Homework 1

MATH 4387/5387 Fall 2020

Subrata Paul

8/24/2020

Problem 1

On canvas there is a data named `simu_hw1.txt` in the data folder. Download the data and read it in R using `read.table` function.

- What are the names of the columns?
- Plot a boxplot of the `pred` variables.
- Do all the `pred` variables look symmetric? Is it clear from the plot? If not what you can try to do to understand the `pred4` variable? Get a closer look on `pred4` variable. Does it look symmetric?
- How would you compare the mean of the `pred` variable? (Are they somewhat equal or some variable has higher mean than other? You don't need to do any analysis. Just your comments on the boxplot.)
- Make a density plot of all the variables with clear legend.

Problem 2

- Using appropriate plot, comment on relationship between the `response` and other variables?
- Normalize (make mean = 0 and sd = 1) the `pred4` and plot densities of the original variable and the normalized version with legend.

Problem 3

Consider the joint distribution (pdf) of two random variables X and Y,

$$f(x, y) = \frac{1}{\sqrt{3}\pi} \exp\left[-\frac{2}{3}(x^2 - xy + y^2)\right]$$

- Find the marginal pdf: f_x and f_y
- Find the expected values and variances of X and Y
- Find the covariance between X and Y.
- (MATH 5387) Use R to get the probability of $P(-0.2 \leq X < 0.1)$?

Problem 4 (MATH 5387)

Prove the identities. Here $\mathbf{x}, \mathbf{y}$ are random vectors and A is a matrix of constants. A^T represents transpose of A.

- $E(A\mathbf{y}) = AE(\mathbf{y})$
- $E(\mathbf{x} + \mathbf{y}) = E(\mathbf{x}) + E(\mathbf{y})$
- $var(A\mathbf{y}) = A\ var(\mathbf{y})A^T$
- $cov(\mathbf{x} + \mathbf{y}, \mathbf{z}) = cov(\mathbf{x}, \mathbf{z}) + cov(\mathbf{y}, \mathbf{z})$
- $cov(\mathbf{x}, \mathbf{y} + \mathbf{z}) = cov(\mathbf{x}, \mathbf{y}) + cov(\mathbf{x}, \mathbf{z})$
- $cov(A\mathbf{x}, \mathbf{y}) = A\ cov(\mathbf{x}, \mathbf{y})$ and $cov(\mathbf{x}, A\mathbf{y}) = cov(\mathbf{x}, \mathbf{y})A^T$.