HAFTA 6

DEĞİŞEN VARYANS (HETEROSCEDASTICITY)

Klasik doğrusal regresyon modelinin önemli varsayımlarından birisi hata terimleri ε_i 'lerin sabit varyanslı olduğudur. Bu varsayımın sağlanmadığı durumda

1. Değişen varyansın niteliği nedir?
2. Doğurduğu sonuçlar nelerdir?
3. Düzeltici önlemler nelerdir?
4. Düzeltici önlemler nelerdir?

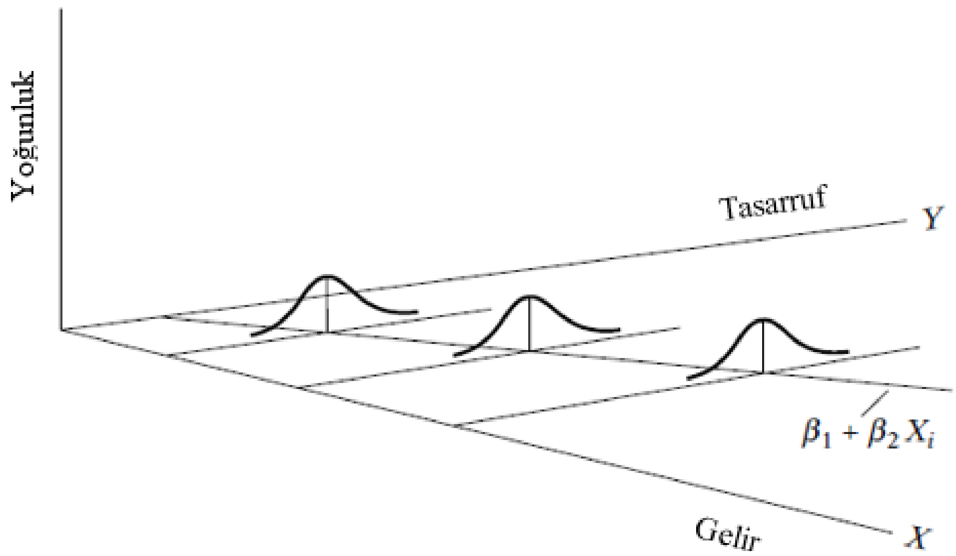
sorularına cevap aranacaktır.

Değişen Varyansın Niteliği:

Sabit varyans (homoscedasticity) ya da eşit (homo), yayıklık (scedasticity) eşit varyans varsayımıdır.

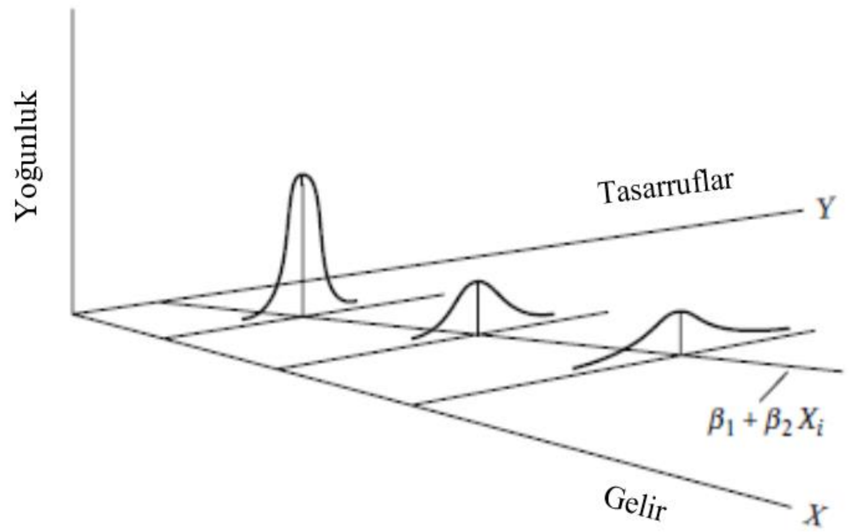
$Var(\varepsilon_i) = \sigma^2; \;\; i = 1, 2, \ldots, n$

Şekil: Sabit varyans dağılımı



X_i 'lere bağlı olarak Y_i 'nin koşullu varyansı X değişkeni hangi değerleri alırsa alsın, aynı kalır.

