Time Series Errors Derivations

Time Series Residuals

Model:

$$y_t = \beta_0 + \beta_1 x_t + e_t$$

Sample Estimates:

$$y_t = \hat{\beta}_0 + \hat{\beta}_1 x_t + \hat{e}_t$$

- Where do standard errors for $\hat{\beta}_0$, $\hat{\beta}_1$ come from?

$$\widehat{se(\hat{\beta}_1)} = \sqrt{\frac{\hat{\sigma}_e^2}{T \cdot \hat{\sigma}_x^2}}$$

This Lecture

- Derivation and explanation of different standard errors:
  1. OLS standard errors ("plain vanilla").
  2. Heteroskedasticity-consistent (HC), robust, White standard errors.
  3. Heteroskedasticity- and autocorrelation-consistent (HAC) or Newey-West standard errors.

Standard Errors in General

$$y_t = \hat{\beta}_0 + \hat{\beta}_1 x_t + \hat{e}_t$$

Stack y_t in a column vector, x_t in a matrix, β_0, β_1 in a vector. The OLS estimate $\underline{\widehat{\beta}}$ for

$$\mathbf{y} = \mathbf{X}\underline{\beta} + \mathbf{e}$$

is (show aside in video):

$$\begin{aligned}
\underline{\widehat{\beta}} &= \frac{Cov(\mathbf{X},\mathbf{y})}{Var(\mathbf{X})} \\
&= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\
&= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\underline{\beta} + \mathbf{e}) \\
&= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\underline{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e} \\
&= \underline{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e}
\end{aligned}$$

Standard Errors in General

Recall

$$\underline{\widehat{\beta}} - \underline{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e}$$

The variance of $\underline{\widehat{\beta}}$ is

$$\begin{aligned} Var(\underline{\widehat{\beta}}) = E\left[\left(\underline{\widehat{\beta}} - \underline{\beta}\right)^2 |\mathbf{X}\right] &= E\left[\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e}\right)^2 |\mathbf{X}\right] \\ &= E\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e}\mathbf{e}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}|\mathbf{X}\right] \end{aligned}$$

"Sandwich Estimator" for the variance: the $(\mathbf{X}'\mathbf{X})^{-1}$'s are the bread, $\mathbf{X}'\mathbf{e}\mathbf{e}'\mathbf{X}$ is the meat.

Standard Errors in General

"Sandwich Estimator" for the variance: the $(\mathbf{X}'\mathbf{X})^{-1}$'s are the bread, $\mathbf{X}'\mathbf{e}\mathbf{e}'\mathbf{X}$ is the meat.

$$Var(\underline{\widehat{\beta}}) = E\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e}\mathbf{e}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}|\mathbf{X}\right]$$

So what is $\underset{n\times n}{\mathbf{e}\mathbf{e}'}$? If e_t is i.i.d.,

$$E\left[\mathbf{e}\mathbf{e}'\right] = \sigma^2\mathbf{I}$$

$$\begin{aligned} Var(\underline{\widehat{\beta}}) &= E\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\sigma^2\mathbf{I}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}|\mathbf{X}\right] \\ &= \sigma^2(\mathbf{X}'\mathbf{X})^{-1} \\ &= \sigma_e^2\left[\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'\right]^{-1} \end{aligned}$$

Standard Errors in General

If e_t is NOT i.i.d.,

$$\underset{n\times n}{\mathbf{ee}'} = \underset{n\times n}{\Omega}$$

$$Var(\underline{\widehat{\beta}}) \quad = E\left[(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\boldsymbol{\Omega}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}|\mathbf{X}\right]$$

Now $\mathbf{X}'\boldsymbol{\Omega}\mathbf{X}$ is the meat of the sandwich.

- Heteroskedasticity: diagonal elements are not identical ($\sigma_i \neq \sigma_j$)
- Serial correlation: off-diagonal elements are not identical ($Cov(e_t, e_{t-s}) \neq 0$)
- Can we estimate $\boldsymbol{\Omega}$ to give different weight to different observations?

Heteroskedasticity

What is the problem? Suppose

$$y_t = \alpha + \beta x_t + e_t$$

$$\sigma^2_{e,t} = \sigma^2 \cdot f(x_t), \;\; \text{unknown } f(\cdot)$$

- $\hat{\beta} = \beta + \frac{\sum x_t e_t}{\sum (x_t - \bar{x})^2}$
- Unbiased as long as $Cov(x, e) = 0$.
- BUT $Var(\hat{\beta}) = \frac{E[x_t^2 \sigma^2_{e,t}]}{SST_x^2}$
- Naive $\widehat{Var(\hat{\beta})} = \frac{\sigma^2}{SST_x}$

Serial Correlation

What is the problem? Suppose

$$y_t = \alpha + \beta x_t + e_t$$

$$e_t = \rho e_{t-1} + a_t$$

- $\hat{\beta} = \beta + \frac{\sum x_t e_t}{\sum (x_t - \overline{x})^2}$
- Unbiased as long as $Cov(x, e) = 0$.
- BUT $Var(\hat{\beta}) = \frac{\sigma^2}{SST_x} + f(Cov(e_t, e_{t-j}))$
- Naive $\widehat{Var(\hat{\beta})} = \frac{\sigma^2}{SST_x}$

HC or White's standard errors

- Assuming constant variance:

$$Var(\hat{\beta}) = \sigma_e^2 \left[\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t' \right]^{-1}$$

- Inflating variance because of noisier observations:

$$Var(\hat{\beta})_{HC} = \left[\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t' \right]^{-1} \left[\sum_{t=1}^{T} \hat{e}_t^2 \mathbf{x}_t \mathbf{x}_t' \right] \left[\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t' \right]^{-1}$$

- where $\hat{e}_t$ is the OLS residual

HAC or Newey-West standard errors

- Heteroskedasticity-autocorrelation-consistent (HAC) standard errors:

$$Var(\hat{\beta})_{HC} = \left[\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'\right]^{-1} \left[\sum_{t=1}^{T} \hat{e}_t^2 \mathbf{x}_t \mathbf{x}_t'\right] \left[\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'\right]^{-1}$$

$$Var(\hat{\beta})_{HAC} = \left[\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'\right]^{-1} \hat{C}_{HAC} \left[\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'\right]^{-1}$$

- $\hat{C}_{HAC}$ is a weighting matrix to inflate variance for either noisier observations, or autocorrelated observations.

HAC or Newey-West standard errors

$$\hat{C}_{HAC} = \left[\sum_{t=1}^{T} \hat{e}_t^2 \mathbf{x}_t \mathbf{x}_t'\right] + \sum_{j=1}^{l} w_j \sum_{t=j+1}^{T} \left(\mathbf{x}_t \hat{e}_t \hat{e}_{t-j} \mathbf{x}_{t-j}' + \mathbf{x}_{t-j} \hat{e}_{t-j} \hat{e}_t \mathbf{x}_t'\right)$$

- First term is for heteroskedasticity
- Second term is for autocorrelation
- You get to choose an autocorrelation lag l, default is a fraction of sample size T.
- w_j is a weighting function