FALL 2023 MATH 484/564 HOMEWORK #2

- Due: September 27, Wednesday, 11:59PM
- How to submit: via Blackboard. If you have multiple files, upload a zipped file
- Submission link will disappear after 48 hours
- Homework solution is not required to be typed, but must be legible.

Problem 1 Find the least square estimate for β_1 when the regression model becomes $Y = \beta_1 X + \epsilon$. This is called regression through the origin (RTO).

Problem 2 First generate the dataset $(x_i, y_i)_{i=1}^n$ from $Y = 2X + \epsilon$, where ϵ follows normal distribution $\mathcal{N}(0, 0.1^2)$, and X is distributed in the range of [-1, 1]. Let $n = 100$. Then use regression through the origin (RTO) model and simple linear regression (SLR) model to fit the same data, and report the value of $\sum\limits_{i=1}^{n} e_i$ and the coefficient of determination for both RTO and SLR. For SLR, report R^2; for RTO, report R^2_{RTO}.

Problem 3 Consider the coefficient of determination R^2 in SLR and MLR.

1) Show that $R^2 = (\mathrm{Cor}(Y, \hat{Y}))^2$ holds for both simple linear regression and for multiple linear regression.

2) Show that $\mathrm{Cor}(Y, \hat{Y}) = |\,\mathrm{Cor}(Y, X)|$ in SLR.

Problem 4 Prove that in the hypothesis testing for the slope β_1 in SLR and MLR,

$$\begin{cases} H_0: & \beta_1 = 0 \\ H_A: & \beta_1 \neq 0 \end{cases}$$

the t-test statistic $t^\star$ has the following relation with the F-test statistic $F^\star$:

$$(t^\star)^2 = F^\star$$

First show the relationship is true for SLR with only one predictor variable, then show it is also true for MLR with $p > 1$ predictor variables.

Problem 5 Exercise 3.14 from the TEXT.