Tutorial 3

Inference in Simple Linear Model

January 26, 2022

Weekly Overview

- Asn 3 is due on Friday.
- We will go over, hypothesis testing and confidence intervals.
- Correlation coefficients and coefficient of determination.
- Prediction intervals and confidence intervals for new values.

Review of Confidence Intervals

i) For confidence intervals of a parameter we need three things: the statistic for the parameter, the standard error of the statistic and the α-level critical value at which to construct the confidence interval.

ii) The basic formula for a CI: $\hat{\beta} \pm t^{df}_{\alpha/2} \cdot \sigma_{\hat{\beta}}$

iii) Recall from MATH 203, that when we said find the 99%-CI for μ when $\bar{x} = 1.2$ and standard error $\sigma_{\bar{x}} = 0.3$, with $n = 24$.

iv) The 99% interval for μ will be,
$1.2 \pm 2.807 \times 0.3 = [0.358, 2.042]$.

v) Think about what the confidence interval tells us.

Hypothesis Test:

i) Suppose we have a hypothesis, $\mathcal{H}_0 : \beta = c$ vs $\mathcal{H}_1 : \beta \neq c$

ii) We need to calculate the appropriate statistic,

$$T = \frac{\hat{\beta} - c}{\sigma_{\hat{\beta}}}$$

iii) And then based on the distribution of T we can decide the results of the hypothesis based either on the p-value or the rejection region.

iv) Exercise: using the previous example, do the hypothesis when $c = 0$ at level $\alpha = 0.05$. (**Ans**: We reject the $\mathcal{H}_0$).

v) Exercise: What do we do when $\mathcal{H}_1 : \beta \geq \ (or \ \leq) \ c$

Correlation Coefficient and Coefficient of Determination

i) Measures the strength of the linear relationship and the direction of the relationship (sign)

ii)

$$r = \frac{SS_{xy}}{\sqrt{SS_{xx}SS_{yy}}}$$

iii) r takes the value between $[-1, 1]$, with the end points showing perfect correlation and $r = 0$ showing no correlation.

iv) $r > 0$ shows positive correlation and $r < 0$ shows negative correlation.

v) At the same time $r^2 = \frac{SSReg}{SST} = 1 - \frac{SSE}{SST}$, is the coefficient of determination. This describes what proportion of the variation in Y is explained by our model.

Hypothesis test for Correlation Coefficient

1. Testing $\mathcal{H}_0 : \rho = 0$ vs $\mathcal{H}_1 : \rho \neq 0$ is equivalent to testing whether $\beta_1 = 0$.
2. Reference to seeing the equivalence of this test can by found in the Chapter 13 of the book or in class notes, Set-6.

Prediction errors:

i) After we have estimated the model we can predict two things. The prediction interval of the mean-response will be at some specific value x_p

ii) The $100(1-\alpha)\%$ confidence interval of the mean-response, $\hat{y}$, will be at some specific value x_p. This is given by, $\hat{y} \pm t_{\alpha/2}^{n-2} \times \hat{\sigma} \times \sqrt{\frac{1}{n} + \frac{(x_p - \bar{x})^2}{SS_{xx}}}$.

iii) The $100(1-\alpha)\%$ prediction interval of y at some new point x_p. This is given by, $\hat{y} \pm t_{\alpha/2}^{n-2} \times \hat{\sigma} \times \sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{SS_{xx}}}$.

iv) You can notice that for both the prediction and confidence term, the error is minimized when x_p is as close to $\bar{x}$.

Practice Problem: Swiss fertility data

Using the data set provided, investigate how education affected the fertility rates in the 47 French speaking provinces in Switzerland around the time period 1888. The dependent variable is fertility and the independent variable is education. The units for education is, % education beyond primary school for draftees.

i) Regress fertility on education.

ii) Test the hypothesis $\mathcal{H}_0 : \hat{\beta}_1 = 0$ vs $\mathcal{H}_1 : \hat{\beta}_1 \neq 0$ at the level $\alpha = 0.05$. State any assumptions made. What are your conclusions.

iii) What does failure to reject this hypothesis imply?

iv) Find the 99% CI for $\hat{\beta}_1$. Do this by hand and use the `confint()` function.

v) Find r and r^2. What can you say about these results?

vi) Plot the line of best fit and the predicted values.

vii) Find the 95% CI and 95% PI for $Education = 15\%$. Do this by hand and use the `predict()` function.