HW 4 - due 05/12 at 11:59 pm.

Math 181B, Spring 22, Rava

Follow closely the 'Hw guide' under Files in the folder 'Course Contents' on how to write, scan and submit your homework.
On any problem involving R, you should include your code and output as part of your answer. You may take a screenshot of the code/output, or write it by hand.
Be careful with notation, remember to define the parameters and the random variables you intend to use.

1 Exercise 1

[3 points] Consider the regression line $\hat{y} = a + bx$. Prove that, if $x = \bar{x} + cs_x$ for a constant $c > 0$, $\hat{y} = \bar{y} + crs_y$. This means that, if x is c standard deviations above the mean, we predict y to be rc standard deviations above/below the mean.

2 Exercise 2

Consider the equation $y = \frac{x}{a+bx}$ where a, b are constants.

a) [3 points] Linearize the equation and derive formulas for a and b.

b) [2 points] Write a function in R that, given a vector x and a vector y, compute a,b for this equation. (Arguments: x,y, Output: a,b). Here don't use the built-in function lm but use the formulas for a, b derived in part a.

3 Exercise 3

[5 points] In this exercise we investigate what happens to the distribution of $\hat{\beta}_1$ in linear regression when the homoscedasticity assumption is not satisfied. Assume that, given $x_1, \ldots, x_n$, $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, where $\epsilon_i \sim N(0, \sigma_i^2)$. Assume that ϵ_i are independent. This is similar to the simple linear model we have used in class with the exception that the assumption of homoscedasticity is not satisfied. Consider the estimator for the slope $\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x}) Y_i}{\sum_{i=1}^n (x_i - \bar{x})^2}$. Prove that $\hat{\beta}_1 \sim N\left(\beta_1, \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \sigma_i^2}{\left\{\sum_{i=1}^n (x_i - \bar{x})^2\right\}^2}\right)$.