Değişen Varyansın Düzeltilmesi

* Uyguladığımız DV testleri sonucunda modelde istatistiki olarak DV olduğu sonucuna varırsak, DV-nin düzeltilmesi ve modelin buna uygun olarak yeniden tahmin edilmesi gerekir.
* DV-nin düzeltilmesi için literatürde sıklıkla kullanılan 4 adet yöntem vardır: logaritmik transformasyon, DV Dirençli (Gürbüz) Standart Hata kullanımı, Ağırlıklandırılmış En Küçük Kareler (Weighted Least Squares - WLS) kullanımı, Tahmin Edilebilir Genelleştirilmiş En Küçük Kareler (Feasible Generalized Least Squares - FGLS) kullanımı
* Bu yöntemler içinde en basiti logaritmik transformasyondur. Fakat en sık kullanılanı dirençli standart hataların kullanımıdır.

① Logaritmik Transformasyon

* DV-nin modelde olması durumunda ilk uygulanabilecek yöntem tüm değişkenlerin logaritmasını (doğal logaritma) almaktır.

ana model → $y = B_0 + B_1 X_1 + B_2 X_2 + \varepsilon$ — logaritmik transformasyon

logaritmik model → $\ln y = \alpha_0 + \alpha_1 \ln X_1 + \alpha_2 \ln X_2 + u$

* Logaritmik transformasyon değişkenleri yeniden ölçeklendirdiği yani ölçek küçülttüğü için dönüşüm yapılmış modelde DV olmayabilir.
* Logaritmik transformasyon varyansı stabilize eder.

Not: $\ln(x) = 1 \rightarrow x = e^1$

$\ln(2) = x \rightarrow 2 = e^x$

ÖRNEK Şimdi bir çoklu doğrusal regresyon modeli üzerinden logaritmik transformasyonun DV-yi ortadan kaldırıp kaldırmadığına bakalım. Düzey formundaki ana modeli öncelikle Breusch Pagan testi ile DV için test edeceğiz daha sonra ana modeli logaritmik transformasyon ile tekrar hesaplayıp White testin ikinci versiyonu ile DV için test edeceğiz.

ana model $\longrightarrow$ $y = B_0 + B_1 X_1 + B_2 X_2 + B_3 X_3 + u$

ÖRF $\longrightarrow$ $\hat{y} = -21{,}77 + 0{,}0021 X_1 + 0{,}123 X_2 + 13{,}853 X_3$

$$\quad (29.475) \quad (0.006) \quad (0.013) \quad (9.010)$$

$$n = 88 \quad R^2 = 0{,}672$$

• Breusch-Pagan test ile DV varlığını araştırıyoruz. Yani yukarıda uyguladığımız modelden kalıntıları çekip $(y_i - \hat{y}_i = \hat{u}_i)$ onların karelerini alıp tüm bağımsız değişkenler üzerine regres ediyoruz.

$\longrightarrow$ yardımcı regresyon ÖRF

$$\hat{u}^2 = -5522.79 + 0{,}202 X_1 + 1{,}691 X_2 + 1041.76 X_3$$

$$\quad (3259.5) \quad (0.071) \quad (1.464) \quad (996.38)$$

$$n = 88 \quad R^2 = 0{,}1601$$

$$\boxed{F = \frac{R^2/k}{(1-R^2)/(n-k-1)}}$$

• Yardımcı regresyonda daha önceden gösterdiğimiz F yada LM testini uyguladığımızda $F = 5.34$ ve $LM = 14.09$ buluyoruz. Sadece f testi üzerinden gittiğimizde %5 anlamlılık düzeyinde $F_{\alpha, k, n-k-1} = 2.71$ ← F-kritik değeri.

• Sonuç olarak H_0 temel hipotezini red edip DV buluyoruz.

