Bölüm 4

Farklıserpilimsellik

4.1 Farklıserpilimselliğin Niteliği

4.1.1 Nedenleri ve Sonuçları

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?

- *"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.

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

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

Bu da farklıserpilimsellik altındaki formülden farklıdır:

Farklıserpilimsellik

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

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

4.1.2 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.
- 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.
- 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.
- 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.

4.1.3 Farklıserpilimsellik Altında SEK

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)

- 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.

4.2 Farklıserpilimselliği Saptamak

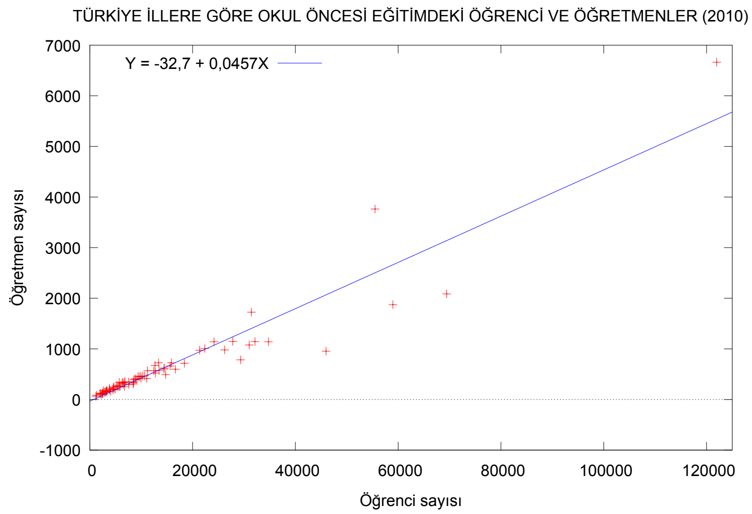
- *"Farklıtürel"* (heterogeneous) birimler içeren yatay kesit verilerinde, farklıserpilimsellik asla sıra dışı değildir.
- Örnek olarak küçük, orta ve büyük firmalar eğer birarada örneklenmişse farklıserpilimsellik genellikle beklenir.
- Belli bir durumda farklıserpilimselliğin varlığını anlamanın kesin bir yolu yoktur. Çeşitli başparmak kuralları vardır.
- Bunun nedeni, σ_i^2'yi bilebilmek için bütün Y anakütlesini bilmenin gerekli olmasıdır.
- Ancak çoğu iktisadi çalışmada belli X değerlerine karşılık tek bir Y örneklem değeri bulunur.
- Bu nedenle, farklıserpilimselliğin varlığını anlamak için kullanılabilecek biçimsel ve biçimsel olmayan yöntemlerin çoğu $\hat{u}_i$ SEK kalıntılarının incelenmesine dayanır.
- Yalnızca $\hat{u}_i$'lar gözlenebildiği için de bunların gerçek u_i'lerin iyi birer tahmincisi oldukları umulur.

