ALSM: Chapter 3

Simultaneous Inferences and Other Topics in Regression Analysis

Darshan Patel

12/27/2020

```
library(tidyverse)
library(latex2exp)
library(gridExtra)
library(wesanderson)
library(broom)
theme_set(theme_minimal())
```

Problem 1:

When joint confidence intervals for β_0 and β_1 are developed by the Bonferroni method with a family confidence coefficient of 90 percent, does tihs imply that 10 percent of the time the confidence interval for β_0 will be incorrect? That 5 percent of the time the confidence interval for β_0 will be incorrect and 5 percent of the time that for β_1 will be correct? Discuss.

Answer:

Problem 2:

Refer to Problem 2.1. Suppose the student combines the two confidence intervals into a confidence set. What can you say about the family confidence coefficient for this set?

Answer:

Problem 3:

Refer to *Copier maintenance* Problem 1.20.

(a) Will $b_0 $ and b_1 tend to err in the same direction or in opposite directions here? Explain.

Answer:

(b) Obtain Bonferroni joint confidence intervals for β_0 and β_1, using a 95 percent familt confidence coefficient.

Answer:

(c) A consultant has suggested that β_0 should be 0 and β_1 should equal 14.0. Do your joint confidence intervals in part (b) support this view?

Answer:

Problem 4:

Refer to *Airfreight breakage* Problem 1.21.

(a) Will b_0 and b_1 tend to err in the same direction or in opposite directions here? Explain.

Answer:

(b) Obtain Bonferroni joint confidence intervals for β_0 and β_1, using a 99 percent family confidence coefficient. Interpret your confidence intervals.

Answer:

Problem 5:

Refer to *Plastic hardness* Problem 1.22.

(a) OBtain Bonferroni join confidence intervals for β_0 and β_1, using a 90 percent family confidence coefficient. Interpret your confidence intervals.

Answer:

(b) Are b_0 and b_1 positively or negatively correlated here? Is this reflected in your joint confidence intervals in part (a)?

Answer:

(c) What is the meaning of the family confidence coefficient in part (a)?

Answer:

Problem 6:

Refer to *Muscle mass* Problem 1.27.

(a) Obtain Bonferroni joint confidence intervals for β_0 and β_1, using a 99 percent family confidence coefficient. Interpret your confidence intervals.

Answer:

(b) Will b_0 and b_1 tend to err in the same direction or in opposite directions here? Explain.

Answer:

(c) A researcher has suggested that β_0 should equal approximately 160 and that β_1 should be between -1.9 and -1.5. Do the joint confidence intervals in part (a) support this expectation?

Answer:

Problem 7:

Refer to *Copier maintenance* Problem 1.20.

(a) Estimate the expected number of minutes spent when there are 3, 5, and 7 copiers to be serviced, respectively. Use interval estimates with a 90 percent family confidence coefficient based on the Working-Hotelling procedure.

Answer:

(b) Two service calls for preventive maintenance are scheduled in which the numbers of copiers to be serviced are 4 and 7, respectively. A family of predicition intervals for the times to be spent on these calls is desired with a 90 percent family confidence coefficient. Which procedure, Scheffé or Bonferroni, will provide tigher prediction limits here?

Answer:

(c) Obtain the family of prediction intervals required in part (b), using the most efficient procedure.

Answer:

Problem 8:

Refer to *Airfreight breakage* Problem 1.21.

(a) It is desired to obtain interval estimates of the mean number of broken ampules when there are 0, 1, and 2 transfers for a shipment, using a 95 percent family confidence coefficient. Obtain the desired confidence intervals, using the Working-Hotelling procedure.

Answer:

(b) Are the confidence intervals obtained in part (a) more efficient than Bonferroni intervals here? Explain.

Answer:

(c) The next three shipments will make 0, 1, and 2 transfers, respectively. Obtain prediction intervals for the number of broken ampules for each of these three shipments, using the Scheffé procedure and a 95 percent family confidence coefficient.

Answer:

(d) Would the Bonferroni procedure have been more efficient in developing the prediction intervals in part (c)? Explain.

Answer:

Problem 9:

Refer to *Plastic hardness* Problem 1.22.

