X_3 değişkenini modele eklemek istediğimizi varsayalım.
- Üçlü bağlanıma ait VARÇÖZ çizelgesi ise şöyle idi:

Değişimin Kaynağı	KT	sd	OKT
Bağlanım	205,022	2	102,511
Kalıntılar	448,080	78	5,74461
Toplam	653,102	80	

- KKT'deki $(493{,}370 - 448{,}080 = 45{,}290)$ birimlik azalmanın istatistiksel olarak anlamlı olup olmadığını bulmak istiyoruz.

Bir Açıklayıcı Değişkenin Marjinal Katkısı

- X_2'nin katkısı biliniyorken, X_3'ün marjinal katkısı şu sınama istatistiği ile ölçülebilir:

$$F = \frac{Q_3/sd}{Q_4/sd} = \frac{(\mathrm{KKT}_{eski} - \mathrm{KKT}_{yeni})/m}{\mathrm{KKT}_{yeni}/(n-k)}$$

- Burada m, yeni modele eklenen değişken sayısını gösterir.
- Elimizdeki örnek için F istatistiği şu şekilde hesaplanır:

$$F = \frac{(493{,}370 - 448{,}080)/1}{448{,}080/78} = 7{,}884$$

- Bulunan istatistik anlamlıdır. İldeki erkek işsizlik oranını eklemek KKT'yi anlamlı biçimde azaltmaktadır.

Bir Açıklayıcı Değişkenin Marjinal Katkısı

- Eldeki F oranı R^2 değerlerini kullanarak da bulunabilir:

$$F = \frac{(R^2_{yeni} - R^2_{eski})/m}{(1 - R^2_{yeni})/(n-k)}$$

- Örneğimiz için:

$$F = \frac{(0{,}3139 - 0{,}2446)/1}{(1 - 0{,}3139)/78} = 7{,}878$$

- Bu da yuvarlama hataları dışında önceki değer ile aynıdır.

Yeni Bir Değişken Ne Zaman Eklenmeli?

- Araştırmacılar çoğu zaman aynı bağımlı değişkeni içeren ama açıklayıcı değişkenleri farklı olan modeller arasında seçim yapmak durumunda kalırlar.
- Böyle durumlardaki genel eğilim en yüksek $\bar{R}^2$'yi seçmek yönündedir.
- Diğer yandan, yeni eklenen bir değişkenin katsayısının t değeri mutlak olarak 1'den büyük olduğu sürece $\bar{R}^2$ artar.
- Diğer bir deyişle yeni eklenen bir değişkene ilişkin $F(= t^2)$ değeri 1'den büyükse, bağlanım $\bar{R}^2$ değeri de yükselir.
- Demek ki $\bar{R}^2$ değerini yükselttiği halde KKT'yi istatistiksel olarak anlamlı ölçüde azaltmayan bir ek değişkenin modele eklenmesi konusunda dikkatli olunmalıdır.

Sınırlamalı Enküçük Kareler Yöntemi

- İktisat kuramı zaman zaman belli bir bağlanım modelindeki katsayılar için bir takım doğrusal sınırlamalar öngörebilir.
- Örnek olarak Cobb-Douglas üretim işlevini ele alalım:

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

- Burada Y üretim, X_2 emek girdisi, X_3 de sermaye girdisidir.
- Modelin log-doğrusal biçimdeki gösterimi şöyledir:

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

- Burada β_1', $\ln \beta_1$'dir.

Sınırlamalı Enküçük Kareler Yöntemi

- Eğer ölçeğe göre sabit getiri söz konusu ise, iktisat kuramı aşağıdaki doğrusal sınırlamayı öngörür:

$$\beta_2 + \beta_3 = 1$$

- $\beta_2 + \beta_3 = 1$ gibi bir doğrusal sınırlamanın geçerli olup olmadığı, t sınaması yöntemi kullanılarak görülebilir.
- Bunun için, önce model tahmin edilir ve $H_0 : \beta_2 + \beta_3 = 1$ önsavı bildik yolla sınanır:

$$t = \frac{(\hat{\beta}_2 + \hat{\beta}_3) - 1}{\sqrt{\text{var}(\hat{\beta}_2) + \text{var}(\hat{\beta}_3) + 2\text{cov}(\hat{\beta}_2 \hat{\beta}_3)}}$$

- Bulunan t değeri eğer seçili anlamlılık düzeyindeki kritik t değerinden büyükse, H_0 reddedilir.