★ Şimdi ana modeli logaritmik transformasyona uğratalım ve bu modelde DV varlığını White testin ikinci versiyonu ile test etmeye çalışalım.

logaritmik model $\longrightarrow$ $\ln y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \varepsilon$

ÖRF $\rightarrow$ $\widehat{\ln y} = -1.30 + 0.17 \ln X_1 + 0.7 \ln X_2 + 0.04 \ln X_3$

$$\quad (0.651) \quad (0.038) \quad (0.093) \quad (0.028)$$

$$n = 88 \quad R^2 = 0.643$$

• White testin ikinci versiyonu ile DV varlığını araştırıyoruz. Yani yukarıda uyguladığımız logaritmik modelden kalıntıları çekip $(\ln y_i - \widehat{\ln y}_i = \hat{\varepsilon}_i)$ onların karelerini alıp $\hat{y}$ ve $\hat{y}^2$ (yani $\widehat{\ln y}$ ve $\widehat{\ln y}^2$) üzerine regres ediyoruz.

doğal logaritmasını aldığımız için artık $\widehat{\ln y}$ kullanmalıyız.

yardımcı regresyon ÖRF $\longrightarrow$ $\hat{\varepsilon} = 5.047 - 1.709\,\widehat{\ln y} + 0.145\,\widehat{\ln y}^2$
$\quad\quad(3.345)\quad(1.163)\quad(0.100)$

$n = 88 \quad R^2 = 0.03917$

$\boxed{LM = n.R^2}$

• Yardımcı regresyonda daha önceden gösterdiğimiz F ya da LM testini uyguladığımızda $LM = 3.447$ ve $F = 1.7325$ buluyoruz. Sadece LM testi üzerinden gittiğimizde %5 anlamlılık düzeyinde $\chi^2_{\alpha,k} = 5.99 \leftarrow$ LM testi için kritik değer.

• Sonuç olarak Ho temel hipotezini red edemiyoruz ve sabit varyans kararı veriyoruz.

★ Görüldüğü üzere logaritmik transformasyon modeldeki DV-nin düzeltilmesini sağlamış oldu.

② DV Dirençli (Gürbüz) Standart Hata Kullanımı

★ Yatay kesit verisi ile yapılan regresyonlarda değişen varyans varlığında takip edilebilen bir diğer yol ise DV Dirençli (Gürbüz) Standart Hatalarının kullanılmasıdır.

★ DV Dirençli (Gürbüz) Standart hataları $\longrightarrow$ dirençli standart hataları White Standart hataları

↳ Heteroskedasticity-robust-standard error

★ Değişen varyansın bilinmeyen bir formu mevcut iken yani $Var(u|X) = \sigma^2 h(x)$ ve $h(x)$ fonksiyonunun formu bilinmiyorken dirençli standart hataları kullanarak daha önceden geçersiz olan t ve F testlerini artık geçerli hale getirmek mümkün olacaktır.

↳ Not: DV durumunda SEKK parametre tahmincileri için hesapladığımız varyans formülleri geçersiz olacağı için t ve F testleri de geçersiz olacaktır. Bu konuyu daha önce ayrıntısı ile görmüştük.

→ Kısacası DV durumunda dirençli standard hataları kullanarak DV-den etkilenmeyen t ve F istatistikleri (hatta LM istatistiği) hesaplamak mümkün olacaktır.

★ Dirençli standart hatalar modelde DV olsun ya da olmasın her iki durumda da geçerlidir. Yani t ve F testleri güvenilirdir.

↳ Peki neden her zaman dirençli standart hataları kullanmıyoruz?

• Eğer DV yoksa ve artıklar normal dağılıyorsa, t istatistiği gözlem sayısına bakılmaksızın tam t dağılımı yapar.

• Dirençli standart hata kullanılarak elde edilen t-istatistiğinin t dağılımı yapabilmesi gözlem sayısının yeterince büyük olmasına bağlıdır.

• Kısacası dirençli standart hata kullanımı sadece büyük örneklem durumunda geçerlidir.

• Küçük örneklemlerde ise değişen varyansın yapısına (formuna) ilişkin belirli varsayımlar altında SEKK dışında daha etkin bir tahmin yönteminin (WLS ya da FGLS) kullanılması gerekir. Çünkü SEKK artık DESTE değildir.

→ Ayrıca unutulmamalıdır ki DV durumunda parametre tahmincilerinin güven aralıkları da geçersiz oluyordu. DV durumunda dirençli standart hataları kullanarak geçerli güven aralıkları elde edebiliriz. Tek yapmamız gereken güven aralığı formülünde bahse konu olan parametre tahmincisinin normal standart hatasını kullanmak yerine aynı parametre tahmincisine ait dirençli standart hatayı kullanmaktır.

