L3: CNLRM Distribution Interval, and Testing

Feng Li
`feng.li@cufe.edu.cn`

School of Statistics and Mathematics
Central University of Finance and Economics

Revision: September 23, 2014

What we have learned last time...

1. Estimating $\hat{\beta}_1$ and $\hat{\beta}_2$ and find their variance via OLS.
2. The properties of OLS (have you done the assignments?).
3. The assumptions of linear model. Are they realistic?
4. The measurement of goodness fit.
5. Remember to use the correct notations:
    1. The population regression function: $Y_i = \beta_1 + \beta_2 X_i + u_i$
    2. We estimate it from sample regression function: $Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i$
    3. And the estimated value of Y_i: $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$

Today we are going to learn...

1. Normal assumptions of u_i
2. Confidence intervals for regression coefficients β_1 and β_2
3. Confidence interval for σ^2
4. Hypothesis testing
5. Predictions
6. Normality tests of residuals
7. Consistency of OLS and MLE

Normal assumptions of u_i

1. The classical normal linear regression model assumes each u_i is distributed normally with

$$\begin{aligned} E(u_i) &= 0 \\ Var(u_i) &= E(u_i - E(u_i))^2 = \sigma^2 \\ cov(u_i, u_j) &= E[(u_i - E(u_i))(u_j - E(u_j))] = E(u_i, u_j) = 0, i \neq j. \end{aligned}$$

2. We write $u_i \sim N(0, \sigma^2)$ for short.
3. Why normal?
    1. Simple.
    2. Central limit theorem.

The properties under normal assumptions of u_i

1. The estimators are unbiased, i.e.,

$$E(\hat{\beta}_1) = \beta_1, \qquad E(\hat{\beta}_2) = \beta_2$$

see Appendix 3A.2

2. The variance of the estimators are minimal.
3. The estimators of the parameters also follows normal distribution (**suppose σ^2 is known** which is of course not true).

$$\hat{\beta}_1 \sim N(\beta_1, \frac{\sum X_i^2}{n \sum x_i^2}\sigma^2), \qquad \hat{\beta}_2 \sim N(\beta_2, \frac{1}{\sum x_i^2}\sigma^2).$$

4. Recall that σ^2 is not known and replaced with its estimator $\hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n-2}$, and

$$(n-2)\frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-2).$$

5. $\hat{\beta}_1$ is independent of $\hat{\sigma}^2$, so does $\hat{\beta}_2$.

Summary: The least estimators $\hat{\beta}_1$ and $\hat{\beta}_2$ are **best unbiased estimators(BUE) in the entire class of unbiased estimators**.

Confidence Intervals for regression coefficients

1. $\frac{\hat{\beta}_i - \beta_i}{se(\hat{\beta}_i)} \sim N(0,1)$ for $i = 1, 2$ when σ^2 is known **which is rare**.
2. $\frac{\hat{\beta}_i - \beta_i}{se(\hat{\beta}_i)} \sim t(n-2)$ when σ^2 is replaced by $\hat{\sigma}^2$.
    1. Given α level of significance,

$$\Pr\left[-t_{\alpha/2} \leqslant \frac{\hat{\beta}_i - \beta_i}{se(\hat{\beta}_i)} \leqslant t_{\alpha/2}\right] = 1 - \alpha$$

       which provides $100(1-\alpha)$ percent confidence interval for β_i

$$\Pr\left[\hat{\beta}_i - t_{\alpha/2} se(\hat{\beta}_i) \leqslant \beta_i \leqslant \hat{\beta}_i + t_{\alpha/2} se(\hat{\beta}_i)\right] = 1 - \alpha$$

       or simply $\hat{\beta}_i \pm t_{\alpha/2} se(\hat{\beta}_i)$.
    2. The interpretation:
       Right way: Given the confidence coefficient of $1-\alpha$, $100(1-\alpha)$ out of 100 cases the interval will contain the true β_i
       Wrong way: The probability of β_i falling into the interval is $100(1-\alpha)$.
       Note: The probability of β_i falling into the interval is either 0 or 1.

Confidence Intervals for σ^2

1. Given $\alpha/2$ level of significance,

$$\Pr\left[\chi^2_{1-\alpha/2} \leqslant (n-2)\frac{\hat{\sigma}^2}{\sigma^2} \leqslant \chi^2_{\alpha/2}\right] = 1-\alpha$$

   which provides $100(1-\alpha)$ percent confidence interval for σ^2

$$\Pr\left[(n-2)\frac{\hat{\sigma}^2}{\chi^2_{\alpha/2}} \leqslant \sigma^2 \leqslant (n-2)\frac{\hat{\sigma}^2}{\chi^2_{1-\alpha/2}}\right] = 1-\alpha$$

   Remember that χ^2 is always positive and skewed.
2. **Exercise Table 3.2**: Construct the confidence intervals for β_2 and σ^2.