Sınırlamalı Enküçük Kareler Yöntemi

- t sınaması yaklaşımı, "sınırlamasız" (unrestricted) bağlanım bulunduktan sonra sınama yapmaya dayandığı için yeğlenmeyen bir yöntemdir.
- Daha doğru bir yaklaşım "sınırlamalı enküçük kareler" (restricted least squares) yöntemidir.
- Bu yönteme göre, $(\beta_2 = 1 - \beta_3)$ denkleme en başta koyulur ve "sınırlamalı" (restricted) model aşağıdaki gibi türetilir:

$$\begin{aligned}
\ln Y_i &= \beta_1' + (1-\beta_3)\ln X_{2i} + \beta_3 \ln X_{3i} + u_i \\
&= \beta_1' + \ln X_{2i} + \beta_3(\ln X_{3i} - \ln X_{2i}) + u_i \\
\ln Y_i - \ln X_{2i} &= \beta_1' + \beta_3(\ln X_{3i} - \ln X_{2i}) + u_i \\
\ln(Y_i/X_{2i}) &= \beta_1' + \beta_3 \ln(X_{3i}/X_{2i}) + u_i
\end{aligned}$$

- Burada (X_{3i}/X_{2i}) sermaye/emek oranını, (Y_i/X_{2i}) ise çıktı/emek oranını gösteren önemli iktisadi büyüklüklerdir.

Sınırlamalı Enküçük Kareler Yöntemi

- Tanımlanan sınırlamanın geçerli olup olmadığı iki bağlanımın karşılaştırılması ile bulunur:

$$F = \frac{(\mathrm{KKT}_s - \mathrm{KKT}_{sz})/m}{\mathrm{KKT}_{sz}/(n-k)} = \frac{(R^2_{sz} - R^2_s)/m}{(1 - R^2_{sz})/(n-k)}$$

- m burada doğrusal sınırlama sayısını, sz ve s ise sınırlamasız ve sınırlamalı bağlanımları göstermektedir.
- Dikkat: Sınırlamasız ve sınırlamalı modellerde bağımlı değişken farklı ise, R^2_{sz} ve R^2_s'nin birlikte kullanılabilmesi için gerekli dönüşümün yapılmış olması önemlidir.

Sınırlamalı Enküçük Kareler Açıklayıcı Örnek

- Örnek olarak, Tayvan tarım kesimi için Cobb-Douglas üretim modelini ölçeğe göre sabit getiri sınırlaması ile tahmin edelim:

$$\begin{array}{llll} \ln(\widehat{Y_i/X_{2i}}) = & 1{,}7086 & + 0{,}61298 \ln(X_{3i}/X_{2i}) & \\ \text{öh} & (0{,}4159) & (0{,}0933) & r^2 = 0{,}7685 \end{array}$$

- Sınırlamasız model için R^2 değeri, gerekli dönüştürmeden sonra 0,8489 olarak bulunur ve şu F istatistiği hesaplanır:

$$F = \frac{(R^2_{sz} - R^2_s)/m}{(1 - R^2_{sz})/(n-k)} = \frac{(0{,}8489 - 0{,}7685)/1}{(1 - 0{,}8489)/12} = 6{,}385$$

- F çizelgesinden, gözlenen değerin %5 düzeyinde anlamlı olduğu görülür ve $H_0 : \beta_2 + \beta_3 = 1$ sıfır önsavı reddedilir.

