

Homework 1

Due: Feb 28, 2026, 11:55pm

Instruction: Submit your answers electronically through Moodle. In Moodle, you must submit under the PDF Submission link with ONE PDF file containing answers to all remaining questions.

Question 1

In class, we solved this problem using two different methods:

We divide a class of 16 students — 4 male, 12 female — into 4 groups of 4 students each. What is the probability that each group has exactly one male student?

Here, we would like to try a third method: *computer simulation*.

Write a program where you plot, for different values of N , the fraction of outcomes where a random permutation of these 16 students results in each group having exactly one male student. Show your plot where N ranges from 100 to 10000. What does the value approach as N increases?

(You can use any programming language you like, though the solution will be given in python.)

A possible implementation in Python is shown below.

```

1  """
2  Computer simulation for:
3  16 students (4 male, 12 female) randomly permuted into 4 groups of 4.
4  Estimate the probability that each group contains exactly one male.
5
6  This script:
7  1) Simulates for different N,
8  2) Saves data to CSV and DAT (gnuplot-friendly),
9  3) Saves a plot to PNG.
10 """
11
12 import numpy as np
13 import pandas as pd
14 import matplotlib.pyplot as plt
15 from pathlib import Path
16
17 def trial_success(rng: np.random.Generator) -> bool:
18     """Run one random permutation trial and check if it is a success."""
19     # 1) Define a sequence [1,2,3,4,0,0,...,0] and perform a random
20     #    permutation
21     arr = np.array([1,2,3,4] + [0]*12, dtype=int)
22     rng.shuffle(arr)
23
24     # 2) Find the indices of where 1 is, where 2 is, etc.
25     idx1 = int(np.where(arr == 1)[0][0])
26     idx2 = int(np.where(arr == 2)[0][0])
27     idx3 = int(np.where(arr == 3)[0][0])
28     idx4 = int(np.where(arr == 4)[0][0])
29
30     # 3) Compute the group of these indices. floor(index / 4) works.
31     g1, g2, g3, g4 = idx1//4, idx2//4, idx3//4, idx4//4
32
33     # 4) Check if they occupy different groups.
34     return len({g1, g2, g3, g4}) == 4
35
36 def estimate_fraction(N: int, rng: np.random.Generator):
37     """Run N trials and return (fraction, successes)."""
38     # 5) If yes, increment a counter. The fraction is then this counter
39     #    divided by N.
40     successes = 0
41
42     for _ in range(N):
43         if trial_success(rng):
44             successes += 1
45
46     frac = successes / N
47     return frac, successes

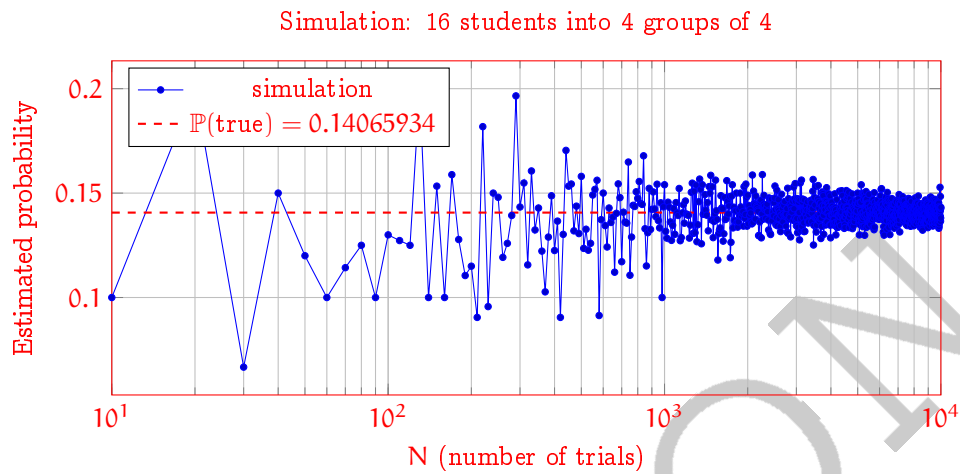
```

