Assignment ON3

Elmer V Villanueva

Due at 3:55 PM on Monday 16 March 2020

Instructions

This assignment consists of many questions. You must attempt all questions. All answers must be submitted on ICE. The preferred submission format is an RMD file, but other typical file types are allowed.

Note that you *must* adhere to reporting requirements.

Data

We will considere the same data as in Assignment ON2. Fifteen specimens of human sera were tested comparatively for tuberculin antibody by a new method. The readings were compared to estimated derived from the old method. The results are shown in the table below.

Table. Tuberculin antibody concentrations estimated from new and old laboratory methods.

OLD	NEW
2041.7	12302.7
257.0	6918.3
524.8	4466.8
257.0	1584.9
128.8	933.3
128.8	1659.6
1023.3	9120.1
134.9	575.4
257.0	2630.3
125.9	2398.8
257.0	1905.5
123.0	851.1
1000.0	3467.4
120.2	1380.4
128.8	575.4

The new method was developed because the old method, while more accurate, used expensive reagents. The investigators wish to see if estimates from the new metehod could be used to predict antibody levels from the old method.

Questions

1. (10 marks) Confirm the relationship between correlation and simple linear regression as described in the notes. Calculate the slope estimate using the two methods described.

2. (10 marks) Regress the readings from the old method on the readings from the new method. Report the SST, SSR and SSE.
3. (5 marks) Estimate the variance of the model.
4. (5 marks) Report the F statistic and its degrees of freedom.
5. (10 marks) Report the p-value and interpret it without recourse to statistical significance.
6. (10 marks) Calculate r^2 and confirm that the two methods of calculation described in the notes are the equivalent.
7. (6 marks) Estimate the mean value of tuberculin antibodies under the old method if the new method gives results of 200, 500 and 750.
8. (6 marks) Construct an appropriate interval for each of your point estimates in (7) above.
9. (12 marks) Interpret each of the three intervals in (8) above.
10. (10 marks) Construct a properly formatted scatterplot showing the individual data points, the regression line, and 90% confidence and prediction intervals for the data.

THE END