Statistical Learning for Engineers (EN.530.641)
Homework 2

Jin Seob Kim, Ph.D.

Out: 09/16/2022
due: 9/23/2022 by midnight EST

1 A ball is drawn randomly from a box that contains 6 black balls, 4 white balls, and 5 red balls.

(a) Calculate the probability that the ball drawn is black.

(b) Calculate the probability that the ball is not black.

(c) Calculate the probability that the ball is either black or white.

2 You toss a fairly well-made coin twice. Let A be the event that at least one head appears, and B be the event that two heads appear.

(a) Compute $P(A \cup B)$.

(b) Are A and B mutually exclusive?

3 Let's consider a family with 2 children. If at least one child in the family is a boy, calculate the probability that both children are boys:

(a) without using conditional probability;

(b) with using conditional probability.

4 (a) Let X be a discrete random variable with the binomial distribution, i.e., $X \sim \text{Bin}(p, n)$ as in the class. Show that

$$E[X] = np$$
$$\text{Var}[X] = np(1-p).$$

(b) Let X be a discrete random variable with the Bernoulli distribution, i.e., $X \sim \text{Ber}(p)$ as in the class. Compute $E[X]$ and $\text{Var}[X]$.

5 Let $\mathbf{x} \in \mathbb{R}^n$ be continuous random variable with the corresponding PDF $f(\mathbf{x})$. Let $\boldsymbol{\mu} = E[\mathbf{x}]$ and $\Sigma = \text{Cov}[\mathbf{x}]$. Now let us consider the transformation $\mathbf{y} = A\mathbf{x} + \mathbf{b}$, where $A \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$ (constant). Show that

$$E[\mathbf{y}] = A\boldsymbol{\mu} + \mathbf{b}$$
$$\text{Cov}[\mathbf{y}] = A\Sigma A^T.$$

6 Use python to generate N sample points from the normal distribution (i.e., Gaussian distribution with zero mean and unity standard deviation). Try with three different N's: $N = 1,000,\ 5,000,\ 100,000$. Then make a 'normalized histogram' (equivalent to PDF) of each case. Also overlap the Gaussian $f(x; 0, 1)$ in the plots together with the histograms to compare.

Submission Guideline

- For analytic parts (e.g., problems 1,2,3,4,5 and plots of problem 6), submit your homework answers in a single pdf format, including plots, to "HW2_analytical" on the gradescope.
  - No more than two (2) homework problems may be on the same page. In other words, for each problem your answers should be on a separate set of pages (e.g., for Problem 1, your answer is on page 1–3, and for Problem 2, your answer is on page 4-5, and so on).
  - Then when submitting, you should assign the pages to each problem on Gradescope. You can scan your answers or use an app (e.g., Adobe Scan) to generate a pdf file. *Show your work.*
- Submit all your python codes in a single .zip file that contains codes for each problem (name them by including the problem number). Name your single zip file submission as "YourName_HW2.zip". For example, "JinSeobKim_HW2.zip" for a single zip file. Submission will be done through "HW2_computational" on the gradescope.
- Just in case you have related separate files, please make sure to include *all the necessary files.* If TAs try to run your function and it does not run, then your submission will have a significant points deduction.
- Make as much comments as possible so that the TAs can easily read your codes.