Homework 2

Spring 2025

Instructions

- The homework is due on **Feb. 14, 2025**, 11:59 pm.
- The homework should be submited electronically via Canvas.
- Please upload a single PDF containing the solutions in the correct order. If you include scanned images, make sure that they are organized and easy to read.

---

1. (20 points) Consider the sample correlation coefficient $\hat{\rho}$ (defined in pg. 43) and the coefficient of determination R^2 (defined in pg. 25) for simple linear regression. Prove that $\hat{\rho}^2 = R^2$ (as claimed on pg. 43).

2. (80 points) The data file `TreeAgeDiamSugarMaple.txt` is a textbook dataset. The data are from 27 maple trees. The first column of the file is `x=tree diameter` and the second column is `y=tree age` (in years). Do the following for these data:

   (i) Determine a good polynomial regression model for this data using the AIC and/or BIC criteria. (Fit all polynomial regression models upto a maximum degree of 8 and then choose the best one.)

   (ii) For the second degree polynomial model, test the null hypothesis that the coefficient of x^2 is 0. Use a level of significance: $\alpha = 0.05$. Report the p-value and make corresponding statistical decision.

   (iii) Produce a plot of the residuals versus the predicted values for the model chosen in part (i) above. Is there anything remarkable about the plot?

   (iv) Using the model in part (i), predict the age of a tree with a diameter of 110 by using an interval in which you have 95% confidence (i.e. obtain a 95% prediction interval for Y at $x = 110$.)