

STATS 500 - Homework 3

Due in class Wednesday, October 4, 2017

1. Using the `sat` data (see `help(sat)` for the description of variables) from "faraway" package:

1. Fit a model with `total` sat score as the response and `takers`, `ratio` and `salary` as predictors. Comment on the coefficients and the goodness of fit. Test the hypothesis $\beta_{salary} = 0$. Test the hypothesis $\beta_{takers} = \beta_{ratio} = \beta_{salary} = 0$.
2. Compute 95% and 99% CIs for the parameter associated with `salary`. Using just these intervals, what can we deduce about the p -value for `salary` in the regression summary?
3. Compute and display a 95% joint confidence region for the parameters associated with `ratio` and `salary`. Add the origin to the plot. The location of the origin on the plot tells us the outcome of a certain hypothesis test. State that test and its outcome.
4. Now add `expend` (current expenditure per pupil) to the model and comment on the coefficients, their significance and the goodness of fit as compared to the model in question 1.
5. In the model of question 4, test the hypothesis $\beta_{salary} = \beta_{expend} = \beta_{ratio} = 0$. Based on your entire analysis, do you feel any of these predictors have an effect on the response?
6. Utilizing the model in 1., generate plot of the predicted value of total sat score as function of salary for `takers` = 35 and `ratio` = 17, and range of salary being 25 to 52. On this plot, also show the 95% confidence interval for the mean and the 95% prediction interval for the actual total sat observation, and provide some comments/discussion on what all of these plots show.

Hints: Useful R functions for the homework: `library()`, `data()`, `lm()`, `summary()`, `anova()`, `confint()`, `ellipse()`, `plot()`, `points()`.

2. Based on Chapter 3, problem 5 (p. 50).

Find a formula relating R^2 and the F -test (statistic) for the regression.

Solutions to this problem homework should be about 3 pages long.