★ Şimdi dirençli standart hataların ve ona bağlı t, F ve LM istatistiklerinin basit ve çoklu doğrusal regresyon için ayrı ayrı gösterelim

Basit Doğrusal Regresyon	Çoklu Doğrusal Regresyon
model → $y = B_0 + B_1 X + u$ (1)	model → $y = B_0 + B_1 X_1 + B_2 X_2 + \dots B_k X_k + u$ (2)

→ DV varsayımı haricinde tüm varsayımlar geçerli ←

sabit varyans varken

(3) $\boxed{Var(\hat{B}_1) = \dfrac{\sigma^2}{\Sigma(x_i-\bar{x})^2}}$ ⟷ (4) $\boxed{Var(\hat{B}_J) = \dfrac{\sigma^2}{SST_J(1-R_J^2)} = \dfrac{\sigma^2}{SSR_J}}$

Değişen varyans varken

(5) $\boxed{Var(\hat{B}_1) = \dfrac{\Sigma(x_i-\bar{x})^2 \sigma_i^2}{\left(\Sigma(x_i-\bar{x})^2\right)^2} = \dfrac{\Sigma(x_i-\bar{x})^2 \sigma_i^2}{SST_x^{\,2}}}$ ⟷ (6) $\boxed{Var(\hat{B}_J) = \dfrac{\Sigma \hat{r}_{iJ}^{\,2} \sigma_i^2}{SSR_J^{\,2}}}$

Dikkat: 3., 4., 5. ve 6. denklemlerdeki varyanslar henüz tahmin değil. Ekonometri I dersinden hatırlayın ki parametre tahmincilerinin varyansını tahmin edebilmek için öncelikle $\hat{\sigma}^2$-yi (ya da DV durumunda $\hat{\sigma}_i^2$-yi) tahmin etmeli ve yukarıdaki formüllerde σ^2 yerine (ya da DV durumunda σ_i^2 yerine) koymalıyız. Bu durumda varyansları tahmin etmiş oluruz.

Not: $SST_x = \Sigma(x_i-\bar{x})^2$

* Şimdi DV durumunda formülünü elde ettiğimiz varyans denklemlerini (5. ve 6. denklem) tahmin etmeye çalışalım.

* White (1980), 5. ve 6. denklemlerdeki varyansların ana modeldeki artıkları ($\hat{u}$) kullanarak tahmin edilebileceğini gösterdi.

* Aşağıdaki 7. ve 8. varyans tahmin formülleri hem sabit varyans hem de değişen varyans durumunda geçerli dirençli varyanslardır.

Basit Doğrusal Regresyon

(7) $$\boxed{Var(\hat{B}_1) = \frac{\sum (x_i-\bar{x})^2 \hat{u}_i^2}{SST_x^2}}$$

Çoklu Doğrusal Regresyon

(8) $$\boxed{Var(\hat{B}_J) = \frac{\sum \hat{r}_{iJ}^2 \hat{u}_i^2}{SSR_J^2}}$$

Dirençli Varyans ?!

Not: 7. ve 8. denklemlerdeki varyansların tahmin olduğunu unutmayın.

(9) $$\boxed{se(\hat{B}_1) = \sqrt{Var(\hat{B}_1)}}$$

(10) $$\boxed{se(\hat{B}_J) = \sqrt{Var(\hat{B}_J)}}$$

Dirençli Standart Hatalar

* Dirençli standart hataları elde ettikten sonra dirençli t-istatistiğini elde etmek çok kolay olacaktır.

$$t = \frac{\text{tahmin - hipotez değeri}}{\text{dirençli standart hata}} = \frac{\hat{B}_J - B_J}{se(\hat{B}_J)}$$

• Yani tek yapmamız gereken daha önceden bildiğimiz t-istatistiği formülünde yeni hesaplanan dirençli standart hataları kullanmaktır (normal standart hata yerine).