Sınırlamalı Enküçük Kareler Açıklayıcı Örnek

- Eğer sınırlamanın geçerli olduğuna karar verilmiş olsaydı, sınırlamalı model için tahmin edilen 0,61298 değeri β_3'ü gösterdiği için β_2 de 0,38702 olarak kolayca bulunurdu.
- Şimdi de sınırlamasız bağlanım bulgularına bir göz atalım:

$$\begin{array}{llllll} \widehat{\ln Y_i} = & -3{,}3384 & + 1{,}4988 \ln X_{2i} & + 0{,}4899 \ln X_{3i} & & R^2 = 0{,}8890 \\ \text{öh} & (2{,}4495) & (0{,}5398) & (0{,}1020) & & \bar{R}^2 = 0{,}8705 \end{array}$$

- Yukarıda emek girdisi esnekliğinin istatistiksel olarak anlamlı olmadığı görülüyor.
- Bu örnek, yalnızca tahmin edilen katsayılar ile yetinmeyip biçimsel sınama da yapmanın daha iyi bir çözümleme için gerekliliğini göstermesi bakımından önemlidir.

Genel F Sınaması

- Bir açıklayıcı değişkenin marjinal katkısı bölümünde söz edilen "yeni" model aslında sınırlamasız modeldir. Buna göre "eski" model de $\beta_3 = 0$ varsayımı ile sınırlamalı olur.
- Aslında, bağlanım bütününün anlamlılığını sınamaya ilişkin formüldeki payın BKT olmasının nedeni de buradaki "süper sınırlamalı" modelin KKT'sinin $\text{TKT}_{sz} = \text{TKT}$ olmasıdır.
- Ele almış olduğumuz örneklerden de anlaşılacağı gibi, F sınaması yöntemi k değişkenli bağlanım modelindeki m anakütle katsayısının sınanması için genel bir yöntemdir.
- Örnek olarak

$$
\begin{array}{ll}
H_0 : \ \beta_2 = \beta_3 & (m = 1), \\
H_0 : \ \beta_3 + \beta_4 + \beta_5 = 3 & (m = 1), \\
H_0 : \ \beta_3 = \beta_4 = 0 & (m = 2), \\
H_0 : \ \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0 & (m = 4)
\end{array}
$$

gibi pek çok farklı önsav F sınaması ile sınanabilir.

Genel F Sınaması

Genel F sınamasının adımları aşağıdaki gibi özetlenebilir:

1. Birincisi geniş sınırlamasız model ve diğeri de daha dar sınırlamalı model olmak üzere iki model vardır.
2. Bunlardan ikincisi, birinciden bazı değişkenler çıkarılarak ya da çeşitli doğrusal sınırlamalar getirilerek elde edilir.
3. Daha sonra; sınırlamasız ve sınırlamalı modeller verilere yakıştırılır ve KKT_{sz} ve KKT_s toplamları ya da R^2_{sz} ve R^2_s belirleme katsayıları bulunur.
4. Eğer bağımlı değişkenler farklıysa, R^2_{sz} ve R^2_s kullanmak için bunları önce birbirleriyle uyumlandırmak gereklidir.
5. KKT_{sz} için serbestlik derecesi $(n-k)$'dir. KKT_s için ise serbestlik derecesi toplam kısıtlama sayısı m'dir.
6. Son olarak, formülü verilen F istatistiği hesaplanır ve bu değer $F_\alpha(m, n-k)$'den büyükse sıfır önsavı reddedilir.

Ders Planı

Yapısal Kararlılığın Sınanması

- Eğer model katsayıları zaman içerisinde sabit kalmayıp değişime uğruyorlar ise bu duruma "yapısal değişim" (structural change) denir.
- Yapısal değişime örnek neden olarak
  2001 yılında dalgalı kur rejimine geçiş,
  1999 yılı vergi yasası reformu,
  1990-1991 Körfez Savaşı,
  1973-1977 OPEC petrol ambargosu
  gibi ulusal ya da küresel etmenler gösterilebilir.
- Yapısal değişim konusu özellikle zaman serileri içeren bağlanım modellerinde önemlidir.

