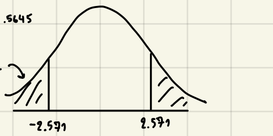
1.a) Based on the regression results provided, write out the estimated coefficients in the form of regression equation (1.1). Interpret the estimated coefficients associated with *avgsen*. Based on Model (1.1), test whether the average sentence served from prior convictions has an impact on the number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

For β_3 ; $H_0 : \beta_3 = 0$, $t_{cal}(\beta_3) = \frac{\hat{\beta}_3 - \beta_3}{se_{\hat{\beta}_3}} = \frac{-.0070 - 0}{.0124} = -.5645$

$H_a : \beta_3 \neq 0$

For $\alpha = .05$; $t_{lower} = -2.571$ $t_{upper} = 2.571$



∴ β_3 reject H_0, so it can be said that the average sentence served from prior convictions has an impact on the number of arrests in the current year.

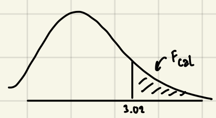
1.b) What is the overall significance of the regression from Model (1.1) and Model (1.2)? What test do you use? (Use α = 0.01)

We can use F-test to find overall significance of the regression from both models.

1) Model (1.1); $H_0 : \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6$, $F_{cal} = \frac{ESS/(k-1)}{RSS/(n-k)} = \frac{85.9532/(6-1)}{1,924.3939/(2,725-6)} = 24.2889$

H_a : otherwise

$\alpha = .01$, $F_{upper,\alpha}(5,2719) = 3.02$

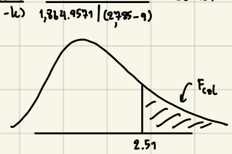


∴ It rejects H_0, so we can make sure that $\beta_2, \beta_3, \beta_4, \beta_5$ & β_6 aren't simultaneously equal to zero 99 times out of 100.

2) Model (1.2); $H_0 : \beta_2 = \beta_3 = \ldots = \beta_9$, $F_{cal} = \frac{ESS/(k-1)}{RSS/(n-k)} = \frac{145.3901/(9-1)}{1,864.9571/(2,725-9)} = 26.4671$

H_a : otherwise

$\alpha = .01$, $F_{upper,\alpha}(8,2716) = 2.51$



∴ It rejects H_0, so we can make sure that $\beta_2, \beta_3, \ldots, \beta_8$ & β_9 aren't simulteneously 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 α = 0.05)

We can use General F-test to test that each pair has an impact or not.

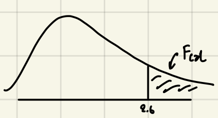
- Unconstrained model : Model 1.2 $R^2_{UR} = .0723$

Constrained model : Model 1.1 $R^2_R = .0428$

- $H_0 : \beta_7 = \beta_8 = \beta_9 = 0$, $F_{cal} = \frac{(R^2_{UR} - R^2_R)/m}{(1 - R^2_{UR})/(n - k_{UR})} = \frac{(.0723 - .0428)/3}{(1 - .0723)/(2,725 - 9)} = 28.7888$

H_a : otherwise

- $\alpha = .05$, $F_{upper,\alpha}(3, 2,716) = 2.6$



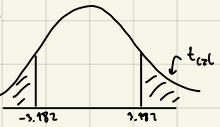
∴ It rejects H_0, so black ethnic background - number of arrests, Hispanic ethnic background - number of arrests or legal income - number of arrests are related simoltaneously to have an impact on the model.

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

For β_2 , $H_0 : \beta_2 = 0$, $t_{cal} = \frac{\hat{\beta}_2 - \beta_2}{se_{\hat{\beta}_2}} = \frac{.587 - 0}{.0072} = 81.5278$

$H_a : \beta_2 \neq 0$

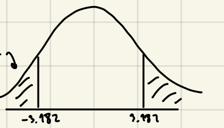
$\alpha = .05$, $t_{upper} = 3.182$, $t_{lower} = -3.182$



For β_3 , $H_0 : \beta_2 = 0$, $t_{cal} = \frac{\hat{\beta}_2 - \beta_2}{se_{\hat{\beta}_2}} = \frac{-.0336 - 0}{.005} = -6.72$

$H_a : \beta_2 \neq 0$

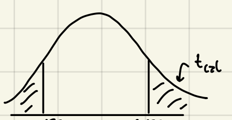
$\alpha = .05$, $t_{upper} = 3.182$, $t_{lower} = -3.182$



For β_4 , $H_0 : \beta_2 = 0$, $t_{cal} = \frac{\hat{\beta}_2 - \beta_2}{se_{\hat{\beta}_2}} = \frac{.0444 - 0}{.0102} = 4.3529$

$H_a : \beta_2 \neq 0$

$\alpha = .05$, $t_{upper} = 3.182$, $t_{lower} = -3.182$



∴ Each parameter rejects H_0, so they're significant different from zero.