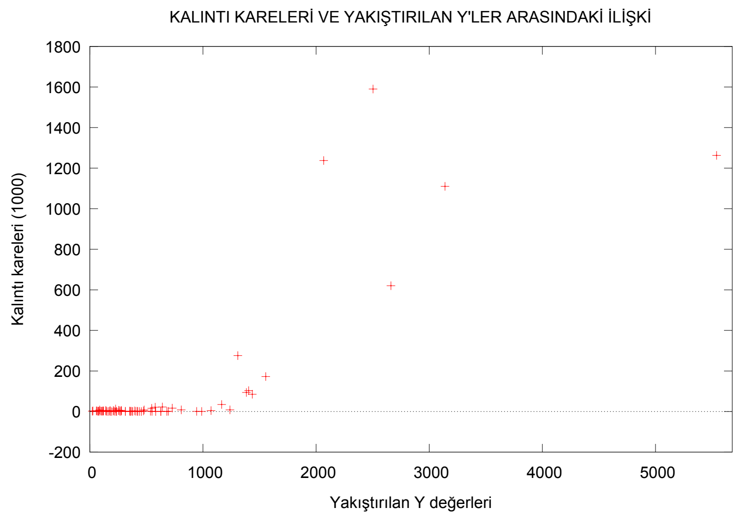
4.2.1 Biçimsel Olmayan Yöntemler

Çizim Yöntemi

- Çoğu durumda farklıserpilimselliği saptamak bir sezgi, eğitimli bir tahmin ya da bir önsel görgül deneyim konusudur.
- Biçimsel bir yöntem olmayan çizim yönteminde, bağlanım çözümlemesi önce aynıserpilimsellik varsayımı ile yapılır.
- Sonra, $\hat{u}_i{}^2$ kalıntı karelerinin düzenli bir görüntü sergileyip sergilemediklerine bakılır.
- Bunun için, $\hat{u}_i{}^2$'lerin $\hat{Y}_i$ ve çeşitli X_i değişkenleri ile ilişkileri çizit üzerinde görüntülenir.
- Eğer örneklem yeterince büyükse, $\hat{u}_i{}^2$'ler u_i^2'lerle aynı şey olmasa da onların yerine kullanılabilirler.

- Eğer bu işlem $\hat{u}_i{}^2$'ler ile $\hat{Y}_i$ ya da X_i arasında doğrusal ya da ikinci derece bir ilişki gösterirse, bu bilgi kullanılarak veriler farklıserpilimsellik sergilemeyecek biçimde dönüştürülür.





4.2.2 Biçimsel Yöntemler

Park Sınaması

- R. E. Park (1966), σ_i^2'nin açıklayıcı değişken X_i'nin bir işlevi olduğunu ileri sürerek çizim yöntemini biçimselleştirir.

- Sınama için öne sürülen iki işlev kalıbı şudur:

$$\begin{aligned} \sigma_i^2 &= \sigma^2 X_i^{\beta} e^{v_i} \\ \text{ya da} \quad \ln \sigma_i^2 &= \ln \sigma^2 + \beta \ln X_i + v_i \end{aligned}$$

- Park, σ_i^2 bilinmediği için yerine $\hat{u}_i{}^2$'yi kullanmayı önerir:

$$\begin{aligned} \ln \hat{u}_i{}^2 &= \ln \sigma^2 + \beta \ln X_i + v_i \\ &= \alpha + \beta \ln X_i + v_i \end{aligned}$$

- Demek ki önce farklıserpilimselliğe bakmadan bir bağlanım bulunur. Sonra, kalıntılardan ikinci bağlanım hesaplanır.
- Eğer β anlamlıysa bu farklıserpilimselliğin göstergesidir. Anlamlı değilse, aynıserpilimsellik sıfır önsavı reddedilmez.

Glejser Sınaması

- H. Glejser (1969) tarafından öne sürülmüş olan Glejser sınaması özünde Park sınamasına benzer.
- Glejser, $\hat{u}_i$ kalıntılarının mutlak değerlerini kullanan şu işlev biçimlerini önerir:

$$\begin{aligned} |\hat{u}_i| &= \beta_1 + \beta_2 X_i + v_{1i} \\ |\hat{u}_i| &= \beta_1 + \beta_2 \sqrt{X_i} + v_{2i} \\ |\hat{u}_i| &= \beta_1 + \beta_2 \tfrac{1}{X_i} + v_{3i} \\ |\hat{u}_i| &= \beta_1 + \beta_2 \tfrac{1}{\sqrt{X_i}} + v_{4i} \\ |\hat{u}_i| &= \sqrt{\beta_1 + \beta_2 X_i} + v_{5i} \\ |\hat{u}_i| &= \sqrt{\beta_1 + \beta_2 X_i^2} + v_{6i} \end{aligned}$$

- v_{ji} $(j = \{1, \ldots, 6\})$ burada hata terimini göstermektedir.
- Görgül olarak çekici görünmelerine karşın Park ve Glejser sınamalarının bazı sorunları da vardır.
- Öncelikle, sınama bağlanımlarında kullanılan hata terimi v_i'nin ortalaması sıfır olmayabilmekte ve bu v_i'ler özilinti ya da farklıserpilimsellik gösterebilmektedir.
- Diğer bir deyişle Park ve Glejser sınamalarının kendileri de SEK varsayımlarını sağlamayabilmektedir.

- Ayrıca, Glejser işlemindeki son iki işlev biçimi katsayılarda doğrusal-dışı olduğundan SEK ile tahmin edilemez.
- Öyleyse uygulamada bu iki sınamanın birer sorgulayıcı yöntem olarak kullanılması daha doğrudur.

