Econometrics 1

Willi Mutschler

Winter 2017/2018

Exercise: Show that $\hat{\sigma}^2 = \frac{1}{T-2}\sum_{t=1}^T \hat{u}_t^2$ is an unbiased estimator of the variance of the error terms
Proof:

The simple linear regression model is

$$y_t = \alpha + \beta x_t + u_t \tag{1}$$

and the least squares estimator for the slope is given by

$$\hat{\beta} = \frac{S_{xy}}{S_{xx}} = \beta + \frac{\sum_{t=1}^T (x_t - \bar{x})(u_t - \bar{u})}{\sum_{t=1}^T (x_t - \bar{x})^2} = \beta + \frac{S_{xu}}{S_{xx}} \tag{2}$$

Summing up (1) and dividing by T yields

$$\bar{y} = \alpha + \beta \bar{x} + \bar{u} \tag{3}$$

Subtracting both equations yields:

$$y_t - \bar{y} = \beta(x_t - \bar{x}) + (u_t - \bar{u}) \tag{4}$$

By definition the residual is equal to

$$\hat{u}_t = y_t - \hat{\alpha} - \hat{\beta} x_t$$

Again summing up and dividing by T yields:

$$\bar{\hat{u}}_t = \bar{y} - \hat{\alpha} - \hat{\beta}\bar{x} \tag{5}$$

Subtracting the equations and noting that due to the first normal equation $\bar{\hat{u}}_t = 0$, we get

$$\hat{u}_t = (y_t - \bar{y}) - \hat{\beta}(x_t - \bar{x}) \tag{6}$$

Inserting equation (4) yields:

$$\hat{u}_t = -(\hat{\beta} - \beta)(x_t - \bar{x}) + (u_t - \bar{u}) \tag{7}$$

Squaring and summing up both sides we have

$$\begin{aligned}
\sum_{t=1}^T \hat{u}_t^2 &= (\hat{\beta} - \beta)^2 \sum_{t=1}^T (x_t - \bar{x})^2 + \sum_{t=1}^T (u_t - \bar{u})^2 - 2(\hat{\beta} - \beta)\sum_{t=1}^T (x_t - \bar{x})(u_t - \bar{u}) && (8)\\
&= (\hat{\beta} - \beta)^2 S_{xx} + S_{uu} - 2(\hat{\beta} - \beta) S_{xu} && (9)\\
&= S_{uu} - (\hat{\beta} - \beta) S_{xx} && (10)
\end{aligned}$$

where we have also taking (2) into account. We make use of the fact that

$$E(\bar{u}) = 0 \text{ and } var(\bar{u}) = E(\bar{u}^2) = \frac{\sigma^2}{T} \tag{11}$$

Then the expectation of the first term in (10) may be expressed as

$$E(S_{uu}) = E\left(\sum_{t_1}^{T}(u_t - \bar{u})^2\right) = E\left[\sum_{t=1}^{T} u_t^2 - T\bar{u}^2\right] \tag{12}$$

$$= \sum_{t=1}^{T} E(u_t^2) - TE(\bar{u}^2) = T\sigma^2 - T\frac{\sigma^2}{T} = (T-1)\sigma^2 \tag{13}$$

For the second term in (10), we make use of the distribution of

$$\hat{\beta} \sim N(0, \sigma^2/S_{xx}) \tag{14}$$

Hence,

$$E(\hat{\beta} - \beta)^2 \cdot S_{xx} = var(\hat{\beta}) \cdot S_{xx} = \frac{\sigma^2}{S_{xx}} S_{xx} = \sigma^2 \tag{15}$$

In sum, the expectation of (10) is given by

$$E\left[\sum_{t=1}^{T} \hat{u}_t^2\right] = (T-1)\sigma^2 - \sigma^2 = (T-2)\sigma^2 \tag{16}$$

$$\Leftrightarrow E\left[\underbrace{\frac{1}{T-2}\sum_{t=1}^{T} \hat{u}_t^2}_{\hat{\sigma}^2}\right] = \sigma^2 \tag{17}$$

Therefore $\hat{\sigma}^2$ is an unbiased estimator of the variance of the error terms.

Exercise: Show that $\sum_{t=1}^{T}\left(\frac{\hat{u}_t}{\sigma}\right)^2 \sim \chi^2(T-2)$
Proof (INCOMPLETE):

We start at equation (10)

$$\sum_{t=1}^{T} \hat{u}_t^2 = S_{uu} - (\hat{\beta} - \beta)^2 S_{xx} \tag{18}$$

Dividing by σ^2 we get

$$\sum_{t=1}^{T} \left(\frac{\hat{u}_t}{\sigma}\right)^2 = \frac{S_{uu}}{\sigma^2} - \frac{(\hat{\beta} - \beta)^2}{\sigma^2/S_{xx}} \tag{19}$$

Note that the second term $\left(\frac{\hat{\beta}-\beta}{\sigma/\sqrt{S_{xx}}}\right)^2$ is the square of a single standard normally distributed variable due to (14) and therefore $\chi^2(1)$ distributed. The first term may be expressed as

$$\frac{S_{uu}}{\sigma^2} = \frac{\sum_{t_1}^{T}(u_t - \bar{u})^2}{\sigma^2} = \frac{\sum_{t_1}^{T} u_t^2}{\sigma^2} - \frac{T\bar{u}^2}{\sigma^2} \tag{20}$$

$$= \sum_{t=1}^{T} \left(\frac{u_t}{\sigma}\right)^2 - \left(\frac{\bar{u}}{\sigma/\sqrt{T}}\right)^2 \tag{21}$$

The first term is the sum of T squared standard normally distributed random variables, hence it is $\chi^2(T)$ distributed, whereas the second term is $\chi^2(1)$ distributed as $\bar{u}$ is the mean of iid random variables and its distribution is equal to $\bar{u} \sim N(0, \sigma^2/T)$.
Going back to (18), we find that the right hand side is the sum of $T - 1 - 1 = T - 2$ squared standard normally and independently distributed random variables, hence

$$\sum_{t=1}^{T} \left(\frac{\hat{u}_t}{\sigma}\right)^2 \sim \chi^2(T) - \chi^2(1) - \chi^2(1) = \chi^2(T-2) \tag{22}$$

WRONG: In general the distribution of the difference of two independently distributed χ^2 variables is not χ^2 in general... Need to express this as some sort of sums maybe instead of differences.