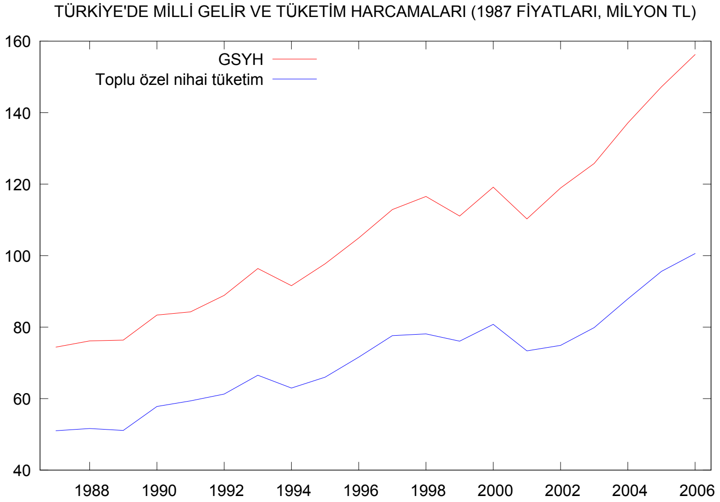
Yapısal Kararlılığın Sınanması

- Bir yapısal değişimin varlığını görebilmeye örnek olarak, 1987 ve 2006 yılları arasında Türkiye'de toplam tüketim harcamaları ve GSYH (1987 fiyatları, milyon TL) örneğimizi 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

Yapısal Kararlılığın Sınanması



Yapısal Kararlılığın Sınanması

- Türkiye'de tüketim harcamaları ve milli gelir arasındaki ilişkiyi incelemek istiyoruz.
- Elimizde 1987 ve 2006 yılları arasını kapsayan bir SEK bağlanımını tahmin etmek için yeterli veriler bulunmaktadır.
- Diğer yandan, tasarruf ve gelir arasındaki ilişkinin 20 yıl boyunca aynı kaldığını varsaymak fazla inandırıcı olmaz.
- Örnek olarak, Şubat 2001 ve öncesinde yaşanan olaylar sonrasında Cumhuriyet tarihindeki en büyük ekonomik krizlerden birinin ortaya çıkmış olduğunu biliyoruz.
- Buna dayanarak, 2001 ve sonrası dönemin yapısal olarak farklı olup olmadığını görmek istediğimizi varsayalım.

Yapısal Kararlılığın Sınanması

- Yapısal kararlılığı sınamak için örneklemi 2001 öncesi ve 2001 ve sonrası olarak iki döneme ayırabiliriz.
- Böylece elimizde tahmin edilebilecek üç ayrı bağlanım olur:

$$
\begin{array}{lll}
\text{1987-2000 dönemi:} & Y_t = \lambda_1 + \lambda_2 X_t + u_{1t} & (n_1 = 14) \\
\text{2001-2006 dönemi:} & Y_t = \gamma_1 + \gamma_2 X_t + u_{2t} & (n_2 = 6) \\
\text{1987-2006 dönemi:} & Y_t = \alpha_1 + \alpha_2 X_t + u_{3t} & (n_3 = 14 + 6 = 20)
\end{array}
$$

- Yukarıdaki üçüncü bağlanım, tüm gözlemleri kapsamakta ve 1987-2006 aralığı içinde yapısal bir değişim olmadığını varsaymaktadır.
- Öyleyse üçüncü model, $\lambda_1 = \gamma_1$ ve $\lambda_2 = \gamma_2$ koşullarından dolayı bir sınırlamalı model olarak düşünülebilir.

