STAT2203/7203 (S2-2023): Assignment 01
Due: 18-August-2023 @16:59

---

1. Comparative Study [**2 marks each**]

An experiment is conducted to determine the optimal time and temperature combination for baking a cake. The response variable of interest is taste ("Great", "Mediocre", or "Terrible"). Six batches of cake will be baked separately at each combination of baking times (25 and 30 minutes) and temperature settings ($275°F$, $300°F$, and $325°F$).

(a) What are the experimental units, i.e., the group of subjects on which measurements are made?

(b) What are the factors in this experiment?

(c) State the levels of each factor.

(d) List all the treatments in this experiment.

(e) Is the response variable qualitative or quantitative?

---

2. Visualization [**10 marks**]

An article reports on the results of a cloud-seeding experiment[a]. The question of interest is whether cloud seeding with silver nitrate increases rainfall. Out of 52 clouds, 26 were randomly selected for seeding, with the remaining 26 serving as controls. The rainfall measurements, in acre-feet, are given in `CloudSeedingData.csv`. Use `R` (or any other programming language of your choice) to construct a comparative boxplot and comment on possible differences in rainfall between seeded and unseeded clouds. **Hint:** If you are using `R`, the command `stack` might be useful.

----
[a]J. Simpson, A. Olsen, and J. C. Eden (1975). Technometrics, 17: 161–166.

---

3. Counting [**5 marks each**]

Independently throw 10 balls into one of three boxes, numbered 1, 2, and 3, with probabilities 1/4, 1/2, and 1/4, respectively.

(a) What is the probability that box 1 has 2 balls, box 2 has 5 balls, and box 3 has 3 balls?

(b) What is the probability that box 1 remains empty?

(c) Write a script in `R` (or any other programming language of your choice) to verify your answers to parts (a) and (b) by repeating the experiments many times.

4. Conditional Probability [25 marks]

Suppose email arriving in you inbox is spam 30% of the time. With probability 0.9 your spam filter will correctly flag an email as spam. However, your spam filter will also flag non-spam emails as spam with probability 0.05. You received an email today which your spam filter has flagged. What is the probability that it really is spam?

5. Discrete Random Variable [5 marks each]

In 3 spot Keno a player picks 3 numbers from $\{1, 2, \ldots, 80\}$. A machine picks 20 numbers from $\{1, 2, \ldots, 80\}$. The payout to the player depends on how many matches the player gets.

Matches	0	1	2	3
Payout(\$)	0	0	1	44
Probability	0.417	0.430	0.139	0.014

(a) Compute the probability that the player wins any money in a single game.

(b) Compute the expectation and standard deviation of the number of matches in a game of 3 spot Keno.

6. Continuous Random Variable [5 marks each]

A continuous random variable is said to have a logistic distribution if its pdf is given by

$$f(x) = \frac{e^{-x}}{(1 + e^{-x})^2}, \quad x \in \mathbb{R}.$$

(a) Plot the graph of the pdf using `R` (or any other programming language of your choice).

(b) Show that $\mathbb{P}(X > x) = \dfrac{1}{1 + e^x}$ for all x.

7. Independence [10 marks]

If A and B are independent, show that A and $B^{\complement}$ are also independent.

8. Expectation [5 marks each]

(a) Let X and Y be any two random variables. Prove that

$$\mathbb{E}(\max\{X, Y\}) = \mathbb{E}(X) + \mathbb{E}(Y) - \mathbb{E}(\min\{X, Y\}).$$

Interestingly, this is analogous to the probability law $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

(b) A random variable X is said to have a discrete uniform distribution on the set $\{a, a+1, \ldots, b\}$ if

$$\mathbb{P}(X = x) = \frac{1}{b - a + 1}, \quad x = a, a+1, \ldots, b.$$

Find the moment generating function of X.

100 marks in total

Note:

- This assignment counts for 10% of the total mark for the course.

- Although not mandatory, if you could type up your work, e.g., `LaTex`, it would be greatly appreciated.

- Show all your work and attach your code and all the plots (if there is a programming question).

- Combine your solutions, all the additional files such as your code and numerical results, **all in one single PDF file**.

- Please submit your single PDF file on Blackboard.