Spearman Sıra İlintisi Sınaması

- *"Spearman sıra ilintisi"* (Spearman's rank correlation) şu şekilde tanımlanır:

  Sıra ilintisi

$$r_s = 1 - 6\left[\frac{\sum d_i^2}{n(n^2-1)}\right]$$

- d_i değeri burada i'inci kişi ya da olgunun iki farklı özelliğine verilen sıra numaraları arasındaki farkı göstermektedir.
- n ise sıralanan kişi ya da olgu sayısıdır.
- Örnek olarak, ders çalışma ve sınavlardaki başarı ilintisini bulmak için öğrenciler haftada kaç saat ders çalıştıklarına göre ve sınavda aldıkları puana göre sıraya sokulur. Daha sonra, iki sıralama arasındaki farkın kareleri hesaplanır.

$Y_i = \beta_1 + \beta_2 X_i + u_i$ ikili bağlanımını alalım. Farklıserpilimselliği Spearman sıra ilintisi ile sınamak için şu adımlar izlenir:

1. Veriler bağlanıma yakıştırılıp kalıntılar elde edilir.
2. $|\hat{u}_i|$ mutlak değerleri bulunur.
3. Hem $|\hat{u}_i|$ hem de X_i yöneyleri artan ya da azalan bir sıraya dizilir ve Spearman sıra ilinti katsayısı r_s hesaplanır.
4. Anakütle sıra ilintisi $\rho_s = 0$ ve $n > 8$ varsayımları altında, $(n-2)$ sd ile t dağılımına uyan şu istatistik hesaplanır:

$$t = \frac{r_s\sqrt{n-2}}{\sqrt{1-r_s^2}}$$

5. Eğer hesaplanan t değeri kritik t değerinden büyük ise, aynıserpilimsellik sıfır önsavı reddedilir.
6. Bağlanım modeli eğer birden çok X değişkeni içeriyorsa, r_s her bir X için ayrı ayrı hesaplanıp sınanmalıdır.

Goldfeld-Quandt Sınaması

$Y_i = \beta_1 + \beta_2 X_i + u_i$ ikili bağlanımını ele alalım. Ayrıca σ_i^2 ve X_i arasında $\sigma_i^2 = \sigma^2 X_i^2$ biçiminde bir ilişki olduğunu da varsayalım. Goldfeld-Quandt sınamasının adımları aşağıdaki gibidir:

1. X_i gözlemleri küçükten büyüğe doğru sıralanır.
2. Ortadaki c sayıda gözlem örneklemden çıkarılır.
3. Kalan $(n - c)$ gözlem ortadan iki eşit öbeğe bölünür.
4. İki öbek ayrı ayrı SEK bağlanımına yakıştırılır, KKT_1 ile KKT_2 elde edilir ve şu oran hesaplanır:
$$F = \frac{\mathrm{KKT}_2/\mathrm{sd}}{\mathrm{KKT}_1/\mathrm{sd}}$$
5. F dağılımına uyan bu istatistiğin pay ve payda sd'si aynıdır ve $[(n - c)/2] - k$'ye eşittir. Eğer hesaplanan değer kritik F'den büyükse, aynıserpilimsellik sıfır önsavı reddedilir.

- Goldfeld-Quandt sınamasında, ortadaki c sayıda gözlemin dışlanma amacı, küçük varyanslı öbek ile büyük varyanslı öbek arasındaki farkı keskinleştirmektir.
- Öyleyse c'nin nasıl seçileceği önemlidir.
- Örneklem büyüklüğü yaklaşık 30 iken c'nin 4 ve örneklem büyüklüğü 60 iken de c'nin 10 olması yeterli sayılmaktadır.
- Modelde birden fazla açıklayıcı değişken var ise, bunlar uygun olduğu düşünülen X'e göre sıralanır ya da sınama her bir X için ayrı ayrı yapılır.

Breusch-Pagan-Godfrey Sınaması

- Goldfeld-Quandt sınamasının bir sakıncası, sonuçların gözlemleri sıralamada kullanılan X değişkeninin seçimine bağlı olmasıdır.
- Bu da Breusch-Pagan-Godfrey sınaması ile giderilebilir.
- Aşağıdaki k değişkenli bağlanım modelini ele alalım:
$$Y_i = \beta_1 + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + u_i$$
- σ_i^2 hata varyansının, olasılıksal olmayan Z değişkenlerinin doğrusal bir işlevi olduğunu varsayalım:

$$\sigma_i^2 = \alpha_1 + \alpha_2 Z_{2i} + \cdots + \alpha_m Z_{mi}$$

- Z değişkeni olarak X'lerin tümü ya da birkaçı kullanılabilir.
- Eğer $\alpha_2 = \alpha_3 = \ldots = \alpha_m = 0$ ise $\sigma_i^2 = \alpha_1$ olur.
- Demek ki BPG sınaması, σ_i^2'nin sabit olup olmadığını anlamak için $\alpha_2 = \alpha_3 = \ldots = \alpha_m = 0$ varsayımını sınar.