Yapısal Kararlılığın Sınanması

- Üç bağlanıma ait bulgular aşağıdaki gibidir:

$$\begin{array}{llll} \hat{Y}_t = & 1{,}5027 & +\ 0{,}6679X_t & R^2 = 0{,}9937 \\ t & (1{,}0156) & (43{,}5400) & \text{KKT}_1 = 8{,}9210 \end{array}$$

$$\begin{array}{llll} \hat{Y}_t = & 1{,}0706 & +\ 0{,}6358X_t & R^2 = 0{,}9835 \\ t & (0{,}1947) & (15{,}4361) & \text{KKT}_2 = 10{,}3487 \end{array}$$

$$\begin{array}{llll} \hat{Y}_t = & 8{,}0344 & +\ 0{,}5934X_t & R^2 = 0{,}9854 \\ t & (4{,}3310) & (34{,}8352) & \text{KKT}_3 = 55{,}0062 \end{array}$$

- Sonuçlar, tasarruf ve gelir arasındaki ilişkinin iki alt döneme ait tahminlerinde farklılıklar olduğunu göstermektedir.
- Buna göre üçüncü bağlanımın uygun ve güvenilir olmadığı düşünülebilir.

Chow Sınaması

- Yapısal değişimin varlığını sınamak için kullanılabilecek yöntemlerden biri Chow sınamasıdır.
- Bu sınama, bildiğimiz F sınamasından farklı olmamakla birlikte geliştiricisi Gregory Chow'un adıyla anılır.
- Chow sınamasının gerisinde iki önemli varsayım vardır:
- Varsayım 1: Birinci ve ikinci modellere ait hata terimleri aynı sabit varyans ile normal dağılmaktadırlar:

$$u_{1t} \sim N(0, \sigma^2) \text{ ve } u_{2t} \sim N(0, \sigma^2)$$

- Varsayım 2: u_{1t} ve u_{2t} aynı zamanda bağımsız dağılırlar.

Chow Sınamasının Adımları

Verilen varsayımlar altında Chow sınaması şöyle yapılır:

1. Birinci modelden sd'si $(n_1 - k)$ olan KKT_1 bulunur.
2. İkinci modelden sd'si $(n_2 - k)$ olan KKT_2 bulunur.
3. İki bağlanıma ait hata terimleri bağımsız kabul edildiği için, $\text{KKT}_{sz} = \text{KKT}_1 + \text{KKT}_2$ olarak hesaplanır.
4. Tüm gözlemlerin kullanıldığı 3. model tahmin edilir ve KKT_3 ya da KKT_s bulunur.
5. Yapısal değişim yoksa KKT_s ve KKT_{sz} istatistiksel olarak farklı olmamalıdır. Sınamak için şu istatistik hesaplanır:

$$F = \frac{(\text{KKT}_s - \text{KKT}_{sz})/k}{(\text{KKT}_{sz})/(n_1 + n_2 - 2k)} \sim F_{[k,(n_1+n_2-2k)]}$$

Chow Sınaması

- Örneğimize dönecek olursak *F* istatistiğini şöyle buluruz:

$$F = \frac{(55{,}0062 - 19{,}2697)/2}{(19{,}2697)/(16)} = 14{,}8363$$

- Gözlenen değer 2 ve 22 sd için yüzde 1 kritik *F* değeri olan 6,23'ten büyük olduğu için; $H_0 : \lambda_1 = \gamma_1, \ \lambda_2 = \gamma_2$ reddedilir.
- Demek ki Chow sınaması 2001 yılında Türkiye'nin bir yapısal değişim geçirdiği savını desteklemektedir.

Chow Sınaması

Chow sınaması ile ilgili şu noktalara dikkat edilmelidir:

- Chow sınaması, birden fazla yapısal değişimin varlığını sınamak için genellenebilir.
- Örnek olarak, örneklemi üç ayrı döneme bölüp dört farklı bağlanım tahmini yapmak ve daha sonra KKT_s'yi de $\text{KKT}_1 + \text{KKT}_2 + \text{KKT}_3$ olarak hesaplamak olanaklıdır.
- Chow sınamasında "yapısal kırılma" (structural break) noktasının hangi dönemde yer aldığının bilindiği varsayılır.
- Chow sınaması iki bağlanımın farklı olup olmadığını söyler ancak farkın sabit terimden mi, X_t'nin katsayısından mı, ya da aynı anda her ikisinden mi kaynaklandığını bildirmez.
- Yapısal değişimin kaynağının ne olduğunu anlamak için kukla değişkenlere dayanan farklı bir yaklaşım gereklidir.
- Ayrı dönemlere ait hata varyanslarının sabit olduğu varsayımının ayrıca sınanması gerekli olabilir.

