Farklıserpilimsellik

Farklıserpilimselliğin Niteliği

Ekonometri 2 – Konu 8

Sürüm 2,0 (Ekim 2011)

UADMK Açık Lisans Bilgisi

A. Talha Yalta
TOBB Ekonomi ve Teknoloji Üniversitesi
Ekim 2011 BY-NC-SA

Ders Planı

Farklıserpilimsellik Kavramı

Klasik doğrusal bağlanım modelinin önemli bir varsayımı, hata teriminin sabit varyans ile dağılmakta olduğudur.
Bu varsayım her zaman geçerli olmayabilir.

Bu bölümde şu sorulara yanıt arayacağız:

1. Farklıserpilimselliğin niteliği nedir?
2. Uygulamada doğurduğu sonuçlar nelerdir?
3. Varlığı nasıl anlaşılabilir?
4. Düzeltmek için ne gibi önlemler alınabilir?

Farklıserpilimsellik Kavramı

- “Aynıserpilimsellik” (homoscedasticity) varsayımına göre verili X_i açıklayıcı değişkenlerine bağlı olarak Y_i'nin koşullu varyansı sabittir:

$$E(u_i^2) = \sigma^2 \qquad i = 1, 2, \ldots, n$$

- “Farklıserpilimsellik” (heteroscedasticity) durumunda ise X_i değiştikçe Y_i'nin koşullu varyansı da değişir:

$$E(u_i^2) = \sigma_i^2$$

- Farklıserpilimselliğe bir örnek olarak tasarrufların varyansının gelirle birlikte artmasını verebiliriz.
- Yüksek gelirli ailelerin tasarrufları, düşük gelirli ailelere oranla hem ortalama olarak daha çoktur hem de değişirliği daha fazladır.

Farklıserpilimselliğin Nedenleri

Hata terimi varyansının değişken olma nedenlerinden bazıları şunlardır:

1. "Hata-öğrenme" (error-learning) modellerine göre insanlar bazı konuları öğrendikçe daha az hata yaparlar. Buna göre σ^2'nin de zamanla küçülmesi beklenir.
   Örnek olarak, daktilo kullanma süresi arttıkça hem daktilo hataları hem de bunların varyansı azalır.
2. Gelir düzeyi arttıkça gelirin harcanabileceği seçenekler de genişler. Böylece, gelir düzeyi ile birlikte hem harcamaların hem de bunların varyansının artması beklenir.
3. Zaman içerisinde veri derleme tekniklerinin gelişmesine koşut olarak σ_i^2 de düşebilir.

(. . . devam)

Farklıserpilimselliğin Nedenleri

4. Farklıserpilimsellik, "dışadüşen" (outlier) gözlemlerin bir sonucu olarak da ortaya çıkabilir.
   Böyle gözlemlerin alınması ya da bırakılması, özellikle de örneklem küçükken sonuçları önemli ölçüde değiştirebilir.
5. Farklıserpilimselliğin bir diğer nedeni model belirtim hatasıdır. Özellikle de önemli bir değişkenin modelden çıkartılması farklıserpilimselliğe yol açabilir.
6. Farklıserpilimsellik sorunu yatay kesit verilerinde zaman serisi verilerine oranla daha fazla görülebilmektedir.
   Bunun nedeni, zaman serilerinde değişkenlerin zaman içerisinde yakın büyüklüklerde olma eğilimidir.

Farklıserpilimselliğin SEK Tahminlerine Etkisi

- σ_i^2 şeklindeki farklıserpilimsellik altında SEK tahmincilerinin varyanslarının ne şekilde etkileneceğini görmek için, iki değişkenli modeli ele alalım:

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

Aynıserpilimsellik durumunda varyans formülü şöyledir:

Aynıserpilimsellik

$$\mathrm{var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_i^2}$$

Bu da farklıserpilimsellik altındaki formülden farklıdır:

Farklıserpilimsellik

$$\mathrm{var}(\hat{\beta}_2) = \frac{\sum x_i^2 \sigma_i^2}{(\sum x_i^2)^2}$$

- Her bir i için, $\sigma_i^2 = \sigma^2$ olması durumunda iki formülün aynı olacağına dikkat ediniz.

Farklıserpilimselliğin SEK Tahminlerine Etkisi

- KDBM'nin tüm varsayımları geçerli olduğu zaman SEK tahmincisi $\hat{\beta}_2$'nın EDYT olduğunu anımsayalım.
- $\hat{\beta}_2$ tahmincisinin aynıserpilimselliğin geçerli olmadığı durumda bile doğrusal ve yansız olduğu gösterilebilir.
- Ancak, böyle bir durumda SEK tahmincileri artık "en iyi" ya da enaz varyanslı olma özelliklerini kaybederler.
- Öyleyse, farklıserpilimsellik durumunda tahmincilerin EDYT olabilmesi için SEK'ten ayrı bir yöntem izlemek gerekir.

