Assignment 1
Assigned on Apr 9th, 2022. Due date Saturday, May 7th, 2022 before midnight.

Instructions

Siraphassorn Noochoochai
6304641415

(1) Please read the instruction carefully.
(2) Please read each question carefully and answer the questions straightforwardly. Always provide economic reasons at lease a paragraph for your analysis, or a graph when necessary, even when the question does not indicate so.
(3) Handing and submitting assignments are only available via BE Moodle.

Answering the questions and preparing answer sheets

(1) Answers are to be handwritten, in either digital or analog form, in a blank canvas or any clean paper. Make sure that your handwriting is clearly visible and readable.
(2) There is no need to rewrite the question. Just indicate the question number clearly for each of the answer, such as 1.a).
(3) Default decimal point is 4.
(4) Choose precise wordings, especially when you want to interpret the meaning of a test, confidence interval, or coefficients.
(5) When done, for the digital case, collage all the pages into a single PDF file. For those who write on sheets of paper, take photo of all pages then convert all of them into a single PDF file as well.
(6) Name your PDF file as StudentID_YourNickname, such as 640123456_Bo.

Submitting your answers

(1) Make sure your file does not exceed 10MB. This is the maximum file size for BE Moodle upload.
(2) Login to BE Moodle, head into the course, then the assignment topic.
(3) Choose your file to submit. Done. There will be timestamp for your upload date and time, so please make sure to not submit later than that.

(1)

1.a) Based on the regression results provided, write out the estimated coefficients in the form of regression equation (1.1). Interpret the estimated coefficients associated with *avgsen*. Based on Model (1.1), test whether the average sentence served from prior convictions has an impact on the number of arrests in the current year (1986). Show your work. (Use α = 0.05)

$\hat{\beta}_3 = -0.007$; If average sentence served from prior convictions ($avgsen_i$) increases by 1 month, the number of arrests in the current year ($narr86_i$) will decrease by 0.007 units, holding other factor remain the same.

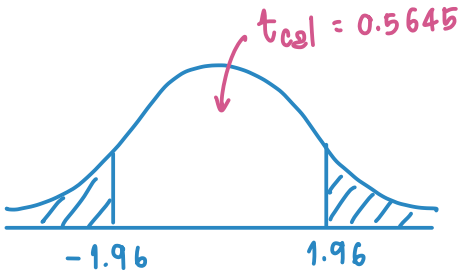
$H_0 : \beta_3 = 0$

$H_1 : \beta_3 \neq 0$

$$t_{cal} = \frac{\hat{\beta}_3 - \beta_3}{se(\hat{\beta}_3)} = \frac{-0.0070 - 0}{0.0124} = 0.5645$$

upper bound : $t_{\frac{\alpha}{2}, n-k} = t_{\frac{0.05}{2}, 2725-6} = 1.96$

lower bound : $t_{\frac{\alpha}{2}, n-k} = -1.96$

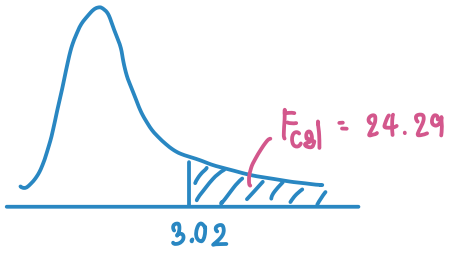


∴ We cannot reject H_0, average sentence served from prior convictions ($avgsen_i$) cannot impact on the number of arrests in the current year ($narr86_i$) at 0.05 level of significence.

1.b) What is the overall significance of the regression from Model (1.1) and Model (1.2)? What test do you use? (Use α = 0.01)

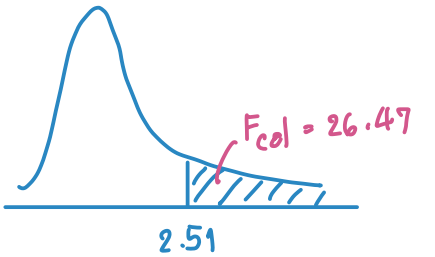
Use F-test

Model (1.1) $F_{\alpha(k-1, n-k)} = F_{0.01(5, 2716)} = 3.02$, $F_{cal} = 24.29$



∴ Reject H_0, we can make sure that slope coefficients are not simultaneously equal to zero, 99 times out of 100.

Model (1.2) $F_{\alpha(k-1, n-k)} = F_{0.01(8, 2716)} = 2.51$, $F_{cal} = 26.47$



∴ Reject H_0, we can make sure that slope coefficients are not simultaneously equal to zero, 99 times out of 100.

1.c) If we are interested in testing whether "ethnic background and legal income" has an impact on the number of arrests in the current year (1986), what kind of null/alternative hypothesis would we be testing? Perform the test and discuss your finding. (Use $\alpha = 0.05$)

$H_0 : \beta_7 = \beta_8 = 0$

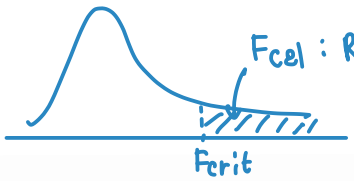
$H_1 :$ otherwise

Unrestriction Model $narr86_i = \beta_1 + \beta_2 pcnv_i + \beta_3 avgsen_i + \beta_4 tottime_i + \beta_5 ptime86_i + \beta_6 qemp86_i + \beta_7 inc86_i + \beta_8 black_i + \beta_9 hispan_i + u_i$

Restriction Model $narr86_i = \beta_1 + \beta_2 pcnv_i + \beta_3 avgsen_i + \beta_4 tottime_i + \beta_5 ptime86_i + \beta_6 qemp86_i + \beta_9 hispan_i + u_i$

Find $F_{cal} = \dfrac{(RSS_R - RSS_{UR})/m}{(RSS_{UR})/(n-k_{UR})}$ or $F_{cal} = \dfrac{(R^2_{UR} - R^2_R)/m}{(1-R^2_{UR})(n-k_{UR})}$

$F_{crit} = F(m, n-k_{UR})$



∴ If we can reject H_0, the restriction model is invalid.
If we cannot reject H_0, the restriction model is valid.

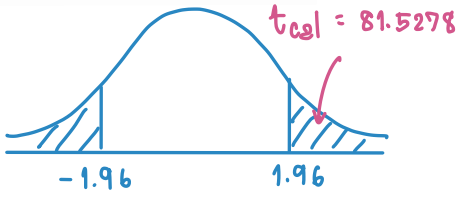
(2)

2.a) Test all the parameters individually if each of them is significantly different from zero of not.

Test β_2 $H_0 : \beta_2 = 0$

$H_1 : \beta_2 \neq 0$

$t_{cal} = \dfrac{\hat{\beta}_2 - \beta_2}{se(\hat{\beta}_2)} = \dfrac{0.587 - 0}{0.0072} = 81.5278$



upper bound : $t_{\frac{\alpha}{2}, n-k} = t_{\frac{0.05}{2}, 97878-4} = 1.96$

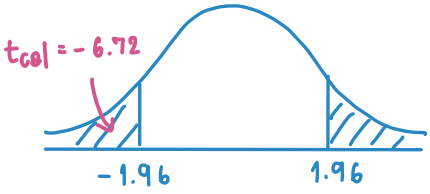
lower bound : $t_{\frac{\alpha}{2}, n-k} = -1.96$

∴ we can reject H_0, slope parameter is not zero at 0.05 level of significance.

Test β_3 $H_0 : \beta_3 = 0$

$H_1 : \beta_3 \neq 0$

$t_{cal} = \dfrac{\hat{\beta}_3 - \beta_3}{se(\hat{\beta}_3)} = \dfrac{-0.0336 - 0}{0.005} = -6.72$



upper bound : $t_{\frac{\alpha}{2}, n-k} = t_{\frac{0.05}{2}, 97878-4} = 1.96$

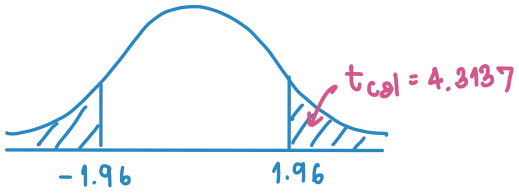
lower bound : $t_{\frac{\alpha}{2}, n-k} = -1.96$

∴ we can reject H_0, slope parameter is not zero at 0.05 level of significance.

Test β_4 $\quad H_0 : \beta_4 = 0$

$H_1 : \beta_4 \neq 0$

$$t_{cal} = \frac{\hat{\beta}_4 - \beta_4}{se(\hat{\beta}_4)} = \frac{0.044 - 0}{0.0102} = 4.3137$$



upper bound : $t_{\frac{\alpha}{2}, n-k} = t_{\frac{0.05}{2}, 97878-4} = 1.96$

lower bound : $t_{\frac{\alpha}{2}, n-k} = -1.96$

∴ we can reject H_0, slope parameter is not zero at 0.05 level of significance.

2.b) How much on average does a civil servant and state employee earns more or less than the others disregarding the year?

$$\ln \widehat{wage}_i = 9.1748 + 0.587\, civil_i$$

civil servant state employee → civil = 1 : $\ln \widehat{wage}_1 = 9.1748 + 0.587(1)$

$= 9.7618$ —①

otherwise → $civil_i = 0$: $\ln \widehat{wage}_2 = 9.1748 + 0.587(0)$

$= 9.1748$ —②

① − ② ; $\ln \widehat{wage}_1 - \ln \widehat{wage}_2 = 9.7618 - 9.1748$

$$\ln\left[\frac{\widehat{wage}_1}{\widehat{wage}_2}\right] = 0.587$$

$$\frac{\widehat{wage}_1}{\widehat{wage}_2} = e^{0.587}$$

$$\frac{\widehat{wage}_1}{\widehat{wage}_2} - 1 = 1.7985 - 1$$

$$\frac{\widehat{wage}_1 - \widehat{wage}_2}{\widehat{wage}_2} = 0.7985$$

∴ Civil servant and state employee earns more than the others 79.85%.

2.c) How much on average does the pandemic affect wage overall?

$$\ln \widehat{wage} = 9.1748 - 0.0336\, year_i$$

year 2020 → $year_i = 1$: $\ln \widehat{wage}_1 = 9.1748 - 0.0336(1)$

$= 9.1412$ —①

otherwise → $year_i = 0$: $\ln \widehat{wage}_2 = 9.1748 - 0.0336(0)$

$= 9.1748$ —②

② − ① ; $\ln \widehat{wage}_2 - \ln \widehat{wage}_1 = 9.1748 - 9.1412$

$$\ln\left[\frac{\widehat{wage}_2}{\widehat{wage}_1}\right] = 0.0336$$

$$\frac{\widehat{wage}_2}{\widehat{wage}_1} - 1 = 1.034 - 1$$

$$\frac{\widehat{wage}_2 - \widehat{wage}_1}{\widehat{wage}_1} = 0.034$$

∴ If the pandemic happen, the wage will reduce 3.4%.