MWD Sınaması

- Doğrusal ve log-doğrusal model arasında bir seçim yapma zorunluluğu, görgül çalışmalarda sık sık ortaya çıkar.
- Böyle bir model seçimi için MacKinnon, White ve Davidson (1983) tarafından önerilen MWD sınaması kullanılabilir.
- MWD sınaması şu sıfır ve almaşık önsavları içerir:

$$H_0: \quad \text{Doğrusal model}$$
$$H_1: \quad \text{Log-doğrusal model}$$

MWD Sınamasının Adımları

MWD sınamasının adımları aşağıdaki gibidir:

1. Doğrusal model tahmin edilir ve $\hat{Y}$ bulunur.
2. Log-doğrusal model tahmin edilir ve $\widehat{\ln Y}$ bulunur.
3. $Z_1 = \ln \hat{Y} - \widehat{\ln Y}$ değişkeni türetilir.
4. Y'nin X'lere ve Z_1'e göre bağlanımı hesaplanır. Eğer Z_1'in katsayısı bilindik t sınaması ile istatistiksel olarak anlamlı çıkarsa, H_0 reddedilir.
5. $Z_2 = \exp(\widehat{\ln Y}) - \hat{Y}$ değişkeni türetilir.
6. $\ln Y$'nin $\ln X$'lere ve Z_2'ye göre bağlanımı hesaplanır. Eğer Z_2'nin katsayısı t sınaması ile anlamlı bulunursa, H_1 savı reddedilir.

MWD Sınaması

Karmaşık gibi görünse de MWD sınamasının mantığı basittir:

- Eğer doğrusal model gerçekten doğru modelse, dördüncü adımda hesaplanan Z_1 değeri anlamlı olmamalıdır.
- Çünkü böyle bir durumda doğrusal modelin $\hat{Y}$ kestirimleri (karşılaştırma yapabilmek için logları alındıktan sonra) ile log-doğrusal modelin kestirimleri farklı çıkmamalıdır.
- Aynı yorum H_1 almaşık önsavı için de geçerlidir.

MWD Sınaması Açıklayıcı Örnek

- Örnek olarak 1971-1975 arası dönem için ABD'nin Detroit şehrindeki gül talebini ele alalım:

  Doğrusal model: $Y_t = \alpha_1 + \alpha_2 X_{2t} + \alpha_3 X_{3t} + u_t$

  Log-log model: $\ln Y_t = \beta_1 + \beta_2 \ln X_{2t} + \beta_3 \ln X_{3t} + v_t$

- Burada
  Y satılan gül miktarını (düzine),
  X_2 ortalama toptan gül fiyatını (dolar),
  X_3 ise ortalama toptan karanfil fiyatını (dolar)
  göstermektedir.
- Beklentiler α_2 ile β_2'nin eksi, α_3 ve β_3'ün ise artı değerli olması yönündedir.

MWD Sınaması Açıklayıcı Örnek

- Bağlanım bulguları aşağıdaki gibidir:

$$\begin{array}{llllll}
\widehat{Y}_t & = 9734{,}2176 & - \; 3782{,}1956 X_{2t} & + \; 2815{,}2515 X_{3t} & R^2 = 0{,}7710 \\
t & (3{,}3705) & (-6{,}6069) & (2{,}9712) & F = 21{,}84
\end{array}$$

$$\begin{array}{llllll}
\widehat{\ln Y}_t = & 9{,}2278 & - \; 1{,}7607 \ln X_{2t} & + \; 1{,}3398 \ln X_{3t} & R^2 = 0{,}7292 \\
t & (16{,}2349) & (-5{,}9044) & (2{,}5407) & F = 17{,}50
\end{array}$$

- Görüldüğü gibi hem doğrusal hem de log-doğrusal model verilere iyi yakışmıştır.
- Katsayılar beklenen işaretleri taşımaktadır ve t değerleri de istatistiksel olarak anlamlıdır.