Şekil: Değişen varyans dağılımı

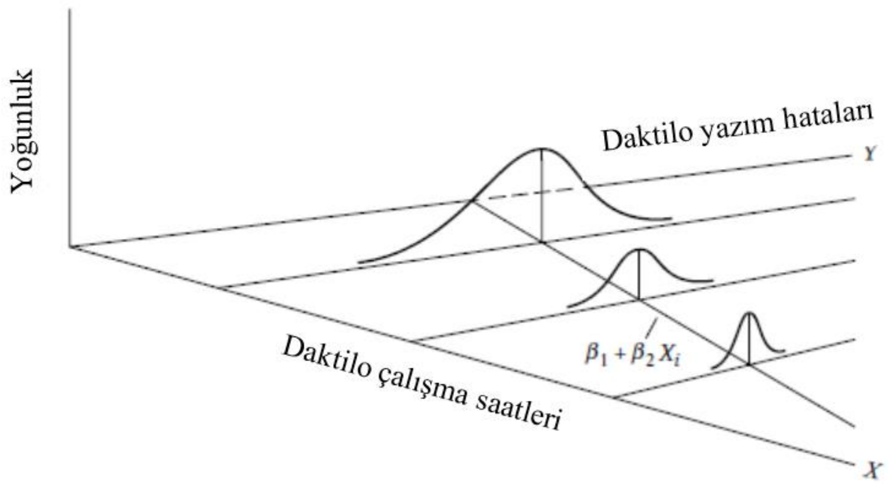


X büyüdükçe Y 'nin koşullu varyasının da büyüdüğü görülüyor. Öyleyse bu değişen varyans durumudur.

$Var(\varepsilon_i) = \sigma_i^2; \;\; i = 1, 2, \ldots, n$

Varyansın değişken olmasının nedenleri:

1. Hatasını öğrenen modellerde davranış hataları zamanla azalır. Bu durumda σ^2 'nin küçülmesi beklenir.
Şekil: Değişen varyansın gösterimi



Daktilo çalışma saatlerinin sayısı arttıkça, hem daktilo hataları hem de bunların varyansı azalmaktadır.

2. Gelir yükseldikçe gelirlerin harcanabileceği seçenekler genişler. Böylelikle σ_i^2 'nin gelirle birlikte büyümesi beklenir.

3. Veri derleme teknikleri geliştikçe σ_i^2 'de düşebilir.
4. Değişen varyans örneklemdeki öteki gözlemlerden çok farklı gözlemlerin varlığının bir sonucu olarak da ortaya çıkabilir.
5. Değişen varyasın bir başka nedeni modelin yanlış kurulması ve bazı önemli değişkenlerin modelden dışlanmasıdır.

Değişen varyans varken EKK tahmini:

Model: $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad Var(\varepsilon_i) = \sigma_i^2;\ i = 1,2,\ldots,n$ olmak üzere varyans-kovaryans matrisi $Cov(\boldsymbol{\varepsilon}) = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix} = V$ dir.

En küçük kareler tahmin edicisi:

$$\hat{\beta} = \left(X'X\right)^{-1} X'Y \Rightarrow \hat{\beta}_1 = \frac{S_{XY}}{S_X^2} = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2} \underset{\substack{x_i = X_i - \bar{X} \\ y_i = Y_i - \bar{Y}}}{=} \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$$

varyansı

$$Var(\hat{\beta}_1) = Var\left(\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}\right) = \frac{\sum_{i=1}^{n} x_i Var(y_i)}{\left(\sum_{i=1}^{n} x_i^2\right)^2} = \frac{\sum_{i=1}^{n} x_i \sigma_i^2}{\left(\sum_{i=1}^{n} x_i^2\right)^2}$$

dir. Her i için $\sigma_i^2 = \sigma^2$ olsaydı

$$Var(\hat{\beta}_1) = \frac{\sum_{i=1}^{n} x_i \sigma^2}{\left(\sum_{i=1}^{n} x_i^2\right)^2}$$

olacaktır. Değişen varyans durumunda en küçük kareler tahmin edicisi hala BLUE mudur? $\hat{\beta}_1$ hala doğrusal ve sapmasız mıdır? $\hat{\beta}_1$ artık en iyi değildir ve min. varyansı da vermez.

Genelleştirilmiş en küçük kareler yöntemi:

Değişen varyans durumunda bulunan EKK tahmin edicisi sapmasız ve minimum varyansı vermediğine göre bu durumu düzeltmek ve sabit varyansı sağlamak için dönüşüm yapılabilir.

$Var(\varepsilon_i) = \sigma_i^2 \Rightarrow Var\left(\frac{\varepsilon_i}{\sigma_i}\right) = 1$ olacaktır. Bu durumda

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

$$\underbrace{\frac{Y_i}{\sigma_i}}_{Y_i^*} = \underbrace{\beta_0\left(\frac{1}{\sigma_i}\right)}_{\beta_0^*} + \beta_1\underbrace{\left(\frac{X_i}{\sigma_i}\right)}_{X_i^*} + \underbrace{\left(\frac{\varepsilon_i}{\sigma_i}\right)}_{\varepsilon_i^*} \Rightarrow Y_i^* = \beta_0^* + \beta_1 X_i^* + \varepsilon_i^*$$

olacak ve $Var(\varepsilon_i^*) = 1$ sabit varyans varsayımı sağlanmış olacaktır. Dönüştürülmüş bu model diğer varsayımları korumaktadır. Modeldeki değişkenlere model varsayımlarını sağlayacak şekilde dönüşüm yapıldıktan sonra en küçük kareler yöntemini uygulamaya *genelleştirilmiş en küçük kareler yöntemi* (GEKK) denir. Bu yolla bulunan parametre tahminlerine de *genelleştirilmiş en küçük kareler tahmin edicisi* denir.

