Solution for the old exam

Model: \( y(\text{contract}) = \beta_0 + \beta_1 \times \text{city} + \beta_2 \times \text{brink} + \epsilon, n = 46 \) and \( p = 3 \)

1. \( H_0 : \beta_1 = \beta_2 = 0 \) v.s. \( H_1 : \) at least one is not zero.

2. \( p - 1 = 2 \)

3. (a) R.S.E = \( \sqrt{\frac{\text{RSS}}{n-p}} \)
   
   (b) \( F = \frac{(\text{TSS} - \text{RSS})/(p-1)}{\text{RSS}/(n-p)} \)

   (c) \( R^2 = 1 - \frac{\text{RSS}}{\text{TSS}} \)

   \( \text{RSS} = 159.3419 \times 10^6, \text{TSS} = 223.8494 \times 10^6 \)

   \( R^2 = 0.2882 \)

4. Total collection is \( 109324.72 \) less when brink collects it. So the amount the employees stole is \( 109324.72 \).

5. This data is collected over 4 years so we may need to check the auto-correlation of the errors.

6. Since log function is almost linear when range of data does not change in magnitudes.

7. From the output, \( t\)-value = \(-1.015 \) with \( p\)-value = 0.31587 which implies we cannot reject the null hypothesis that the regression parameter corresponding to the Examination variable in this model is zero. We can say the effect of the Examination on Fertility in this model is not significant when other predictors (Agriculture, Education, Catholic and Infant Mortality) are included in the model.

8. R.S.E = \( \sqrt{\frac{\text{RSS}}{(n-p)}} \). So RSS = 2104.8262.


10. We can perform lack of fit test in two cases. One, when we know \( \sigma^2 \) or two, when we have repeated \( y \) for one or more fixed \( x \). Neither of this information is given in this data.

11. \( \text{TSS} = \frac{\text{RSS}}{(1 - R^2)} = 71.78.8070 \)

    \( \text{RSS} = 2104.8286 \)

    \( F = \frac{(\text{TSS} - \text{RSS})/(p-1)}{\text{RSS}/(n-p)} = 19.7673 \)

12. No. The magnitude itself of the coefficient is not important.

13. No. This is an ecological regression, the observation is aggregated, so the conclusion can not be extended to individuals.

14. No. Normal assumption is made on error in the model, not on the data value (response).