Breusch-Pagan-Godfrey sınamasının adımları şöyledir:

1. Bağlanım SEK ile tahmin edilir ve kalıntılar elde edilir.
2. σ^2'nin EO tahmincisi $\tilde{\sigma}^2 = \sum \hat{u_i}^2/n$ değeri hesaplanır.
3. $p_i = \hat{u_i}^2/\tilde{\sigma}^2$ değişkeni oluşturulur.
4. p_i'lerin Z'lere göre bağlanımı bulunur:

$$p_i = \alpha_1 + \alpha_2 Z_{2i} + \cdots + \alpha_m Z_{mi} + v_i$$

5. Bağlanım kareleri toplamı (BKT) bulunarak şu hesaplanır:

$$\Theta = (\text{BKT})/2$$

6. u_i'nin normal dağıldığı ve aynıserpilimsellik varsayımı altında ve örneklem büyüklüğü sonsuza doğru artarken, Θ değeri de $(m-1)$ sd ile ki-kare dağılımına uyar.
7. Buna göre, hesaplanan Θ eğer kritik χ^2 değerini aşıyorsa aynıserpilimsellik sıfır önsavı reddedilir.

White Genel Farklıserpilimsellik Sınaması

$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$ üçlü bağlanımını ele alalım. White genel farklıserpilimsellik sınaması şöyle yapılır:

1. Veriler SEK bağlanımına yakıştırılır ve kalıntılar alınır.
2. Aşağıdaki yardımcı bağlanım hesaplanır:

$$\hat{u_i}^2 = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + \alpha_4 X_{2i}^2 + \alpha_5 X_{3i}^2 + \alpha_6 X_{2i} X_{3i} + v_i$$

   Kısaca kalıntı karelerinin X'ler, X'lerin kareleri ve çapraz çarpımlarına göre bağlanımı bulunur. İlk bağlanımda sabit terim olmasa bile burada sabit terim kullanılır.

3. Yardımcı bağlanıma ait R^2 ve örneklem büyüklüğü çarpılır:

$$R^2 \times n$$

4. Bu istatistik, yardımcı bağlanımdaki (sabit terim hariç) açıklayıcı değişken sayısı kadar sd ile χ^2 dağılımına uyar.

5. Eğer bulunan χ^2 değeri seçili anlamlılık düzeyindeki kritik değerden büyükse, aynıserpilimsellik sıfır önsavı reddedilir.

- Goldfeld-Quandt sınamasının sakıncası, gözlemlerin hangi X değişkenine göre sıraya sokulduğuna bağlı olmasıdır.
- BPG sınamasının sakıncası da hata teriminin normalliği varsayımına duyarlı olmasıdır.
- White sınaması ise hem normallik varsayımına dayanmaz hem de uygulama yönünden basittir.
- Ancak bu sınama da dikkatli uygulanmalıdır.
- Eğer modelde çok sayıda değişken varsa bunlar, bunların kareleri ve çapraz çarpımları serbestlik derecesini tüketir.
- Ayrıca bazı durumlarda test istatistiğinin anlamlı olmasının nedeni farklıserpilimsellik olmayıp, model belirtim hatası olabilmektedir.
- Öyleyse White sınaması farklıserpilimselliği, model belirtim hatasını ya da her ikisini birden sınamada kullanılabilir.

4.3 Farklıserpilimselliği Düzeltmek

- Farklıserpilimsellik, SEK tahmincilerinin yansızlık ve doğrusallık özelliklerini bozmamaktadır.
- Ancak bu tahmincilerin etkinlik yoksunluğu önsav sınama işlemlerini kuşkulu duruma sokar.
- Dolayısıyla düzeltici önlemlerin gerekli olduğu açıktır.
- Sorunu düzeltmede izlenecek yaklaşım, farklıserpilimsel σ_i^2 varyanslarının bilinip bilinmediğine bağlıdır.
- Eğer σ_i^2 biliniyorsa ağırlıklı en küçük kareler kullanılır.
- σ_i^2 bilinmediği zaman ise White varyansları ya da çeşitli veri dönüştürme işlemleri uygulanır.

