

Introduction to R

Session 7: Fitting models

1. Using the built-in `mtcars` dataset:
 - a. Implement a t-test of the null hypothesis that the average miles per gallon is equal in automatic and manual cars. Check that your output is sensible by plotting a boxplot of the same data.
 - b. Implement analysis of variance, to assess whether the mean miles per gallon is equal in cars with different numbers of forward gears. Again, check your output using a boxplot, and be careful to use a factor representation of the `gear` variable.
2. Again using the `mtcars` dataset, implement linear regression of miles per gallon on weight. How does this compare to your “eyeball” estimate in Session 3? Obtain a p-value assessing the hypothesis that the linear trend in this dataset is flat – how does it compare to the permutation p-value from Session 5?
3. Obtain a 95% confidence interval for the linear trend between LSAT and GPA, in the `lawschool` dataset. Compare this with what you got in Session 6.
4. [For keen people!] The `titaniclong` dataset on the course site contains individual Titanic survival data – each row of the dataset represents one person. Implement logistic regression of survival (1/0) on class, and produce 95% confidence intervals for the odds ratios comparing survival in 1st and 2nd classes to 3rd class.