1

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

Model: $\ln wage_i = 9.1748 + .587\, civil_i - .0336\, year_i + .0444\, civil_i \cdot year_i$

$*_1$ civil & pan: $\ln wage_i = 9.1748 + .587 - .0336 + .0444 = 9.7726$

$wage_i = e^{9.7726}$

$*_2$ otherwise & pan: $\ln wage_i = 9.1748 - .0336 = 9.1412$

$wage_i = e^{9.1412}$

∴ The civil servant and state employee earns more than the other disregarding which is

17,546.3284 - 9,331.9568 = 8,214.3717 THB

1

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

Model: $\ln wage_i = 9.1748 + .587\, civil_i - .0336\,(1) + .0444\, civil_i\,(1)$

$= 9.1412 + .6314\, civil_i$

For an overall during the pandemic, it affects wage for $e^{9.1412} = 9,331.9568$ THB on an average in the base case.

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.

The control group and the treatment group are better-off during the pandemic because both groups are protected by the labour law, and from the model $\ln wage_i = 9.1748 + .587\, civil_i - .0336\, year_i + .0444\, civil_i \cdot year_i$ all the dummy variables are pluged with 1 which they still have all coefficiances whereas the other group is the private and individual sectors who have tough life because of the pandemic. They gain less wage that can show on the model too. For the otherwise, it's pluged by 0 that losses 2 components [$.587\, civil_i$ & $.0444\, civil_i \cdot year_i$]. That is why the second group receives less wage than the first group.

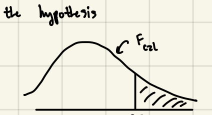
3.a) Is there evidence of multicollinearity in the data? How do you know? Explain your answers in detail and state the critical value for hypothesis testing to receive full points.

No, there's no evidence of multicollinearity. We can use the definition of variance-inflating factor or VIF and tolerance or TOL to find it. Firstly applying VIF, not exceed 10. Secondly, TOL should be closed to 1. Looking at the VIF and 1/VIF table, all values pass the definition.

To state the critical value, we apply the F-test to test the hypothesis

H_0 : no multicollinearity $F_{cal} = \frac{.0633/(4-1)}{(1-.0633)/(59-4+1)} = 1.2614$

H_a : multicollinearity $F_{cri}(\alpha = .05, 3, 59) = 2.8$ => critical value



∴ F_{cal} falls in acceptance region. It cannot reject H_0. So, we can still say that this model has no multicollinearity.

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?

The property of BLUE is linear unbiased which has the least variance among all linear and unbiased estimators. No, if there's the multicollinearity problem, OLS estimators cannot retain the property of BLUE due to high variance of the variables.

4.a) Interpret the intercept and slope coefficients.

Model : $\widehat{Inf}_t = 1.0108 + .5055\, unem_t$

=> interupt 1.0108 means if there's no unemployment rate the inflation rate equals to 1.0108

=> slope coefficient .5055 means if unemployment rate is increasing 1% the inflation rate will increase .5055 % simoultenceously

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 $\alpha = 0.05$)

There's enough evidence to conclude that it's not heteroscedasticity from the white test.

H_0: homoscedasticity $\quad LM_{cal} = 1.0266$

H_2: heteroscedasticity $\quad \chi^2_{.05, k-1} = \chi^2_{.05, 1} = 3.84$

From this test, if $LM_{cal} > \chi^2_{k-1}$, it presents heteroscedasticity. However, from our test, $LM_{cal} < \chi^2_{k-1}$ which cannot reject H_0 and can be said that this model should be homoscedasticity.

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

According to the result we got, the OLS estimators still retain the property of BLUE because it showed that it was homoscedasticity in which variance still keep it position in the property.