

FALL 2021 MATH 484/564 HOMEWORK #1

Due: September 11, 11:59PM, submit in blackboard

Homework solution is not required to be typed, but must be legible.

Problem 1 The purity of oxygen (Y) produced by a fractional distillation process is thought to be related to the percentage of hydrocarbons (X) in the main condensor of the processing unit. Twenty samples were measured and shown in the attached data sheet.

- 1) Fit a simple linear regression model to the data, provide the linear regression line equation $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$.
- 2) Test the hypothesis $H_0 : \beta_1 = 0$ against the alternative $H_A : \beta_1 \neq 0$, and conclude if there is significant linear relationship between the purity of oxygen and the percentage of hydrocarbons.
- 3) Calculate the coefficient of determination, r^2 .
- 4) Find a 95% confidence interval on the slope.
- 5) Find a 95% confidence interval on the mean purity when the hydrocarbon percentage is 1.05.
- 6) What is the correlation coefficient between Y and X ?
- 7) Test the hypothesis: $H_0 : \rho = 0$ against $H_A : \rho \neq 0$ using a t-test based on the correlation coefficient computed from the previous step.

Problem 2 Consider the simple linear regression model $Y = \beta_0 + \beta_1 X + \epsilon$. Assume that $E(\epsilon) = 0$ and $Var(\epsilon) = \sigma^2$, and ϵ_i are independent of each other. Show that $Cov(\bar{Y}, \hat{\beta}_1) = 0$.

Problem 3 Show that the sample correlation coefficient r between X and Y is a value between -1 and 1, i.e., $-1 \leq r \leq 1$. Use two different methods to show it, with one using only the data, and the other one in the context of simple linear regression.

Problem 4 Consider the simple linear regression model $Y = \beta_0 + \beta_1 X + \epsilon$. Assume that $\epsilon_i \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$. Find the variance of \hat{Y}_h , where $\hat{Y}_h = \hat{\beta}_0 + \hat{\beta}_1 x_h$.

Problem 5 Exercise 2.9 from the TEXT.