ÖR: model → $y = B_0 + B_1 x_1 + B_2 x_2 + u$

ÖRF → $\hat{y} = 0.321 + 0.213 x_1 - 0.198 x_2$

(0.100)	(0.055)	(0.058)
[0.109]	[0.053]	[0.058]

(…) → Normal standart hatalar

[…] → Dirençli standart hatalar (köşeli parantez ile verildi)

Dirençli standart hatalar normal standart hatalara göre bazen büyük bazen küçük çıkmış!

Not: DV durumunda, t-istatistiği ve güven aralığı hesaplamasında dirençli standart hatalar kullanılmalıdır.

* Dirençli t-istatistikleri elde ettiğimiz gibi dirençli F-istatistiği ve LM-istatistiği de mümkündür.

* Fakat dirençli F-istatistiğinin kapalı bir formu (belirli bir formu-formülü) yoktur. Bu nedenle elle çözmek zordur. Zaten bir çok istatistik programında dirençli standart hataları ve istatistikleri elde etmek için özel paketler olduğundan, araştırmacılar bu değerleri genellikle kendileri elle hesaplamazlar.

* Dirençli F-istatistiğinin aksine dirençli LM-istatistiğinin bulunması biraz daha fazla uğraşla mümkündür.

Dirençli Wald istatistiği de matrikslerle yapılır.

Not: LM testi aynen F testinin test ettiği durumlarda (modelin genelinin testi ya da çoklu kısıtların testi gibi) uygulanabilir.

derin!

* DV durumunda dirençli LM-istatistiğini hesaplamak için aşağıdaki ana modeli düşünelim.

n= gözlem sayısı

(11) ana model $\rightarrow$ $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + u$

• ve test etmek istediğimiz hipotez ise

(12) $H_0: \beta_1 = \beta_2 = 0$ olsun $\longrightarrow$ q= kısıt sayısı

• Eğer DV olmasaydı yukarıdaki hipotez LM testi ile şu aşamalarla test edilecekti:

1) Temel hipoteze göre olan kısıtlanmış modeli tahmin edin (yani yukarıda örnek olarak verilen hipotezdeki kısıtları modele uygulayın)

kısıtlanmış model $\longrightarrow$ $y = \delta_0 + \delta_1 X_3 + \varepsilon \longrightarrow \hat{\varepsilon}$ topla

Kısıtlanmış modeldeki artıkları ($\hat{\varepsilon}$) toplayın.

2) Kısıtlanmış modeldeki

3) $\hat{\varepsilon}$-leri tüm bağımsız değişkenler (kısıtlanmış olsun ya da olmasın) üzerine regres et.

$\hat{\varepsilon} = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + v \rightarrow R^2_{\hat{\varepsilon}}$ hesapla

$LM = n R^2_{\hat{\varepsilon}} \longrightarrow \chi^2_q \Rightarrow$ ki-kare dağılım yapar

α = anlamlılık seviyesinde eğer

$LM\text{-ist} > \chi^2_{\alpha,q}$ (kritik değer) $\Big\}$ H_0 red edilir ve model istatistiki olarak anlamlı sonucuna varılır.

- Fakat modelde DV varsa, biraz daha uğrasla, 11. denklemde verilen model 12. denklemde verilen hipotez için LM testi ile (dirençli LM testi) test edilebilir.

1) Temel hipoteze göre olan kısıtlanmış modeli tahmin edin (yani 12. denklemde verilen hipotezdeki kısıtları modele uygulayın)

kısıtlanmış model $\rightarrow y = \delta_0 + \delta_1 X_3 + \varepsilon \longrightarrow \hat{\varepsilon}$ topla

2) Kısıtlanmış modeldeki artıkları ($\hat{\varepsilon}$) toplayın.

3) Temel hipotezdeki kısıtlara dahil olan tüm bağımsız değişkenleri ayrı ayrı kısıtı olmayan tüm değişkenler üzerine regres edin ve artıkları toplayın.

$$x_1 = \alpha_0 + \alpha_1 X_3 + r_1 \longrightarrow \hat{r}_1 \text{ topla}$$

$$X_2 = \gamma_0 + \gamma_1 X_3 + r_2 \longrightarrow \hat{r}_2 \text{ topla}$$

4) 3. adımda elde ettiğiniz artıkları, 1. aşamadaki artıkla çarpıp yeni değişkenler oluştur ve bunları tüm değerleri 1 olan bağımlı değişken üzerine regres et.