(a) Management wishes to obtain interval estimates of the mean hardness when the elapsed time is 20, 30, and 40 hours, respectively. Calculate the desired confidence intervals using the Bonferroni procedure and a 90 percent family confidence coefficient. What is the meaning of the family confidence coefficient here?

Answer:

(b) Is the Bonferroni procedure employed in part (a) the most efficient one that could be employed here? Explain.

Answer:

(c) The next two test items will be measured after 30 and 40 hours of elapsed time, respectively. Predict the hardness for each of these two items, using the most efficient procedure and a 90 percent family confidence coefficient.

Answer:

Problem 10:

Refer to *Muscle mass* Problem 1.27.

(a) The nutritionist is particularly interested in the mean muscle mass for women aged 45, 55 and 65. Obtain joint confidence intervals for the means of interest using the Working-Hotelling procedure and a 95 percent family confidence coefficient.

Answer:

(b) Is the Working-Hotelling procedure the most efficient one to be employed in part (a)? Explain.

Answer:

(c) Three additional women aged 48, 59 and 74 have contacted the nutritionist. Predict the muscle mass for each of these three women using the Bonferroni procedure and a 95 percent family confidence coefficient.

Answer:

(d) Subsequently, the nutrionist wishes to predict the muscle mass for a fourth woman aged 64, with a family confidence coefficient of 95 percent for the four predictions. Will the three prediction intervals in part (c) have to be recalculated? Would this also be true if the Scheffé procedure had been used in constructing the prediction intervals?

Answer:

Problem 11:

A behavioral scientist said, "I am never sure whether the regression line goes through the origin. Hence I will not use such a model." Comment.

Answer:

Problem 12:

Typographical errors. A firm specializing in technical manuscripts has a random sample of recent orders, where X is the number of galleys for a manuscript and Y is the total dollar cost of correcting typographical errors. Since Y involves variable costs only, an analyst wished to determine whether regression-through-the-origin model is appropriate for studying the relationship between the two variables.

(a) Fit the regression model prescribed above and state the estimated regression function.

Answer:

(b) Plot the estimated regression function and the data. Does a linear regression function thought the origin appear to provide a good fit here? Comment.

Answer:

(c) In estimating costs of handling prospective orders, management has used a standard of $17.50 per galley for the cost of correcting typographical errors. Test whether or not this standard should be revised; use $\alpha = .02$. State the alternatives, decision rules and conclusion.

Answer:

(d) Obtain a prediction interval for the correction cost on a forthcoming job involing 10 galleys. Use a confidence coefficient of 98 percent.

Answer:

Problem 13:

Refer to *Typographical errors* Problem 4.12.

(a) Obtain the residuals e_i. Do they sum to zero? Plot the residuals against the fitted values $\hat{Y}_i$. What conclusions can be drawn from your plot?

Answer:

(b) Conduct a formal test for lack of fit of linear regression through the origin; use $\alpha = .01$. State the alternatives, decision rule and conclusion. What the P-value of the test?

Answer:

Problem 14:

Refer to *Grade point average* Problem 1.19. Assume that linear regression through the origin model is appropriate.

(a) Fit regression model and state the estimated regression function.

Answer:

(b) Estimate β_1 with a 95 percent confidence interval. Interpret your interval estimate.

Answer:

(c) Estimate the mean freshman GPA for students whose ACT test score is 30. Use a 95 percent confidence interval.

Answer:

Problem 15:

Refer to *Grade point average* Problem 4.14.

(a) Plot the fitted regression line and the data. Does the linear regression function through the origin appear to be a good fit here?

Answer:

(b) Obtain the residuals e_i. Do they sum to zero? Plot the residuals against the fitted values $\hat{Y}_i$. What conclusions can be drawn from your plot?

Answer:

(c) Conduct a formal test for lack of fit of linear regression through the origin; use $\alpha = .005$. State the alternatives, decision rule and conclusion. What is the P-value of the test?

Answer:

Problem 16:

Refer to *Copier maintenance* Problem 1.20. Assume that linear regression through the origin mdodel is appropriate.

(a) Obtain the estimated regression function.

Answer:

(b) Estimate β_1 with a 90 percent confidence interval. Interpret your interval estimate.