Genellemeli En Küçük Kareler

- X_i'nin farklı düzeylerinde farklıserpilimsellik gözleniyorsa, tahmin sürecinde bu bilgiden yararlanmak gerekir.
- Örnek olarak, düşük gelir sınıflarına ait harcamalar daha düşük varyanslı ise bu gruplardan gelen gözlemlere daha çok ağırlık verilmesi istenir.
- Bunun nedeni, düşük varyanslı grupların kendi ortalamaları çevresine daha yakın dağılarak ABİ'nin daha doğru tahmin edilmesini sağlamalarıdır.
- SEK yöntemi, tüm gözlemlere eşit ağırlık verdiği için farklıserpilimsellik durumunda etkin tahminciler üretemez.
- Varyanstaki değişim bilgisinden yararlanan ve bu nedenle tahmincileri EDYT olan yöntem ise "genellemeli en küçük kareler" (generalized least squares) ya da kısaca "GEK" (GLS) yöntemidir.

Genellemeli En Küçük Kareler

- GEK'in farklıserpilimsellik bilgisini nasıl kullandığını görmek için iki değişkenli modeli şöyle yazalım:

$$Y_i = \beta_1 X_{0i} + \beta_2 X_i + u_i$$

- Burada her bir i için $X_{0i} = 1$'dir.
- Farklıserpilimsel varyanslar (σ_i^2) biliniyor olsun. Yukarıdaki denklemi σ_i'ye bölersek şunu elde ederiz:

$$\frac{Y_i}{\sigma_i} = \beta_1 \left(\frac{X_{0i}}{\sigma_i} \right) + \beta_2 \left(\frac{X_i}{\sigma_i} \right) + \left(\frac{u_i}{\sigma_i} \right)$$

- Bunu da gösterim kolaylığı bakımından şöyle yazabiliriz:

$$Y_i^* = \beta_1^* X_{0i}^* + \beta_2^* X_i^* + u_i^*$$

- Buradaki yıldız işaretli dönüştürülmüş değişkenler, baştaki değişkenlerin σ_i'ye bölünmüş halleridir. İkinci modele ait katsayılar farklı olacağı için β'lar da yıldız ile gösterilmiştir.

Genellemeli En Küçük Kareler

- Özgün modeli neden dönüştürdüğümüzü görmek için hata terimi u_i^*'ın şu özelliğine dikkat edelim:

$$\begin{aligned} \text{var}(u_i^*) = E(u_i^*)^2 &= E\left(\frac{u_i}{\sigma_i}\right)^2 \\ &= \frac{1}{\sigma_i^2} E(u_i^2) \qquad (\sigma_i^2 \text{ bilindiği için}) \\ &= \frac{1}{\sigma_i^2}(\sigma_i^2) = 1 \end{aligned}$$

- Demek ki dönüştürülen hata terimi u_i^*'ın varyansı sabittir ve 1'e eşittir.

Genellemeli En Küçük Kareler

- Görüldüğü gibi GEK, KDBM varsayımlarını sağlayan dönüştürülmüş değişkenlere uygulanan SEK yöntemidir.
- Uygulamada, β_1^* ve β_2^*'ı tahmin etmek için dönüştürülen modelin ÖBİ'si kullanılır:

$$Y_i^* = \hat{\beta}_1^* X_{0i}^* + \hat{\beta}_2^* X_i^* + \hat{u}_i{}^*$$

- Daha sonra hata kareleri toplamı $\sum \hat{u}_i{}^{2*}$ enazlanır.
- GEK tahmincisi $\hat{\beta}_2^*$ ve bunun varyansı var$(\hat{\beta}_2^*)$ şöyledir:

$$\begin{aligned}
\hat{\beta}_2^* &= \frac{(\sum w_i)(\sum w_i X_i Y_i) - (\sum w_i X_i)(\sum w_i Y_i)}{(\sum w_i)(\sum w_i X_i^2) - (\sum w_i X_i)^2} \\
\text{var}(\hat{\beta}_2^*) &= \frac{\sum w_i}{(\sum w_i)(\sum w_i X_i^2) - (\sum w_i X_i)^2}
\end{aligned}$$

- Burada $w_i = 1/\sigma_i^2$'yi göstermektedir.

SEK ile GEK Arasındaki Fark

- Bilindiği gibi SEK, hata kareleri toplamını enazlar:
$$\sum \hat{u}_i^2 = \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2$$
- GEK yöntemi ise aşağıdaki dönüştürmeli hata kareleri toplamını enazlamaktadır:
$$\sum w_i \hat{u}_i^2 = \sum w_i (Y_i - \hat{\beta}_1^* - \hat{\beta}_2^* X_i)^2$$
- Buna göre GEK $w_i = 1/\sigma_i^2$ büyüklüğü ile ağırlıklandırılan kalıntı kareleri toplamını enazlarken, SEK de ağırlıksız ya da eşit ağırlıklı KKT'yi enazlamaktadır.
- GEK'te her gözleme verilen ağırlık σ_i ile ters orantılıdır.
- Elimizdeki yöntem ağırlıklandırmalı bir KKT'yi enazlamaya dayandığına göre "ağırlıklı en küçük kareler" (weighted least squares) ya da "AEK" (WLS) diye de adlandırılabilir.
- Demek ki AEK, daha genel bir tahmin yöntemi olan GEK'in özel bir durumudur.

Farklıserpilimselliği Göz Önüne Alan SEK

- Farklıserpilimsellik altında SEK tahmini, farklıserpilimselliği göz önüne alarak ya da göz ardı ederek yapılabilir.
- Yapılan tahminler iki şekilde de hatalı ya da yanıltıcı olabilir.
- Farklıserpilimselliği göz önüne alan SEK tahmincisini ele alalım.
- Göstermiş olduğumuz gibi, SEK tahmincisi $\hat{\beta}_2$'nın ve GEK tahmincisi $\hat{\beta}_2^*$'ın her ikisi de yansız tahmincilerdir.
- Ancak enaz varyanslı olan tahminci GEK tahmincisi $\hat{\beta}_2^*$'dır.
- Bu durum SEK tahmincisine dayanan güven aralıklarının gereksiz yere büyük çıkacağı anlamına gelir.
- Demek ki SEK tabanlı anlamlı olmayan bir katsayı, GEK ile hesaplanmış doğru bir güven aralığı kurulursa anlamlı çıkabilir.

Farklıserpilimselliği Göz Ardı Eden SEK

- Farklıserpilimsellik altında, aynıserpilimselliğe ait varyans formülünü kullanmayı sürdürmek ciddi sorunlar yaratabilir.
- $\hat{\beta}_2$'nın varyansının aynıserpilimsellik ve farklıserpilimsellik varsayımı altındaki formüllerini anımsayalım:

Aynıserpilimsellik

$$\text{var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_i^2}$$

Farklıserpilimsellik

$$\text{var}(\hat{\beta}_2) = \frac{\sum x_i^2 \sigma_i^2}{(\sum x_i^2)^2}$$

- Farklıserpilimsellik durumunda, yukarıda verilen formüllerden sağdaki soldakinin yanlı bir tahmincisi olur.
- Genellikle bu yanlılığın yukarı doğru mu yoksa aşağı doğru mu olduğu da bilinemez.
- Demek ki farklıserpilimsellik altında bildik sınama sürecini kullanmada ısrar etmek yanıltıcı sonuçlara yol açabilir.

SEK Kullanmanın Sonuçları

- Örnek olarak, Davidson ve MacKinnon'ın yapmış oldukları bir Monte Carlo çalışmasını ele alalım.
- Yazarlar, iki değişkenli bağlanım modelini kullanarak ve $\beta_1 = \beta_2 = 1$ ve $u_i \sim N(0, X_i^{\alpha})$, diğer bir deyişle hata teriminin açıklayıcı değişken X'in α üssü değeriyle ilişkili olduğu varsayımını yaparak şu sonuçları elde etmişlerdir:

	$\hat{\beta}_1$'nın ölçünlü hatası			$\hat{\beta}_2$'nın ölçünlü hatası		
α	SEK	SEK_{fs}	GEK	SEK	SEK_{fs}	GEK
0,5	0,164	0,134	0,110	0,285	0,277	0,243
1,0	1,142	0,101	0,048	0,246	0,247	0,173
2,0	0,116	0,074	0,0073	0,200	0,220	0,109
3,0	0,100	0,064	0,0013	0,173	0,206	0,056
4,0	0,089	0,059	0,0003	0,154	0,195	0,017

- SEK_{fs} burada farklıserpilimselliği göz önüne alan SEK'tir.

(. . . devam)

SEK Kullanmanın Sonuçları

- Bulgular, farklıserpilimselliği göz önüne alan SEK_{fs} ve almayan SEK ölçünlü hatalarının GEK'inkilerden yüksek olduğunu, kısaca GEK'in en düşük varyanslı olduğunu göstermektedir.
- Ayrıca, farklıserpilimselliği göz ardı eden yanlı SEK'in varyansı SEK_{fs}'nin varyansından büyük ya da küçük olabilmektedir.
- Buna göre, farklıserpilimsellik durumunda GEK yönteminin üstünlüğü açıktır.
- Ancak GEK'i uygulayabilmek her zaman kolay değildir.

Önümüzdeki Dersin Konusu

Önümüzdeki ders

Farklıserpilimselliği saptamak