Dikkat $\longrightarrow$ Sabit terim yok

$$1 = \theta_1 \underbrace{\hat{r}_1 \hat{\varepsilon}}_{\text{yeni değişken 1}} + \theta_2 \underbrace{\hat{r}_2 \hat{\varepsilon}}_{\text{yeni değişken 2}} \Rightarrow \text{SSR topla} \Rightarrow SSR_1 \text{ diye tanımlayalım}$$

$LM = n - SSR_1 \rightarrow \lambda^2_q \Rightarrow$ ki-kare dağılımı yapar

α = anlamlılık seviyesinde eğer $LM\text{-ist} > \lambda^2_{\alpha,q}$ (kritik değer) } Ho red edilir ve model istatistiki olarak anlamlı sonucuna varılır.

Dirençli LM istatistiği

n büyük olduğu sürece DV olsun ya da olmasın geçerlidir.!

<u>Not:</u> n = gözlem sayısı ; q : kısıt sayısı

③ Ağırlıklandırılmış En Küçük Kareler Yöntemi (WLS)

★ Weighted Least Squares (WLS) yöntemi olarak da bilinmektedir.

★ Bazen dirençli standart hataları kullanmak yerine modeli değişen varyanstan arındırarak yeniden tahmin etmek isteyebiliriz. Bunun için SEKK değil de WLS yöntemini kullanabiliriz.

★ DV durumunda WLS, SEKK-den daha etkin parametre tahmincileri sunar. Yani artık parametre tahmincileri sapmasız ve tutarlı olmalarının yanı sıra etkin de olurlar. Kısacası parametre tahmincileri DESTE olur.

→ WLS-de t ve F istatistikleri geçerlidir
yani t ve F istatistikleri
tam olarak t ve F dağılımı yapar.

★ WLS yöntemi değişen varyans formuna ilişkin bilgiyi gerektirir. Yani değişen varyansı oluşturan fonksiyonun tam olarak bilinmesi gerekir.

★ Örnek olarak aşağıdaki çoklu doğrusal regresyonu ele alalım.

(13) ana model → $y = B_0 + B_1 X_1 + B_2 X_2 + \ldots + B_k X_k + u$ → GDR1-6 ✓
↳ GDR7 ✗

• sadece GDR7 varsayımının (DV yok) sağlanmadığını ve DV-nin formunun aşağıdaki gibi olduğunu düşünelim

$$Var(u_i|X_i) = \sigma^2 h(x) = \sigma^2 h_i$$

• Burada $h(x) > 0$ x'lerin herhangi bir fonksiyonudur.
↳ Neden her zaman sıfırdan büyük olmalı?

• h(x)-in bilindiğini düşünelim. ⇒ yani tüm h_i'ler

• Şimdi, 13. denklemde verilen ana modeli belli bir değer kullanarak dönüştüreceğiz ve sonuç olarak dönüşüme uğrayan bu yeni model GDR1- GDR6 varsayımlarını sağlamasının yanı sıra GDR7 (DV YOK) varsayımını da sağlar.

• Dönüşüm için hangi değeri kullanalım?
↳ DV-ye neden olan h(x) fonksiyonunun kare kökünü kullanmayı deneyelim ve sabit varyansa ulaşıp ulaşamadığımıza bakalım.

* Dönüşümü $\sqrt{h(x)}$ terimi ile yapacağız yani model de her gördüğümüz değişken (hata terimi de dahil) $\sqrt{h(x)}$ -e böleceğiz. Bu işlemi her gözlem için yapacağımızdan indeksli olan versiyonu yani $\sqrt{h_i}$ 'yi kullanalım.

* Dönüşüme uğramış u_i-leri tanımlayalım ve beklenen değer ve varyanslarını inceleyelim.

$$\underbrace{u_i^*}_{\text{dönüştürülmüş hata terimi}} = \frac{u_i}{\sqrt{h_i}}$$

$$E[u_i^* | X] = E\left[u_i/\sqrt{h_i} \,\middle|\, X\right] = \frac{1}{\sqrt{h_i}} \underbrace{E[u_i | X]}_{=0 \rightarrow \text{GRR 5}} = 0$$

Dönüştürülmemiş modelde

$$Var(u_i | x_i) = \sigma^2 h_i = E[u_i^2 | X]$$

$$Var(u_i^* | X) = E[u_i^{*2} | X] = E\left[\left(u_i/\sqrt{h_i}\right)^2 \,\middle|\, X\right] = \left(\frac{1}{\sqrt{h_i}}\right)^2 E[u_i^2 | X] = \frac{1}{h_i}\sigma^2 h_i = \sigma^2$$

$\Rightarrow$ sabit varyans elde edildi yani GDR7 sağlanmış oldu.

* Modeldeki hata terimini (u_i), $\sqrt{h_i}$ ile dönüştürdüğümüzde tüm varsayımlar sağlandığına göre artık 13. denklemde verilen ana modeli $\sqrt{h_i}$ ile dönüştürebiliriz.

ana model indeks ile yazıldı $\rightarrow$ $y_i = B_0 + B_1 X_{i1} + B_2 X_{i2} + \dots + B_k X_{ik} + u_i$

dönüştürülmüş model $\rightarrow$ $\dfrac{y_i}{\sqrt{h_i}} = B_0 \dfrac{1}{\sqrt{h_i}} + B_1 \dfrac{X_{i1}}{\sqrt{h_i}} + B_2 \dfrac{X_{i2}}{\sqrt{h_i}} + \dots + B_k \dfrac{X_{ik}}{\sqrt{h_i}} + \dfrac{u_i}{\sqrt{h_i}}$

(14) yeni model $\rightarrow$ $y_i^* = B_0 X_{i0}^* + B_1 X_{i1}^* + B_2 X_{i2}^* + \dots + B_k X_{ik}^* + u_i^*$

$X_{i0}^* = \dfrac{1}{\sqrt{h_i}}$; $Var(u_i^* | X) = \sigma^2$; $E(u_i^* | X) = 0$

* 14. denklemde verilen yeni model tüm Gauss-Markov varsayımlarını sağlar. Parametre tahmincileri $\rightarrow$ DESTE ✓

$\hat{B}_0^*, \hat{B}_1^*, \hat{B}_2^*, \dots, \hat{B}_k^*$ $\rightarrow$ karışıklık olmaması için "*" ile belirtildi.

* Burada dikkat edilmesi gereken parametre tahmincilerinin yorumlaması 13. denklem üzerinden yapılırken. Bu tahmincilere ait, t ve F testlerini uygulamak için ise 14. modeldeki sonuçlar kullanılır. → Çünkü bunlar geçerli t ve F istatistikleridir.

* Diğer bir önemli nokta ise 14. denklemden (dönüştürülmüş) elde edilen parametre tahmincileri $\{\hat{B}_0^*, \hat{B}_1^*, \hat{B}_2^*, \ldots, \hat{B}_k^*\}$, 13. denklemden elde edilen parametre tahmincilerinden $\{\hat{B}_0, \hat{B}_1, \hat{B}_2, \ldots \hat{B}_k\}$ farklıdır.

* Not: 14. denklemdeki gibi dönüştürülmüş bir modele SEKK yöntemi uygulanırsa buna genel olarak Genelleştirilmiş En Küçük Kareler (Generalized Least Squares - GLS) yöntemi denir. Eğer GLS, değişen varyans amacıyla yapılırsa buna da WLS denir.
* Bu nedenle $\{\hat{B}_0^*, \hat{B}_1^*, \hat{B}_2^*, \ldots, \hat{B}_k^*\}$ parametre tahmincilerine bazen GLS tahmincileri de denilmektedir.

* GLS parametre tahmincileri DESTE özelliklerini sağlar. Ayrıca doğru standart hatalara sahip oldukları için t ve F istatistikleri geçerlidir.

Çünkü hi → x-e göre artan bir fonksiyon

* WLS-deki temel fikir hata varyansı büyük olan gözleme daha az ağırlık $\left(\frac{u_i}{\sqrt{h_i}}\right)$ verilmesi ve hata varyansı küçük olan gözleme daha fazla ağırlık verilmesidir. Böylelikle her gözlem için hata varyansı eşitlenebilir.
Böylelikle her gözleme eşit ağırlık verilir.