Not: Yukarıdaki modelde varyans sabitlemek için yapılan dönüşüm matrislerle ifade edilecek olursa, önce hataların varyans-kovaryans matrisi

$$Cov(\boldsymbol{\varepsilon}) = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix} = V$$

olmak üzere öyle bir Γ matrisi tanımlanır ki $V = \Gamma'\Gamma$ ve $V^{-1} = \Gamma^{-1}\left(\Gamma'\right)^{-1}$ olsun. Bu koşulları sağlayan matris $\Gamma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n \end{bmatrix}$ dir. Bu Γ matrisi ile

$Y = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ modelinde ağırlıklandırma yapılırsa,

$$\underbrace{\Gamma^{-1}Y}_{Y^*} = \underbrace{\Gamma^{-1}\mathbf{X}}_{\mathbf{X}^*}\boldsymbol{\beta} + \underbrace{\Gamma^{-1}\boldsymbol{\varepsilon}}_{\boldsymbol{\varepsilon}^*} \Rightarrow Y^* = \mathbf{X}^*\boldsymbol{\beta} + \boldsymbol{\varepsilon}^*$$

klasik lineer modeline dönecektir. EKK tahmin edicisi ise

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^{*\prime}\mathbf{X}^*\right)^{-1}\mathbf{X}^{*\prime}Y^*$$

olacaktır. Buradan açıkça yazılırsa;

$$\hat{\boldsymbol{\beta}} = \left(X'\underbrace{(\Gamma^{-1})'\Gamma^{-1}}_{V^{-1}}X\right)^{-1} X'\underbrace{(\Gamma^{-1})'\Gamma^{-1}}_{V^{-1}}Y = \left(X'V^{-1}X\right)^{-1} X'V^{-1}Y$$

olarak elde edilir. $\hat{\boldsymbol{\beta}}$ ’nın varyans-kovaryans matrisi

$$Cov(\hat{\boldsymbol{\beta}}) = \left(X'V^{-1}X\right)^{-1} X'V^{-1}\underbrace{Cov(Y)}_{V}V^{-1}X\left(X'V^{-1}X\right)^{-1}$$

$$= \left(X'V^{-1}X\right)^{-1} X'V^{-1}\underbrace{VV^{-1}}_{I}X\left(X'V^{-1}X\right)^{-1}$$

$$= \left(X'V^{-1}X\right)^{-1} \underbrace{X'V^{-1}X\left(X'V^{-1}X\right)^{-1}}_{I} = \left(X'V^{-1}X\right)^{-1}$$

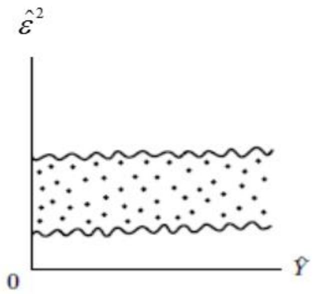
dir. Genelleştirilmiş en küçük karelerde de $w_i = \dfrac{1}{\sigma_i}$ ile ağırlıklandırılmış hata kareler toplamını (WSSE) en küçüğe indirirken EKK de ağırlıklandırılmamış ya da eşit ağırlıklandırılmış SSE'yi en küçüğe indirir.

Değişen varyansın varlığı:

– **<u>Biçimsel olmayan yöntemler:</u>**

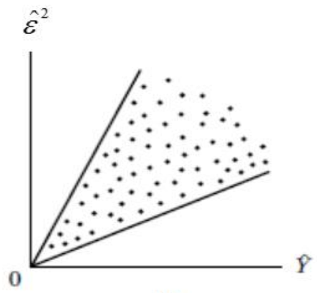
a) <u>Sorunun niteliği</u>: Yatırım harcamalarının satışlar, faiz oranları gibi değişkenlerle ilgisini konu alan kesit çözümlemelerinde eğer küçük, orta, büyük firmalar bir arada örneklenmiş ise genellikle değişen varyans beklenir.

b) <u>Çizim yöntemi:</u> Değişen varyansa ilişkin önsel ve görsel bilgi yoksa modelden tahmin edilen $\hat{\varepsilon}_i$ hatalar bulunur. Daha sonra $\hat{\varepsilon}_i^2$ 'lerin nasıl bir yapı sergilediğine bakılır.



