Spring 2023

Quantitative Analysis
Chung-Ming Kuan

Problem Set 3

Due: 03/20

Part One: Hand-Written Exercise

1. We mentioned that the F statistic is given by:

$$F = \frac{(\mathrm{SSR}_r - \mathrm{SSR}_{ur})/q}{\mathrm{SSR}_{ur}/(n-k-1)},$$

where SSR_r and SSR_{ur} are the residual sums of squares of restricted and unrestricted regressions respectively. $(\mathrm{SSR}_r - \mathrm{SSR}_{ur})$ and SSR_{ur} are independent of each other.

(a) Given the fact that:

$$\frac{(n-k-1+q)\hat{\sigma}_r^2}{\sigma^2} - \frac{(n-k-1)\hat{\sigma}_{ur}^2}{\sigma^2} \sim \chi^2(q),$$

where $\hat{\sigma}_r^2$ and $\hat{\sigma}_{ur}^2$ are the OLS estimators of σ^2 of the restricted and unrestricted regressions respectively. Please show that

$$\frac{(\mathrm{SSR}_r - \mathrm{SSR}_{ur})/q}{\mathrm{SSR}_{ur}/(n-k-1)} \sim F(q, n-k-1).$$

(b) Show that the F statistic can also be written as the R-squared form

$$F = \frac{(R_{ur}^2 - R_r^2)/q}{(1-R_{ur}^2)/(n-k-1)},$$

where R_r^2 and R_{ur}^2 are the R^2s of the restricted and unrestricted regressions.

2. Abby and Bob are trying to understand the difference of the health expenditure, y, of a smoker and a non-smoker with different models. Abby adopts Model A while Bob adopts Model B:

$$\text{Model A: } E[y] = \beta_0 + \beta_1 x_1, \text{ where } x_1 = \begin{cases} 1, & \text{for smokers,} \\ 0, & \text{for non-smokers} \end{cases}$$
$$\text{Model B: } E[y] = \alpha_0 + \alpha_1 x_2, \text{ where } x_2 = \begin{cases} 0, & \text{for smokers,} \\ 1, & \text{for non-smokers} \end{cases}$$

(a) Please express β_0 and β_1 with α_0 and α_1.

(b) Are predictions, $\hat{y}$, the same for Model A and Model B? Discuss both $\hat{y}$ for a smoker and a non-smoker.

(c) Chris combines Model A and Model B and get Model C:

$$\text{Model C:}\ \ E[y] = \delta_0 + \delta_1 x_1 + \delta_2 x_2,$$

$$\text{where } x_1 = \begin{cases} 1, & \text{for smokers,} \\ 0, & \text{for non-smokers} \end{cases}, \ x_2 = \begin{cases} 0, & \text{for smokers,} \\ 1, & \text{for non-smokers} \end{cases}$$

Chris claims that Model C has more explaining power than Model A and Model B since it includes more explanatory variables. Is his statement true? Explain it.

3. The following model can be used to study whether campaign expenditures affect election outcomes:

$$\text{voteA} = \beta_0 + \beta_1 ln(\text{expendA}) + \beta_2 ln(\text{expendB}) + \beta_3 \text{prtystrA} + u,$$

where "voteA" is the percentage of the vote received by candidate A, "expendA" and "expendB" are campaign expenditures by candidates A and B, and "prtystrA" is a measure of party strength for candidate A (the percentage of the most recent presidential vote that went to A's party).

(a) What is the interpretation of β_1?

(b) In terms of the parameters, state the null hypothesis that the effect of the increase in A's expenditure will be offset by the increase in B's expenditure.

(c) Write the detailed procedure to do the hypothesis testing in (b).

(d) If someone claims that both candidates' expenditures do not have any effect on the outcome, how can you specify a testing null hypothesis?

(e) Write the detailed procedure to do the hypothesis testing in (d).

Part Two: Computer Exercise

Following Question 2 of the computer exercise in Problem Set 2, consider the following model:

$$\texttt{drat}_i = \beta_0 + \beta_1 \texttt{wt}_i + \beta_2 \texttt{hp}_i + \beta_3 \texttt{qsec}_i + \beta_4 \texttt{vs}_i + u_i,$$

1. Test the hypothesis $H_0 : \beta_1 = 0$.

(a) Please construct the t statistic without the function `lm()`.

(b) Use the function `lm()` to directly obtain the t statistic. Verify that it's identical to (a).

2. Test the hypothesis $H_0 : \beta_1 = \beta_2 = 0$.

   (a) Please construct the constrained and unconstrained model, obtain R^2_{ur} and R^2_r and construct the F statistic.

   (b) Instead of R^2_{ur} and R^2_r, please obtain SSR_{ur} and SSR_r and recalculate the F statistic. Verify that it's identical to (a).

   (c) Use the function `linearHypothesis()` to directly obtain the F statistic. Verify that it's identical to (a).