1

Correlation, Heteroscedasticity, GLS

H-tests, p-vals, CIs, t-stats: rely on $SE(\hat{\beta}_j)$

Recall, $SE(\hat{\beta}_j) = \widehat{SD}(\hat{\beta}_j) = \left(V(\hat{\beta}_j) \right)^{1/2}$

How did we get $V(\hat{\beta}_j)$?

<u>Def</u>: $V(\underset{\sim}{\hat{\beta}}) = E[\underset{p \times 1}{(\hat{\beta} - \beta)} \underset{1 \times p}{(\hat{\beta} - \beta)^T}]$

$$\begin{aligned}\hat{\beta} &= (X^T X)^{-1} X^T y \\ &= (X^T X)^{-1} X^T [X\beta + \epsilon] \\ &= \beta + (X^T X)^{-1} X^T \epsilon \quad \Rightarrow \quad \hat{\beta} - \beta = (X^T X)^{-1} X^T \epsilon\end{aligned}$$

$$\begin{aligned}V(\hat{\beta}) &= E[(X^T X)^{-1} X^T \epsilon \epsilon^T X (X^T X)^{-1}] \\ &= (X^T X)^{-1} X^T E[\epsilon \epsilon^T] X (X^T X)^{-1} \\ &= (X^T X)^{-1} X^T \underline{V(\epsilon)} X (X^T X)^{-1}\end{aligned}$$

Now recall, in the model $\underset{n \times 1}{\underset{\sim}{y}} = \underset{n \times p}{X} \underset{p \times 1}{\beta} + \underset{n \times 1}{\underset{\sim}{\epsilon}}$, $E(\underset{\sim}{\epsilon}) = 0$

<u>Def</u>: $$\begin{aligned}V(y) &= E[(y - E(y)) (y - E(y))^T] \\ &= E((y - X\beta)(y - X\beta)^T) \\ &= E(\epsilon \epsilon^T) \\ &= E((\epsilon - 0)(\epsilon - 0)^T) \\ &= V(\epsilon)\end{aligned}$$

So $V(y) = V(\epsilon)$.

Up until now, we have assumed $V(\epsilon) = \sigma^2 I \Rightarrow$ "iid errors"

$$V(\epsilon) = \sigma^2 I \Rightarrow V(\hat\beta) = (X^T X)^{-1} X^T (\sigma^2 I) X (X^T X)^{-1}$$
$$= \sigma^2 (X^T X)^{-1} X^T X (X^T X)^{-1}$$
$$= \underline{\underline{\sigma^2 (X^T X)^{-1}}}$$

If $V(\epsilon) = \Sigma$

$$\begin{cases} \Sigma = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & \ddots & \\ \rho^2 & & \ddots \end{pmatrix} & \text{autocorrelation} \\ \Sigma = \begin{pmatrix} v_1 & & 0 \\ & \ddots & \\ 0 & & v_n \end{pmatrix} \quad v_i = f(x_i^T \beta) & \underline{\text{mean-var relat}} \\ \Sigma = \sigma^2 \begin{pmatrix} 1/m_1 & & 0 \\ & \ddots & \\ 0 & & 1/m_n \end{pmatrix} & \text{different measurement variances.} \end{cases}$$

then in general, $V(\hat\beta) = \underline{\underline{(X^T X)^{-1} (X^T \Sigma X) (X^T X)^{-1}}}$

How to get standard errors?

1) estimate Σ somehow from $\hat{\epsilon} = y - \hat{y}$ (will depend on covariance model)

2) Compute $\widehat{V(\hat\beta)} = (X^T X)^{-1} X^T \hat\Sigma X (X^T X)^{-1}$

3) obtain $SE(\hat\beta_j) = \left(\widehat{V(\hat\beta)}_{jj}\right)^{1/2}$

Typically need to do this "by hand".

Generalized least squares (includes weighted least squares)

Suppose $\underset{\sim}{y} = X\beta + \underset{\sim}{\epsilon}$, $E[\underset{\sim}{\epsilon}] = 0$, $V(\underset{\sim}{\epsilon}) = \Sigma$, Σ known

An aside: if $V(\epsilon) = \sigma^2$ (scalar), $V((\sigma^2)^{-1/2}\epsilon) = V(\epsilon/\sigma) = \frac{1}{\sigma^2}\sigma^2 = 1$

similarly, if $V(\underset{\sim}{\epsilon}) = \Sigma$, $V(\Sigma^{-1/2}\underset{\sim}{\epsilon}) = \Sigma^{-1/2}\,\Sigma\,\Sigma^{-1/2} = \underline{\underline{I}}$

$$\underset{\sim}{y} = X\beta + \epsilon$$

$$\Sigma^{-1/2}y = \Sigma^{-1/2}X\beta + \Sigma^{-1/2}\epsilon$$

$$\tilde{y} = \tilde{X}\beta + \tilde{\epsilon}, \quad \underline{\underline{E[\tilde{\epsilon}] = 0}}, \quad V(\tilde{\epsilon}) = V(\Sigma^{-1/2}\epsilon) = \underline{\underline{I}}$$

OLS estimate based on $\tilde{y}, \tilde{X}$ is

$$\hat{\beta} = (\tilde{X}^T\tilde{X})^{-1}\tilde{X}^T\tilde{y}$$

$$= (X^T\Sigma^{-1/2}\Sigma^{-1/2}X)^{-1}X^T\Sigma^{-1/2}\Sigma^{-1/2}y$$

$$= \underline{\underline{(X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}y}} = \underline{\underline{\hat{\beta}_{GLS}}}$$

$\hat{\beta}_{GLS}$ is the generalized least-squares estimator of β.

Note: ① if $\Sigma = I\sigma^2$, $\hat{\beta}_{GLS} = \hat{\beta}_{OLS}$

if $\Sigma \neq I\sigma^2$, $\hat{\beta}_{GLS} \neq \hat{\beta}_{OLS}$

② $V(\hat{\beta}_{OLS}) \geq V(\hat{\beta}_{GLS})$ in general, $\hat{\beta}_{GLS}$ is the BLUE "best linear unbiased estimator"