The significance of coefficients: the t test

1. **Significant of a statistic**: If the value of the test statistic lies in the **critical region**.
2. Significance testing: Find the critical region
3. Procedures:
   1. Write down the null hypothesis (H_0) and alternative hypothesis (H_α)
   2. Calculate the test statistic e.g., $t = (\hat{\beta}_2 - \beta_2^*)/se(\hat{\beta}_2)$
   3. Look up the table and find the critical value
   4. Make decision.
4. **One-side test** vs **two-sided test** Table

Type of hypothesis	H_0: the null hypothesis	H_1: the alternative hypothesis	Decision rule: reject H_0 if
Two-tail	$\beta_2 = \beta_2^*$	$\beta_2 \neq \beta_2^*$	$\|t\| > t_{\alpha/2,\text{df}}$
Right-tail	$\beta_2 \leq \beta_2^*$	$\beta_2 > \beta_2^*$	$t > t_{\alpha,\text{df}}$
Left-tail	$\beta_2 \geq \beta_2^*$	$\beta_2 < \beta_2^*$	$t < -t_{\alpha,\text{df}}$

The significance of σ^2: the χ^2 test

1. The testing purpose: if $\sigma^2 = \sigma_0^2$ or not.
2. The decision rule.

H_0: the null hypothesis	H_1: the alternative hypothesis	Critical region: reject H_0 if
$\sigma^2 = \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$\frac{\text{df}(\hat{\sigma}^2)}{\sigma_0^2} > \chi^2_{\alpha,\text{df}}$
$\sigma^2 = \sigma_0^2$	$\sigma^2 < \sigma_0^2$	$\frac{\text{df}(\hat{\sigma}^2)}{\sigma_0^2} < \chi^2_{(1-\alpha),\text{df}}$
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$\frac{\text{df}(\hat{\sigma}^2)}{\sigma_0^2} > \chi^2_{\alpha/2,\text{df}}$ or $< \chi^2_{(1-\alpha/2),\text{df}}$

The ANOVA table for the two-variable regression model

1. We arrange the sums of squares in the following table (aka ANOVA table.)

ANOVA TABLE FOR THE TWO-VARIABLE REGRESSION MODEL

Source of variation	SS*	df	MSS†
Due to regression (ESS)	$\sum \hat{y}_i^2 = \hat{\beta}_2^2 \sum x_i^2$	1	$\hat{\beta}_2^2 \sum x_i^2$
Due to residuals (RSS)	$\sum \hat{u}_i^2$	$n-2$	$\frac{\sum \hat{u}_i^2}{n-2} = \hat{\sigma}^2$
TSS	$\sum y_i^2$	$n-1$	

*SS means sum of squares.
†Mean sum of squares, which is obtained by dividing SS by their df.

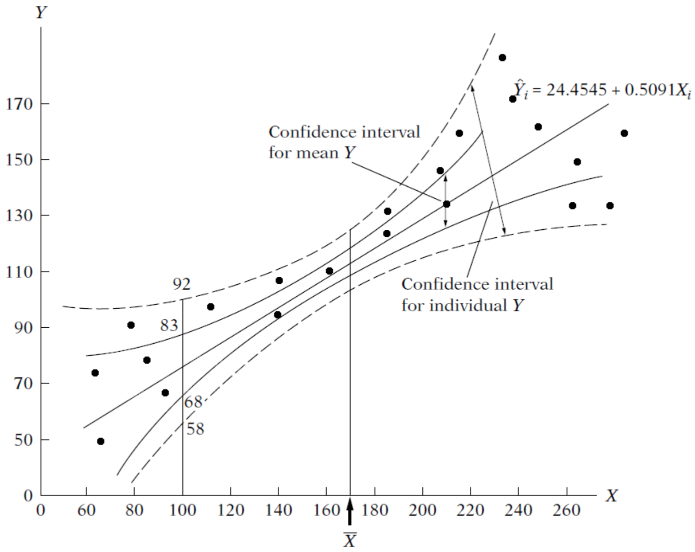
2. Then we consider

$$\frac{\mathrm{ESS}/\mathrm{df}_{\mathrm{ESS}}}{\mathrm{RSS}/\mathrm{df}_{\mathrm{RSS}}} = \frac{\hat{\beta}_2^2 \sum x_i^2}{\sum \hat{u}_i^2/(n-2)} \sim F(1, n-2)$$

which can be used to test the overall significance of the model. In particular the null hypothesis of $\beta_2 = 0$ can also be tested (**How**?).

Predictions

1. **Mean predictions**: Predict Y_0 given X_0 which is from the observations that on the population regression line (see next figure).
   1. $$var(\hat{Y}_0) = \sigma^2 \left[\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum x_i^2} \right]$$
   2. $$t = \frac{\hat{Y}_0 - (\beta_1 + \beta_2 X_0)}{se(\hat{Y}_0)}$$
2. **Individual prediction**: Predict Y_0 given X_0 which is **not** from the population regression line.
   1. $$var(Y_0 - \hat{Y}_0) = \sigma^2 \left[1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum x_i^2} \right]$$
   2. $$t = \frac{Y_0 - \hat{Y}_0}{se(Y_0 - \hat{Y}_0)}$$


Y
X
170
150
130
110
90
70
50
0
0
60
80
100
120
140
160
180
200
220
240
260
$\bar{X}$
$\hat{Y}_i = 24.4545 + 0.5091X_i$
Confidence interval
for mean Y
Confidence interval
for individual Y
92
83
68
58

Normality tests of residuals

1. Histogram of residuals
2. Anderson–Darling test:
   H_0: the variable is normal distributed.
3. Jarque-Bera test:
   H_0 : the variable is normal distributed (skewness(S)=0, kurtosis(K)=3).
   test statistic: $JB = n[S^2/6 + (K-3)^2/24]$

Take home questions

1. Read **Appendix A.8, p.831** if you have problems of hypothesis testing.
2. Do the example in **p. 133**.

Consistency of an estimator

- **Consistency**: An estimator is unbiased and its variance tends to zero as the sample size goes to infinity.
- The OLS and MLE estimators are consistent.