4.3.1 Ağırlıklı En Küçük Kareler

- AEK yöntemini göstermek için aşağıdaki iki değişkenli ÖBİ'yi ele alalım:

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i, \qquad E({\hat{u}_i}^2) = \sigma_i^2$$

- Farklıserpilimsel σ_i^2 varyansları biliniyor olsun. Yukarıdaki denklemin her iki yanını ağırlık değişkeni $1/\sigma_i$ ile çarpalım:

$$\tfrac{Y_i}{\sigma_i} = \hat{\beta}_1^* \left(\tfrac{1}{\sigma_i}\right) + \hat{\beta}_2^* \left(\tfrac{X_i}{\sigma_i}\right) + \left(\tfrac{\hat{u}_i}{\sigma_i}\right)$$

- Yukarıdaki dönüştürmeli modelin hata teriminin varyansı artık sabit ve 1'e eşittir.
- AEK yöntemi, bu dönüştürmeli modelin SEK ile tahmin edilmesi demektir.
- İlk modelin sabit terimli, dönüştürmeli modelin ise sıfır noktasından geçen bir bağlanım olduğuna dikkat ediniz.

4.3.2 Verilerin Dönüştürülmesi

Düzeltmeli White Varyansları

- Gerçek σ_i^2 çoğu zaman bilinemez.
- Böyle durumlarda SEK tahmincilerinin *"farklıserpilimsellik tutarlı"* (heteroscedasticity consistent) White varyansları kullanılabilir.
- Farklıserpilimsellik tutarlı White ölçünlü hatalarına *"sağlam ölçünlü hatalar"* (robust standard errors) da denmektedir.
- Birçok ekonometri yazılımı, SEK ölçünlü hataları yanında sağlam ölçünlü hataları da vermektedir.
- Sağlam ölçünlü hatalar SEK ölçünlü hatalarına göre daha büyük ya da daha küçük olabilmektedirler.
- White sürecinin sakıncası ise bunun kavuşmazsal olarak geçerli (büyük örnekleme dayalı) bir süreç olmasıdır.
- Ayrıca, White tahmincileri farklıserpilimselliği düzeltecek şekilde dönüştürülen verilerle elde edilen tahminciler kadar etkin olamayabilmektedirler.

Verilerin Dönüştürülmesi

- Verilerin dönüştürülmesi işlemi, SEK kalıntıları kullanılarak farklıserpilimselliğin gösterdiği *"örüntü"* (pattern) biçiminin incelenmesine dayanır.
- Yöntemi açıklamak için ikili bağlanım modelini ele alalım:

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

- Hangi dönüştürme işleminin yapılacağına ve bunun hangi X değişkenine göre yapılacağına çizim yöntemi ya da Park ya da Glejser sınamaları sonucunda karar verilebilir.

$1/X_i$ Dönüştürmesi

- Hata varyansının X_i^2 ile doğru orantılı olduğunu varsayalım:

$$E(u_i^2) = \sigma^2 X_i^2$$

- Bu durumda ilk model X_i'ye bölünerek dönüştürülebilir:

$$\begin{aligned} \frac{Y_i}{X_i} &= \frac{\beta_1}{X_i} + \beta_2 + \frac{u_i}{X_i} \\ &= \beta_1 \frac{1}{X_i} + \beta_2 + v_i \end{aligned}$$

- Böylece dönüştürülen hata terimi v_i'nin varyansı sabit olur:

$$E(v_i)^2 = E\left(\frac{u_i}{X_i}\right)^2 = \frac{1}{X_i^2} E(u_i^2) = \sigma^2$$

- Artık dönüştürmeli modele SEK uygulanabilir. İlk modele dönmek için ise tahmin edilen model yeniden X_i ile çarpılır.