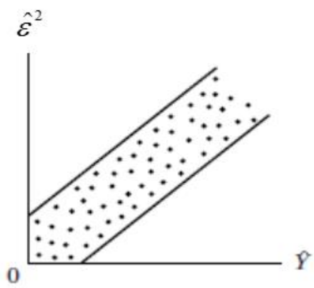
(a)

İki değişken arasında düzenli bir örüntü yok ve değişen varyans bulunmamaktadır.



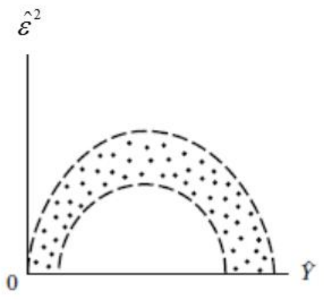
(b)

Belirli bir örüntü var. Artan varyans bulunmaktadır.



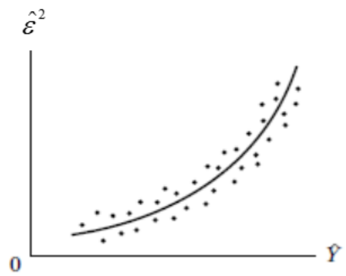
(c)

Belirli bir örüntü, Doğrusal ilişki ve Değişen varyans bulunmaktadır.



(d)

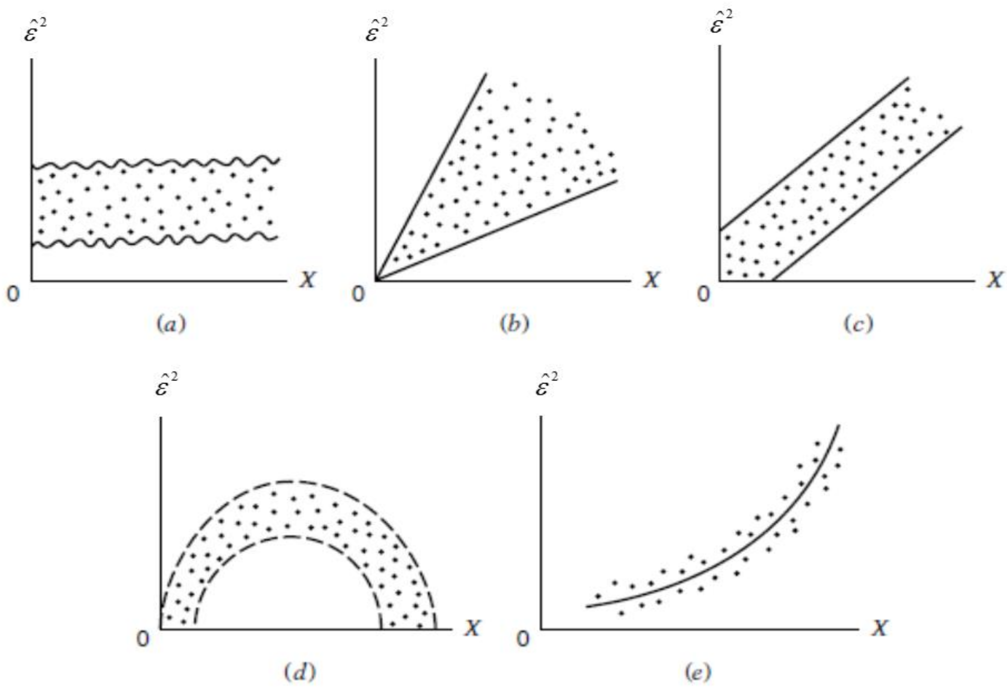
$\hat{\varepsilon}^2$ ile $\hat{Y}$ arasında ikinci dereceden bir ilişki var. Değişen varyans sergilemeyecek biçimde dönüşüm yapılır.



(e)

$\hat{\varepsilon}^2$ ile $\hat{Y}$ arasında ikinci dereceden bir ilişki olması sebebiyle değişen varyans sergilemeyecek biçimde dönüşüm yapılır.

Artık kareler $\hat{\varepsilon}_i^2$ ile $\hat{Y}_i$ kestirimleri arasındaki ilişki yerine $\hat{\varepsilon}_i^2$'nin X'lerden birisiyle ilişkisi çizilebilir. Örneğin $\hat{\varepsilon}_i^2$ ile X arasındaki grafik şekil (c)'deki gibi ise $\hat{\varepsilon}_i$'nin X ile doğrusal ilişkili olduğunu gösterir. Bu bilgide varyansı sabitlemede kullanılabilir.



$\hat{\varepsilon}^2$
0
X
(a)
$\hat{\varepsilon}^2$
0
X
(b)
$\hat{\varepsilon}^2$
0
X
(c)
$\hat{\varepsilon}^2$
0
X
(d)
$\hat{\varepsilon}^2$
0
X
(e)