Mid-Term Exam in Inferential Statistics and Causal Reasoning

Assignment: Probability Theory Fundamentals

Submission Deadline: 13/02/2025 - 16:00 (GMT)

Instructions

- This coursework is to be completed **individually**.
- Provide clear and concise answers to all theoretical questions.
- For coding tasks, include well-commented code and any necessary explanations.
- Submit your answers in a single **PDF** document. Include code snippets within the document where applicable.
- Ensure all sources are properly cited if references are used.
- Make sure that you CID is visible at the beginning of the submission. DO NOT put your name in the submission as we will grade each submission anonymously.

Assignment Overview

This assignment is designed to assess your understanding of key concepts in Statistics and Probability Theory that were covered in the first part of the module.

Questions

1. Simulation: Sum of Two Dice [5 points]
You have two fair 6-sided dice. Each die has faces {1,2,3,4,5,6} with equal probability.

(a) Write a Python snippet to simulate rolling these two dice **10,000** times and record the sum of the two faces each time.

(b) Estimate the probability that the sum is at least 9. Compare to the exact theoretical probability computed by enumerating all outcomes.

(c) If you repeat your entire simulation multiple times, about how much does this estimate fluctuate around the theoretical value?

2. Combinatorial Reasoning [5 points]

A company has 12 employees, and it wants to form a 4-person project team from these 12.

(a) How many ways can the 4-person team be chosen if we only care about *which* employees are on the team (i.e. order does not matter)?

(b) Suppose now that each of the 4 chosen employees must also have a distinct "role rank" among them: leader, co-leader, secretary, or observer. How many ways can we form the team *and* assign these 4 roles?

3. Bernoulli Distribution [5 points]

A coin is slightly biased so that it lands heads with probability $p = 0.55$. Let X be the random variable for heads ($X = 1$ if heads, $X = 0$ if tails).

(a) Write down the pmf of X.

(b) Compute $E[X]$ and $\mathrm{Var}(X)$, explicitly or using Python.

(c) Show a short Python snippet (using `scipy.stats.bernoulli`) to compute $\Pr(X = 1)$.

4. Binomial Distribution [10 points]

A factory produces items, each of which is "good" with probability 0.9, independently of others. Let Y be the number of good items out of 30 produced in a day.

(a) State the distribution of Y.

(b) Write $\Pr(Y = k)$.

(c) Compute $\Pr(Y \geq 28)$ explicitly or using a binomial cdf. Show briefly how to do it using `scipy.stats.binom`.

(d) Find $E[Y]$ and $\mathrm{Var}(Y)$, explicitly or using Python.

5. Discrete Uniform Distribution [5 points]

Let Z be a random integer chosen uniformly from $\{1, 2, \ldots, 20\}$.

(a) State the pmf $p_Z(k)$ for $k = 1, \ldots, 20$.

(b) Find $E[Z]$ and $\mathrm{Var}(Z)$, explicitly or using Python.

(c) Compute $\Pr(Z \leq 5)$.

6. Continuous Uniform Distribution [10 points]
A random variable W is uniformly distributed on the interval $[2, 5]$.

(a) Explain why W is called Uniform$(2, 5)$.

(b) Write the pdf of W.

(c) Find $E[W]$ and Var(W), explicitly or using Python.

(d) Compute $\Pr(W > 4)$.

7. Normal (Gaussian) Distribution [5 points]
Suppose exam scores in a certain course follow a normal distribution $N(\mu = 50,\ \sigma^2 = 16)$. Let S be a student's exam score.

(a) Write the pdf of S.

(b) Find $\Pr(S \leq 45)$ by standardizing or using `scipy.stats.norm`.

(c) Compute $\Pr(48 \leq S \leq 55)$.

8. Simulation: Averages of Bernoulli(0.4) [5 points]
Consider a Bernoulli random variable with $p = 0.4$. We draw 100 independent samples of this variable and compute their average. We repeat that experiment 2,000 times and store all 2,000 sample means.

(a) Write Python code to simulate this process, collecting 2,000 averages.

(b) Plot or describe the distribution of these 2,000 means. By the Central Limit Theorem, how does it compare to a normal distribution with mean 0.4 and variance $\frac{0.4 \times 0.6}{100}$? Plot the results and comment.

Grading Criteria

- **Clarity and Precision:** Answers should be clear, well-organized, and precise.
- **Depth of Understanding:** Demonstrate a thorough understanding of the concepts.
- **Mathematical Rigor:** Provide complete derivations and calculations where requested.
- **Code Quality (for coding tasks):** Code should be correct, efficient, and well-documented.
- **Critical Thinking:** Show insightful analysis and original thought, especially in extended responses.

Total Points: 50

Submission Instructions

- Submit your PDF document via Blackboard by the deadline indicated at the top of this document.
- Ensure that your student ID is included on the first page of your document. Do not add your name to the PDF.
- Make sure that you submit your work in advance of the time of the deadline. Late submissions will be penalized according to the college late submission policy, which can be found here. The policy states that any work submitted up to one (1) day after the assessment deadline (date and time) will be marked but capped at the passmark. Work submitted more than one (1) day late will not be accepted as a valid attempt and mark of zero will be recorded.

Academic Integrity

Remember to adhere to Imperial Code of Conduct, available here. All work must be your own, and any sources used must be properly cited.

Good Luck!

Note: If you have any questions or need clarifications regarding the assignment, please contact Pietro or the teaching assistants well before the deadline.