Answer:

(c) Predict the service time on a new call in which six copiers are to be serviced. Use a 90 percent prediction interval.

Answer:

Problem 17:

Refer to *Copier maintenance* Problem 4.16.

(a) Plot the fitted regression line and the data. Does the linear regression through the origin appear to be a good fit here?

Answer:

(b) Obtain the residuals e_i. Do they sum to zero? Plot the residuals against the fitted values $\hat{Y}_i$. What conclusions can be drawn from your plot?

Answer:

(c) Conduct a formal test for lack of fit of linear regression through the origin; use $\alpha = .01$. State the alternatives, decision rule and conclusion. What is the P-value of the test?

Answer:

Problem 18:

Refer to *Plastic hardness* Problem 1.22. Suppose that errors arrise in X because the laboratory technician is instructed to measure the hardness of the ith specimen (Y_i) at a prerecorded elapsed time (X_i), but the timing is imperfect so the true elapsed time varies at random from the prerecorded elapsed time. Will ordinary least squares estimates be biased here? Discuss.

Answer:

Problem 19:

Refer to *Grade point average* Problem 1.19. A new student earned a grade point average of 3.4 in the freshman year.

(a) Obtain a 90 percent confidence interavl for the students' ACT test score. Interpret your confidence interval.

Answer:

(b) Is criterion (4.33) as to the appropriateness of the approximate confidence interval met here?

Answer:

Problem 20:

Refer to *Plastic hardness* Problem 1.22. The measurement of a new test item showed 238 Brinell units of hardness.

(a) Obtain a 99 percent confidence interval for the elapsed time before the hardness was measured. Interpret your confidence interval.

Answer:

(b) Is criterion (4.33) as to the appropriateness of the approximate confidence interval met here?

Answer:

Problem 21:

When the predictor variable is coded that $\overline{X} = 0$ and the normal error regression model applies, are b_0 and b_1 independent? Are the joint confidence intervals for β_0 and β_1 then independent?

Answer:

Problem 22:

Derive an extension of the Bonferroni inequality (4.2a) for the case of three statements, each with statement confidence coefficients $1 - \alpha$.

Answer:

Problem 23:

Show that for the fitted least squares regression line through the origin (4.15), $\sum X_i e_i = 0$.

Answer:

Problem 24:

Show that $\hat{Y}$ as defined in (4.15) for linear regression through the origin is an unbiased estimator of $E[Y]$.

Answer:

Problem 25:

Derive the formula for $s^2[\hat{Y}_h]$ for linear regression through the origin.

Answer:

Problem 26:

Refer to the *CDI* dataset in Appendix C.2 and Project 1.43. Consider the regression relation of number of active physicians to total population.

(a) Obtain Bonferroni joint confidence intervals for β_0 and β_1, using a 95 percent family confidence coefficient.

Answer:

(b) An investigator has suggested that β_0 should be -100 and β_1 should b .0028. Do the joint confidence intervals in part (a) support this view? Discuss.

Answer:

(c) It is desired to estimate the expected number of active physicians for countries with total population of $X = 500,\ 1,000,\ 5,000$ thousands with family confidence coefficient .90. Which procedure, the Working-Hotelling or the Bonferroni, is more efficient here?

Answer:

(d) Obtain the family of interval estimates required in part (c), using the more efficient procedure. Interpret your confidence intervals.

Answer:

Problem 26:

Refer to the *SENIC* dataset in Appendix C.1 and Project 1.45. Consider the regression relation of average length of stay to infection risk.

(a) Obtain Bonferroni join confidence intervals for β_0 and β_1, using a 90 percent family confidence coefficient.

Answer:

(b) A researcher suggested that β_0 should be approximately 7 and β_1 should be approximately 1. Do the joint intervals in part (a) support this expectation? Discuss.

Answer:

(c) It is desired to estimate the expected hospital stay for persons with infection risks $X = 2,\ 3,\ 4,\ 5$ with family confidence coefficient .95. Which procedure, the Working-Hotelling or the Bonferroni, is more efficient here?

Answer:

(d) Obtain the family of interval estimates required in part (c), using the more efficient procedure. Interpret your confidence intervals.

Answer: