

8. DEĞİŞEN VARYANS SORUNU (HETEROSCEDASTICITY)

8.1. Değişen Varyans Sorunu Nedir?

Matrislerle $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$

yani $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + u_i, \quad i = 1 \dots n$

genel doğrusal modelini ele alalım.

Hata terimi için yapılan varsayımlardan birisi $E(\mathbf{u}\mathbf{u}') = \sigma_u^2 \mathbf{I} = \sigma^2 \mathbf{I}$ (hata terimlerinin varyansı sabittir ve aralarındaki kovaryans sıfırdır) varsayımıdır:

$$E(\mathbf{u}\mathbf{u}') = \begin{bmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{bmatrix}$$

Değişen varyans sorunu ise hata terimlerinin varyansları birbirinden farklı olduğu durumda vardır:

$$\text{Var}(u_i) = E(u_i^2) \neq \sigma^2$$

Sorunun varlığı durumunda hata terimi varyans-kovaryans matrisi $\text{Var, Cov}(\mathbf{u}) = E(\mathbf{u}\mathbf{u}') = \sigma^2 \mathbf{I}_n$ şeklinde yazılamıyor fakat

$$\text{Var, Cov}(\mathbf{u}) = E(\mathbf{u}\mathbf{u}') = \boldsymbol{\Omega} = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{bmatrix}$$

olarak yazılabiliyor ise değişen varyans sorunu var demektir.

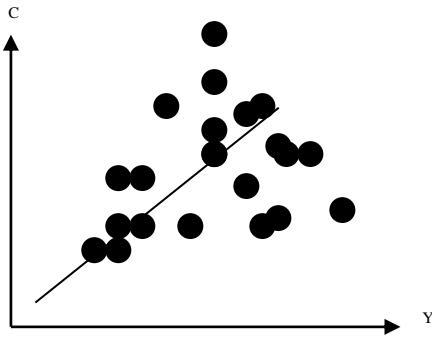
Her bir hata terimi için

$$E(u_i^2) = \sigma_i^2 \quad i = 1, 2, \dots, n$$

dir ve her bir hata terimi varyansı farklıdır.

Değişen varyans sorunu genellikle yatay kesit verileriyle tahmin yapıldığında ortaya çıkar. Hata terimlerinin varyanslarının değişken olmasının bazı nedenleri aşağıdaki gibidir.

- i) Hataların öğrenildiği durumlar: insanlar öğrendikleri için hataları zaman içinde azalır. Ör. Klavye kullanımı arttıkça yazım hataları sayısı azalır.
- ii) Gelir arttıkça insanların gelirini harcamak için daha fazla seçim alanı olur. Örneğin bağımlı değişken gıda harcamaları olsun ve açıklayıcı değişkenler bir sabit ve harcanabilir gelir olsun. Gıda için Engel eğrisinin pozitif eğimli olması beklenir. Yani ortalamada daha yüksek gelirlilerin gıda harcamalarının daha yüksek olması beklenir. Aynı zamanda yüksek gelirliler arasında harcama farklılıklarının düşük gelirliler arasında olduğundan daha yüksek olması beklenir. Dolayısıyla hata teriminin (u_i) varyansı gelire birlikte artar.



Gerçekleşen tüketim değerleriyle tahmin edilen (çizgi üzerindeki) tüketim değerleri arasındaki fark, hata terimi tahminini vermektedir ve gelir arttıkça bunlar da büyümektedir. Bu durumda hata teriminin varyansı da giderek artmaktadır.

Benzer biçimde firmaların karları veya firma büyüklüğü (çalışan sayısı) arttıkça kar payı dağıtımını gibi kararlarında daha fazla değişkenlik gösterirler.

- iii) Veri toplama teknikleri geliştikçe hata varyansları azalır. Ör. Daha gelişmiş veri işleme teknikleri olan bankaların müşterileri ile ilgili verdikleri bilgiler daha az hata içerir.
- iv) Aşırı uçlar (outliers): örneklemden farklı gözlemlere göre çok farklı olan gözlemler değişen varyansa neden olabilir.
- v) Spesifikasyon hataları olması durumunda, özellikle dışlanan değişken varsa değişen varyans sorunu ortaya çıkabilir.

vi) ARCH : Otoresif koşullu deęişen varyans
(Autoregressive Conditional Heteroscedasticity)

Deęişen varyans sorunu zaman serileriyle yapılan tahminlerde de ortaya çıkabilir. Özellikle enflasyon, hisse senedi fiyatları, döviz kurları gibi volatilitenin zaman içinde deęiştii verilerde gözlenmektedir. Hata terimi varyansları hem geçmiş dönemlerin hata terimleri ile ilişkilidir hem de dalgalanmalar gösterir.

Hata terimlerinin t dönemindeki koşullu varyansı:

$$E(u_t^2 \mid u_{t-1}^2, u_{t-2}^2, \dots, u_{t-p}^2) = \sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \dots + \alpha_p u_{t-p}^2 \quad (8.1)$$

Bu modelinin makul olması için (8.1) nin pozitif olması, bunun için de tüm katsayıların pozitif olması gerekir. Çünkü varyans negatif değer alamaz.

(8.1) deki genel model ARCH(p) sürecini yansıtır. Bu durumda örneğin ARCH(1)

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 \text{ dir.}$$

8.2. Değişen Varyans Sorunu EKK Tahmin Edicilerini Nasıl Etkiler?

1. $E(\mathbf{u}) = 0$ ve $E(\mathbf{X}'\mathbf{u}) = 0$ varsayımları geçerli olmayı sürdürdükleri için sapmasızlık özelliğini korunur.

2. EKK tahmin edicisi etkinlik özelliğini kaybeder. Çünkü

$$\begin{aligned}\text{Var,Cov}(\hat{\boldsymbol{\beta}}) &= E\{[\hat{\boldsymbol{\beta}}-E(\hat{\boldsymbol{\beta}})][\hat{\boldsymbol{\beta}}-E(\hat{\boldsymbol{\beta}})]'\} \\ &= E\{[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}-E((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u})][(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}-E((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u})]'\} \\ &= E\{[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}][(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}]'\} \\ &= E\{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\mathbf{u}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\} \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\mathbf{u}\mathbf{u}')\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\sigma^2\boldsymbol{\Omega}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \\ &= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Omega}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\end{aligned}$$

Bu durumda değişen varyans sorunu daha büyük varyansa neden olur ve daha küçük varyansa sahip tahmin ediciler vardır.

$$(E(\mathbf{u}\mathbf{u}') = \sigma^2\mathbf{I}_n \text{ olsaydı } \text{Var,Cov}(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{I}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} = \sigma^2(\mathbf{X}'\mathbf{X})^{-1} \text{ olurdu})$$

3. EKK tahmin edicisinin hesaplanan varyans ve standart hataları yanlış bir ifadeye dayanacak, dolayısıyla t ve F istatistikleri sapmalı olacak, testler güvenilir olmaktan çıkacaktır:

Hata terimlerinin varyansının (σ_u^2) tahmin edicisi $\frac{\sum \hat{u}_i^2}{n-k}$ aşağı doğru sapmalı olur. Dolayısıyla $\text{Var}(\hat{\beta}_j)$ aşağı doğru, t istatistikleri yukarı doğru sapmalı olur.

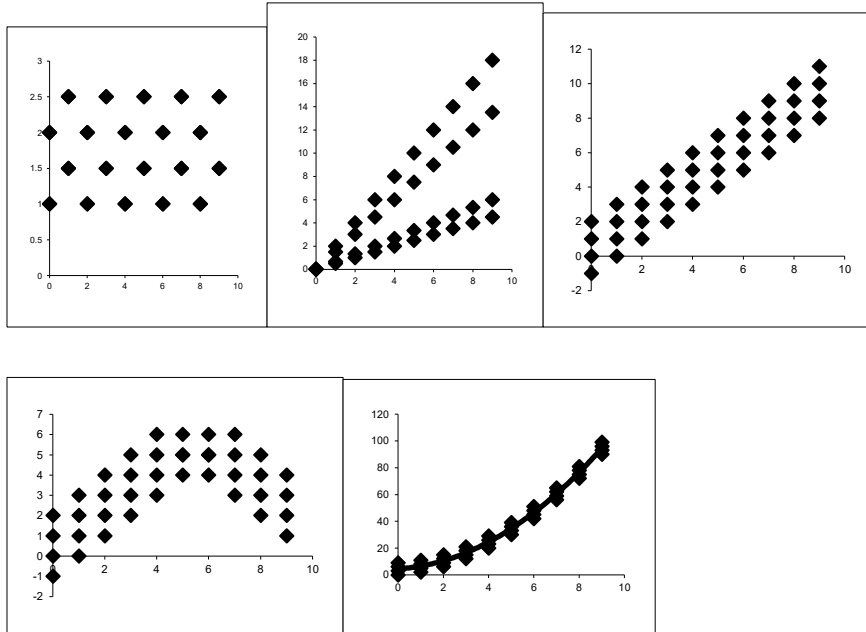
Diğer yandan R^2 ve F istatistiği de yukarı doğru sapmalıdır.

8.3. Değişen Varyans Sorununun Varlığı Saptanabilir mi?

8.3.1. Biçimsel Olmayan Yöntemler

Elimizdeki çalışmanın niteliği bir ipucu verebilir: Daha önceki çalışmalar göstermiştir ki tüketicimin gelirle açıklandığı tahminlerde değişen varyans sorunu vardır. Bu nedenle tüketim denklemi tahmininde bu sorunun olmasını bekleriz. Kesit verileri tahminlerinde heterojen birimler varsa bu sorun sözkonusu olacaktır: örneğin yatırımların bağımlı çıktı, faiz vs. nin açıklayıcı değişken olduğu denklemde küçük, orta ve büyük ölçekli işletmeler aynı örnekleme yer alıyorsa bu sorundan şüphelenmeliyiz.

Grafik yöntemi: Değişen varyansla ilgili önceden elimizde bilgi yoksa hata tahmin karelerinin grafiği incelenerek sistematik bir şekil verip vermediğine bakılabilir. Hata tahminleri hata terimleri ile aynı değildir fakat özellikle örneklem büyüklüğü yeterince genişse iyi bir tahminini verir. Dikey ekseninde hata tahmin kareleri, yatay ekseninde Y'nin tahmin değerleri varken:



Birinci grafikte \hat{Y}_i ile \hat{u}_i^2 arasında sistematik bir ilişki görünmemektedir. Ama diğerlerinde vardır: örneğin 3. de doğrusal bir ilişki 4. ve 5. de karesel bir ilişki vardır.

Grafik özellikle 1. grafikte olduğu gibi bir ilişki göstermiyorsa yatay ekseninde açıklayıcı değişkenlerden birinin olduğu grafik de kullanılabilir.