2.d) Are the control group and the treatment group better-off or worse-off during the pandemic. Discuss each group separately, show your work and explain with economic reasons according to the intention of this model.

$$\widehat{\ln wage_i} = 9.1748 + 0.587\, civil_i - 0.0336\, year_i + 0.0444\, civil_i \cdot year$$

control group → $year_i = 1$ and $civil_i = 1$

$$\widehat{\ln wage_1} = 9.1748 + 0.587(1) - 0.0336(1) + 0.0444(1)(1)$$
$$= 9.7726 \quad -①$$

treatment group → $year_i = 1$ and $civil_i = 0$

$$\widehat{\ln wage_2} = 9.1748 + 0.587(0) - 0.0336(1) + 0.0444(0)(1)$$
$$= 9.1412 \quad -②$$

① − ② ; $\widehat{\ln wage_1} - \widehat{\ln wage_2} = 9.7726 - 9.1412$

$$\ln\left[\frac{\widehat{wage_1}}{\widehat{wage_2}}\right] = 0.6314$$

$$\frac{\widehat{wage_1}}{\widehat{wage_2}} = e^{0.6314}$$

$$\frac{\widehat{wage_1}}{\widehat{wage_2}} - 1 = 1.88 - 1$$

$$\frac{\widehat{wage_1} - \widehat{wage_2}}{\widehat{wage_2}} = 0.88$$

∴ In covid situation, civil servant and state employee receive wage more than private firm or freelance 88%.
control group is better than treatment group.

(3)

3.a) Is there evidence of multicollinearity in the data? How do you know? Explain your answers in detail and state the critical value for hypothesis testing to receive full points.

There is no correlation of multicollinearity in the data

1. Pair-wise correlation ⇒ $|r| < 0.8$,
If coefficient of correlation exceed 0.8 it will cause multicollinearity

	acetic	h2s	lactic
acetic	1.0000		
h2s	0.2700	1.0000	
lactic	0.3607	0.6448	1.0000

∴ From the table, it doesn't have coefficient of correlation that exceed 0.8 so the multicollinearity problem is not happen.

2. VIF and TOL ⇒ VIF shoud not exceed 10 and TOL should be closer to 1 rather than 0

Variable	VIF	1/VIF
lactic	1.83	0.546648
h2s	1.72	0.582609
acetic	1.15	0.867477
Mean VIF	1.57	

∴ VIF from the table aren't exceed 10 so the multicollinearity problem is not happen.

3.b) What is the property of BLUE? If there is the multicollinearity problem, is the OLS estimators still retain the property of BLUE? If not, which properties are violated?

property of BLUE (Best Linear Unbias estimators)
- It is linear.
- It is unbiased, expected value is equal to the true value $E(\hat{\beta}_i) = \beta_i$.
- It has minimum variance.

If there is the multicollinearity problem, OLS estimators is still retain the property of BLUE, the number of each coefficient are still holding characteristics of the BLUE.

④

4.a) Interpret the intercept and slope coefficients.

$\widehat{Inf}_t = 1.0108 + 0.5054\,unem_t$

$\hat{\beta}_1 = 1.0108$, $\hat{\beta}_1$ represents the intercept, if the unemployment rates is 0%, the inflation rates will be 1.0108%.

$\hat{\beta}_2 = 0.5054$, if unemployment rate change by 1%, the inflation rates will change 0.5054% in the same direction.

4.b) According to the test statistics given after Table 4.1 below, is there any sufficient evidence to conclude that there is heteroscedasticity problem? Show your work on the hypothesis testing. (Use $\alpha = 0.05$)

1. Using white test

White's general test statistic: 1.0266 Chi-sq (2)

critical value of $\chi^2_{k-1} = \chi^2_{2-1} = 3.8415$

Chi cal < Chi crit : $1.0266 < 3.8415$

∴ We cannot reject H_0 at 0.05 level of significance. Therefore, there is no heteroscedasticity problem in our model.

2. Using BP test

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

chi2(1) = 1.12

critical value of $\chi^2_{k-1} = \chi^2_{2-1} = 3.8415$

Chi cal < Chicrit : $1.12 < 3.8415$

∴ We cannot reject H_0 at 0.05 level of significance. Therefore, there is no heteroscedasticity problem in our model.

4.c) Given your test results in a), do the OLS estimators still retain the property of BLUE? If not, which properties are violated?

There is no heteroscedasticity problem, so the OLS estimators still retain the property of BLUE.