* SEKK-de ise her zaman her gözlem için 1-dir.
↳ Ağırlık her gözlem için 1-dir.

* Ağırlıklandırmayı daha iyi anlayabilmek için çoklu doğrusal regresyonda parametre tahmincilerini hesaplamak için kullandığımız amaç fonksiyonunu hatırlayalım. (13. denklemdeki model için)

amaç fonksiyonu →

$$\min \sum_{i=1}^{n} \hat{u}_i^2 = \sum_{i=1}^{n} \frac{1}{1}\left(y_i - \hat{B}_0 - \hat{B}_1 X_{i1} - \hat{B}_2 X_{i2} - \ldots - \hat{B}_k X_{ik}\right)^2 \quad (15)$$

amacımız artıkların karesini minimize etmek

• Bu amaç fonksiyonunu 14. denklemde belirtilen dönüştürülmüş model için yazalım

$$\min \sum_{i=1}^{n} \left(\frac{\hat{u}_i}{\sqrt{h_i}}\right)^2 = \sum_{i=1}^{n} \left(\frac{y_i}{\sqrt{h_i}} - \frac{\hat{B}_0}{\sqrt{h_i}} - \hat{B}_1 \frac{X_{i1}}{\sqrt{h_i}} - \hat{B}_2 \frac{X_{i2}}{\sqrt{h_i}} - \ldots - \hat{B}_k \frac{X_{ik}}{\sqrt{h_i}}\right)^2$$

$$(16) \quad \min \sum_{i=1}^{n} \frac{1}{h_i} \hat{u}_i^2 = \sum_{i=1}^{n} \frac{1}{h_i}\left(y_i - \hat{B}_0 - \hat{B}_1 X_{i1} - \hat{B}_2 X_{i2} - \ldots - \hat{B}_k X_{ik}\right)^2$$

* 16. denklemden görüldüğü gibi WLS amaç fonksiyonunu h_i ile ağırlıklandırıyor. Daha önce gösterdiğimiz hata terimi varyansının formunu hatırlayın:

$$Var(u_i|X_i) = \sigma^2 h_i$$ → yani amaç fonksiyonu $\frac{1}{h_i}$ ile ağırlıklandırılırken aslında tüm değişkenler h_i-nin kare kökü $(1/\sqrt{h_i})$ ile ağırlıklandırılmış.

* 15. denklemden de SEKK-nın tüm değişkenleri eşit (1 ile) ağırlıklandırdığı görülüyor. Bu yönüyle aslında SEKK, WLS-nin özel bir halidir.

Not: Dönüştürülmüş modelin (14. denklem ile verilmiş WLS) R^2-si uyum iyiliği ölçütü için kullanılamaz. Ancak test istatistiklerinin hesaplanmasında kullanılabilir.

Not: WLS yöntemini kullanabilmek için DV-nin formunun yani örneğimizdeki h_i-nin fonksiyonel formunun bilinmesi gerekir. Eğer bilinmiyorsa FGLS kullanılmalıdır.

Tahmin Edilebilir Genelleştirilmiş En Küçük Kareler Yöntemi (FGLS)

* Uygulamada değişen varyansın fonksiyonel formu (yani h_i-nin fonksiyonel formu) genellikle bilinmez.
* Bu durumda değişen varyansın yapısının (yani h_i) eldeki verilerle tahmin edilmesi gerekir.
* DV-nin formunun tahmin edilerek değişkenlerin aynı WLS-deki gibi ağırlıklandırması yöntemine FGLS yöntemi denir.
* $Var(u_i|X_i) = \sigma^2 h_i$ → bilinmez ve data ile tahmin edilir yani h_i yerine $\hat{h}_i$ kullanılarak ağırlıklandırma yapılırken.

