Assignment8

OZONE

March 14, 2018

Testing for the ratio of coefficients

Question

1. The exact variance of the ratio is 0.025103, and the asymptotic and bootstrap one is 0.0347249. The asymptotic and bootstrap one is bigger, so the Wald statistics is relatively smaller. I expect the rejection to be bewlow 5%.
2. The 95% quantile of the Wald test is3.2718247, which is lower than the asymptotic 95% quantile:3.8414588.
3. The average rejection rate is 0.04.
4. The average rejection rate is 0.03. Underreject the null hypothesis.
5. The average rejection rate when comparing original Wald test with 95% quantile of the bootstrapped Wald centered around $\frac{\beta_1}{\beta_2}$ is 0. The average rejection rate when comparing original Wald test with 95% quantile of the bootstrapped Wald centered around $\frac{\hat{\beta_1}}{\hat{\beta_2}}$ is 0.035.
6. From 3, the rejection rate of Wald test is 0.04, which indicates asymptotically, the Wald test statistic converges in distribution to Chi-square distribution with degree of freedom of 1. From 4, rejection rate of 0.03 indicates that we underreject the null hypothesis, which is what we expected in Question 1. For question 5, the first rejection rate is false since we centered around the true value of β_1/β_2; the second rejection rate is 0.035, which is close to 5%.
7. Under H0: $\frac{\hat{\beta_1}}{\hat{\beta_2}}$, $\hat{\theta}-\theta = \frac{\hat{\beta_1}}{\hat{\beta_2}} - \frac{0.3}{1.1}$ and $R' = (0\ \frac{1}{\hat{\beta_2}}\ \ -\frac{\hat{\beta_1}}{\hat{\beta_2}^2})$. Under the H0: $\hat{\beta_1} = \frac{0.3}{1.1}\hat{\beta_2}$, $\hat{\theta}-\theta = \hat{\beta_1} - \frac{0.3}{1.1}\hat{\beta_2} - 0$ and $R' = (0\ 1\ -\frac{0.3}{1.1})$. Denote the Wald test statistic under first hypothesis as W_1 and the corresponding Wald test for second hypothesis as W_2. After calculation and rearrangement:

$$\frac{W_1}{W_2} = \frac{1+(\frac{0.3}{1.1})^2}{1+(\frac{\hat{\beta_1}}{\hat{\beta_2}})^2}$$

. If $\frac{\hat{\beta_1}}{\hat{\beta_2}} > \frac{0.3}{1.1}$, then $W_1 < W_2$. If $\frac{\hat{\beta_1}}{\hat{\beta_2}} < \frac{0.3}{1.1}$, then $W_1 > W_2$

Cross-validation

model	mse	MSE
reg1	mse_in	1.198144
reg1	mse_out	1.249883
reg2	mse_in	1.163637
reg2	mse_out	1.241117

Question

1.

- crossval_df is a 100×3 tidy dataframe. For cell, the object is in list type for train and test columns and character for id column. The dimension for the cells in train is 99×3 and for the cells in test is 1×3, for the celss in .id is 1×1.

- mse contains mse_in and mse_out, which indicates in-sample and out-of-sample MSE. model contains reg1 and reg2, which indicates wether we include x_2 in the regression. value contains the value of MSE. -.x refers to the reg column, .y refers to the test column.

2.

- For in-sample MSE, adding one more variable will decrease the MSE. Intuitively, adding one more variable will better fit the data, which will decrease the MSE.
- For out-sample MSE, we still observes that the MSE decreases when adding one more variable. However, this is because the randomness of the data generating process. If we delete seed(2018), most of the time, we observe the out-sample MSE is smaller when we have the correct regression model (without x_2).
- For each model, the out-sample MSE is bigger than the in-sample MSE. Intuitively, OLS is trying to minimize the the in-sample MLE. Therefore, the out-of-sample MLE tends to be bigger than the in-sample MLE.

3. Using Hansen's formula, we calculate the out-of-sample MSE for regression one and two (with x_2 in the regression), the values are 1.2498829 and 1.2411171. They are the same as by simulation.