MWD Sınaması Açıklayıcı Örnek

- Önce modelin doğrusal olup olmadığını sınayalım:

$$\begin{array}{llllll} \widehat{Y}_t & = 9727{,}5685 & - 3783{,}0623X_{2t} & + 2817{,}7157X_{3t} & + 85{,}2319Z_{1t} \\ t & (3{,}2178) & (-6{,}337) & (2{,}8366) & (0{,}0207) \\ R^2 = & 0{,}7707 & & & F = 13{,}44 \end{array}$$

- Z_1'in katsayısı anlamlı olmadığına göre, modelin gerçekte doğrusal olduğunu öne süren önsavı reddetmiyoruz.
- Şimdi de gerçek modelin log-doğrusallığını sınayalım:

$$\begin{array}{llllll} \widehat{\ln Y_t} = & 9{,}1486 & - 1{,}9699 \ln X_{2t} & + 1{,}5891 \ln X_{3t} & - 0{,}0013Z_{2t} \\ t & (17{,}0825) & (-6{,}4189) & (3{,}0728) & (-1{,}6612) \\ R^2 = & 0{,}7798 & & & F = 14{,}17 \end{array}$$

- Z_2'ye ait t değeri $-1{,}6612$'dir. Dolayısıyla log-doğrusallık varsayımı da %5 anlamlılık düzeyinde reddedilemez.
- Örneğin de gösterdiği gibi bazı durumlarda modellerin ikisi de reddedilmeyebilmektedir.

Almaşık Sınamalar

- Görüldüğü gibi doğrusal bağlanım modelleri çerçevesinde çeşitli önsavları sınamak için t ve F sınamalarından yararlanılabilmektedir.
- Doğrusal modellerin basit dünyasından çıkıldığında ise doğrusal ve doğrusal-dışı her modelde kullanılabilecek önsav sınamalarına gereksinim duyulur.
- Bu amaç için sıkça kullanılan üç yöntem şunlardır:

  "Wald sınaması" (Wald test)

  "Olabilirlik oranı" (likelihood ratio), kısaca "OO" (LR)

  "Lagrange çarpanı" (Lagrange multiplier), kısaca "LÇ" (LM)
- Bu üç sınama kavuşmazsal olarak eşdeğerdir ve üçünün de sınama istatistiği χ^2 dağılımına uyar.
- Diğer yandan, doğrusal modellerdeki her türlü sınama için F yeterlidir ve Wald, OO ve LÇ'ye bakmaya gerek yoktur.
- Dolayısıyla bu sınama üçlüsünü şimdilik ele almayacağız.

Çoklu Bağlanım ve Kestirim

- Tahmin edilen bir bağlanım işlevi, belli bir X_0 değerine karşılık gelen Y'yi kestirmek için kullanılabilir.
- İki farklı kestirim türü vardır: "Ortalama kestirimi" (mean prediction) ve "bireysel kestirim" (individual prediction).
- Ortalama kestirimi, belli X_0 değerlerine karşılık gelen $E(Y|X_0)$ koşullu olasılık değerinin kestirilmesini içerir.
- Bireysel kestirim ise X_0'ın karşılığı olan tekil $Y|X_0$ değerinin kestirilmesi demektir.
- Ortalama kestirimi, anakütle bağlanım işlevindeki noktanın kestirimidir ve varyansı bireysel kestirimden daha küçüktür.
- Çoklu bağlanımda kestirim değerlerinin varyans ve ölçünlü hata formülleri karışık olduğu için bunları daha sonra dizey gösterimi ile ele alacağız.

Önümüzdeki Dersin Konusu ve Ödev

Ödev

Kitaptan Bölüm 8 “Multiple Regression Analysis: The Problem of Inference” okunacak.

Önümüzdeki Ders

Kukla Değişkenlerle Bağlanım