* FGLS-de oldukça genel bir çerçeve sunan varyans formu aşağıdaki gibidir

(17) $$Var(u|X) = \sigma^2 \underbrace{\exp(\delta_0 + \delta_1 X_1 + \delta_2 X_2 + \dots + \delta_k X_k)}_{h(x)}$$

$$h(x) = \exp(\delta_0 + \delta_1 x_1 + \delta_2 x_2 + \dots + \delta_k x_k)$$

★ Not: h(x) formunda exponansiyonel fonksiyon kullanılmıştır çünkü varyansın her zaman pozitif ya da en azından sıfır olmalıdır. Fakat doğrusal modellerde tahmin edilen değerlerin (bizim örneğimizde $\hat{h}i$) negatif değerler almaması garanti değildir. Exponansiyonel fonksiyon kullanılarak negatif değerler almaması garanti edilmiş olur.

★ $h(x) = \exp(\delta_0 + \delta_1 x_1 + \delta_2 x_2 + \dots + \delta_k x_k)$

Burada δ_J parametrelerini bilmediğimiz için h(x) tahmin etmeye çalışıyoruz. Eğer bilseydik direkt olarak WLS uygulayabilirdik. Bu parametreleri bilmediğimiz için önce δ_J-leri tahmin edeceğiz daha sonra hi-yi tahmin edeceğiz yani $\hat{h}i$-yi bulacağız.

★ 17. denklemde verilen varyans formunu, varyansın temsilcisi olan u^2-yi kullanarak yeniden yazalım.

$$Var(u|x) = \sigma^2 \underbrace{\exp(\delta_0 + \delta_1 x_1 + \delta_2 x_2 + \dots + \delta_k x_k)}_{h(x)}$$

$$u^2 = \sigma^2 \exp(\delta_0 + \delta_1 x_1 + \delta_2 x_2 + \dots + \delta_k x_k) v$$

→ x-lere göre ortalaması 1 olan terim eklendi

doğal logaritmasını al

$$\ln u^2 = \underbrace{\ln\sigma^2 + \delta_0}_{=\delta_0} + \delta_1 x_1 + \delta_2 x_2 + \dots + \delta_k x_k + \underbrace{\ln v}_{=\varepsilon}$$

$$\ln u^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \dots + \delta_k x_k + \varepsilon$$

ana modelden elde edilen $\hat{u}^2$ kullan u^2-nun temsilcisi olarak

$$\boxed{\ln \hat{u}^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \dots + \delta_k x_k + \varepsilon} \quad (18)$$

→ $\hat{h}i$-yi bulabilmek için tahmin etmemiz gereken model budur. 18. modeli tahmin edip bağımlı değişkenin tahmin edilen değerlerini yani $\widehat{\ln \hat{u}_i^2}$ yi hesaplamamız gerekir. Daha sonra bu değerin exponansiyel fonksiyonunu alırsak $\hat{h}i$-yi elde ederiz.

$$(19) \quad \boxed{\hat{h}_i = \exp(\widehat{\ln \hat{u}_i^2})}$$

* $\hat{h}_i$'leri hesapladıktan sonra aynen WLS-de uyguladığımız prosedürleri bu sefer $\hat{h}_i$-leri kullanarak uygulayabiliriz.

↳ Yani tüm değişkenleri ağırlıklandırmak için $\frac{1}{\sqrt{\hat{h}_i}}$ ağırlıklarını kullanacağız yada amaç fonksiyonunu $\frac{1}{\hat{h}_i}$ ile ağırlıklandıracağız.

* $\hat{h}_i$-ler yerine WLS-deki gibi h_i-leri kullanabilseydik parametre tahmincileri sapmasız olacaktı, Fakat FGLS-de hem modeli hem de $\hat{h}_i$-leri aynı datayı kullanarak tahmin ettiğimiz için artık sapmalı olur. Buna rağmen FGLS-deki parametre tahmincileri hala tutarlıdır ve SEKK-ye göre asimptotik parametre tahmincileri olarak daha etkindir.

↳ FGLS-de parametre tahmincileri sapmalı olduğu için parametre tahmincileri artık DESTE değildir.

↳ Fakat DV durumunda ve gözlem sayısı yeterince büyükse SEKK yöntemine iyi bir alternatiftir.

↳ FGLS-de $n\uparrow$ ise t ve F istatistikleri tam olarak t ve F dağılımı yapar.

Not: Bu konuda DV-yi düzeltmek için 4 farklı yöntem gördük. Uygulamada en sık kullanılan yöntem dirençli standart hataların kullanılmasıdır.