Karekök Dönüştürmesi

- Hata varyansının X_i ile doğru orantılı olduğunu varsayalım:

$$E(u_i^2) = \sigma^2 X_i$$

- Bu durumda ilk model $\sqrt{X_i}$'ye bölünerek dönüştürülebilir:

$$\begin{aligned} \frac{Y_i}{\sqrt{X_i}} &= \frac{\beta_1}{\sqrt{X_i}} + \beta_2\sqrt{X_i} + \frac{u_i}{\sqrt{X_i}} \\ &= \beta_1 \frac{1}{\sqrt{X_i}} + \beta_2\sqrt{X_i} + v_i \end{aligned}$$

- Burada $E(v_i^2) = \sigma^2$ olduğu, diğer bir deyişle v_i teriminin aynıserpilimselliği doğrulanabilir.
- β_1 ve β_2'yi tahmin etmek için sıfır noktasından geçen SEK bağlanımı kullanılır.
- Daha sonra, bağlanımı yorumlamak için tüm değişkenler $\sqrt{X_i}$ ile çarpılarak ilk modele dönülür.

$1/E(Y_i)$ Dönüştürmesi

- Hata varyansının Y_i'nin ortalama değerinin karesiyle ilişkili olduğunu varsayalım:

$$E(u_i^2) = \sigma^2[E(Y_i)]^2$$

- Bu durumda ilk model aşağıdaki gibi dönüştürülebilir:

$$\begin{aligned}\frac{Y_i}{E(Y_i)} &= \frac{\beta_1}{E(Y_i)} + \beta_2\frac{X_i}{E(Y_i)} + \frac{u_i}{E(Y_i)} \\ &= \beta_1\frac{1}{E(Y_i)} + \beta_2\frac{X_i}{E(Y_i)} + v_i\end{aligned}$$

- Ancak bu dönüştürme uygulanabilir değildir çünkü $E(Y_i)$ değerleri, bilinmeyen β_1 ve β_2'ye bağlıdır.
- Bu yüzden, $E(Y_i)$ yerine tahmincisi $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$ alınır. Örneklem büyüklüğü artarken $\hat{Y}_i$'lar da gerçek $E(Y_i)$'lere yakınsayacağı için, bu yöntem uygulamada yeterli olabilir.

Log Dönüştürmesi

- Aşağıda verilen alışıldık log dönüştürmesini ele alalım:

$$\ln Y_i = \beta_1 + \beta_2 \ln X_i + v_i$$

- Farklıserpilimsellik sorunu burada $Y_i = \beta_1 + \beta_2 X_i + u_i$ gibi bir modelde olduğu kadar önemli değildir.
- Bunun nedeni, log dönüştürmesinin verilerin ölçeğini daraltarak değişkenler arası farkı azaltmasıdır.
- Log dönüştürmesinin sağladığı diğer bir yarar da β_2 eğim katsayısının Y'nin X'e göre esnekliğini vermesidir.
- Bu iki özellik, log modellerinin uygulamalı ekonometride yaygın olarak kullanılmasının nedenlerindendir.
- Öte yandan, log dönüştürmesi yapılırken hata teriminin ne şekilde ele alınacağı konusuna özen gösterilmelidir.

Önemli Bazı Noktalar

Ele alınan dönüştürmelerle ilgili önemli bazı noktalar şunlardır:

- İkiden fazla değişken olduğu zaman, dönüştürme için hangi X'in seçileceği konusuna dikkat edilmelidir.
- Eğer Y ya da X değişkenlerinden bazıları sıfır ya da eksi değerli olursa, log dönüştürmesi uygulanamaz.
- Bu durumda tüm gözlemleri artı yapacak şekilde seçilen artı değerli bir k sayısından yararlanılabilir.
- Zaman zaman, değişkenler ilişkisiz olsalar bile bunların oranları arasında bir *"düzmece"* (spurious) ilinti oluşabilir. Örnek olarak, Y_i ve X_i ilişkisizken Y_i/X_i ve $1/X_i$ ilişkili olur.
- σ_i^2'ler bilinmeyip de çeşitli dönüştürmeler ile tahmin edildiği zaman t sınaması ve F sınaması gibi tüm sınama işlemleri yalnızca büyük örneklemlerde geçerlidir. Bu nedenle küçük örneklem tabanlı bulgular yorumlanırken dikkat edilmelidir.

Önümüzdeki Dersin Konusu ve Ödev

Ödev
Kitaptan *Bölüm 11* "Heteroscedasticity: What Happens if the Error Variance Is Non-constant?" okunacak.

Önümüzdeki Ders
Özilinti

UADMK Açık Lisans Bilgisi

A. Talha Yalta
TOBB Ekonomi ve Teknoloji Üniversitesi
Ekim 2011