(continued on next page)

```

47 def main():
48     # 6) Iterate through different values of N from N_start to N_end.
49     N_start = 10
50     N_end = 10000
51
52     N_values = list(range(N_start, N_end + 1, 10))
53
54     rng = np.random.default_rng(12345) # fixed seed for reproducibility
55
56     rows = []
57     for N in N_values:
58         frac, k = estimate_fraction(N, rng)
59         rows.append({"N": N, "successes": k, "fraction": frac})
60
61     df = pd.DataFrame(rows)
62
63     out_dir = Path(".")
64     df.to_csv(out_dir / "fund_group_simulation.csv", index=False)
65
66     # gnuplot-friendly data file (space-separated)
67     with open(out_dir / "fund_group_simulation.dat", "w", encoding="utf-8") as f:
68         f.write("# N fraction\n")
69         for _, r in df.iterrows():
70             f.write(f"{int(r['N'])} {r['fraction']:.8f}\n")
71
72     # Plot (log scale for N)
73     plt.figure(figsize=(8, 4.8))
74     plt.plot(df["N"], df["fraction"], marker="o", linewidth=1)
75     plt.xscale("log")
76     plt.xlabel("N (number of trials, log scale)")
77     plt.ylabel("Estimated probability (success fraction)")
78     plt.title(
79         "Simulation: 16 students (4 male, 12 female) into 4 groups of 4\n"
80         "Each group has exactly one male"
81     )
82     plt.tight_layout()
83     plt.savefig(out_dir / "fund_group_simulation.png", dpi=300)
84     plt.close()
85
86 if __name__ == "__main__":
87     main()

```



Exact (true) probability. We randomly permute the 16 students and split the permutation into 4 consecutive blocks (groups) of size 4. We compute the probability that each group has exactly one male by a sequential argument.

Group 1. Choose exactly 1 male from 4 and 3 females from 12. The probability is

$$\frac{\binom{4}{1} \binom{12}{3}}{\binom{16}{4}}.$$

Group 2. After Group 1, there remain 3 males and 9 females. The probability that Group 2 contains exactly one male is

$$\frac{\binom{3}{1} \binom{9}{3}}{\binom{12}{4}}.$$

Group 3. After Groups 1 and 2, there remain 2 males and 6 females. The probability that Group 3 contains exactly one male is

$$\frac{\binom{2}{1} \binom{6}{3}}{\binom{8}{4}}.$$

Group 4. The remaining group is then forced to contain the last male.

Multiplying the above probabilities, we obtain

$$\mathbb{P}(\text{true}) = \frac{\binom{4}{1} \binom{12}{3}}{\binom{16}{4}} \cdot \frac{\binom{3}{1} \binom{9}{3}}{\binom{12}{4}} \cdot \frac{\binom{2}{1} \binom{6}{3}}{\binom{8}{4}} = \frac{64}{455} \approx 0.14065934.$$

Question 2

Two study rooms at the Learning Commons have different entrances: Entrance A and Entrance B. Each arriving student independently chooses Entrance A with probability 0.6 and Entrance B with probability 0.4.

At each entrance, a mask-detection system is used, and it outputs either “Mask” or “No Mask”. Unfortunately, the detector is imperfect:

- If a student is actually wearing a mask, the system outputs “Mask” with probability 0.95.
- If a student is actually not wearing a mask, the system outputs “Mask” with probability 0.08.

Let M denote the event that a student is actually wearing a mask, and G denote the event that the system outputs “Mask”. Furthermore, it is known that the proportions of students actually wearing masks differ between entrances, with

$$\mathbb{P}(M|A) = 0.70, \quad \mathbb{P}(M|B) = 0.55.$$

1. Find $\mathbb{P}(G)$.
2. Find $\mathbb{P}(M^c|G)$.
3. Find $\mathbb{P}(A|G)$.

From the question, we know

$$\mathbb{P}(A) = 0.6 \quad \mathbb{P}(B) = 0.4 \quad \mathbb{P}(G|M) = 0.95 \quad \mathbb{P}(G|M^c) = 0.08$$

1. First, let us find $\mathbb{P}(M)$ using total probability. Note that A and B are disjoint and cover all possibilities.

$$\begin{aligned} \mathbb{P}(M) &= \mathbb{P}(M|A)\mathbb{P}(A) + \mathbb{P}(M|B)\mathbb{P}(B) \\ &= 0.7 \times 0.6 + 0.55 \times 0.4 \\ &= 0.64 \\ \mathbb{P}(M^c) &= 1 - 0.64 = 0.36 \end{aligned}$$

Then, we can find $\mathbb{P}(G)$ using total probability again, where

$$\begin{aligned} \mathbb{P}(G) &= \mathbb{P}(G|M)\mathbb{P}(M) + \mathbb{P}(G|M^c)\mathbb{P}(M^c) \\ &= 0.95 \times 0.64 + 0.08 \times 0.36 \\ &= 0.6368. \end{aligned}$$

2. Using Bayes' rule,

$$\begin{aligned} \mathbb{P}(M^c|G) &= \frac{\mathbb{P}(G|M^c)\mathbb{P}(M^c)}{\mathbb{P}(G)} \\ &= \frac{0.08 \times 0.36}{0.6368} \\ &\approx 0.04523. \end{aligned}$$

3. First, let us try to find $\mathbb{P}(G|A)$. Although strictly speaking, we have not derived a total probability theorem for conditional probability, we can rather easily combine what we know to have

$$\mathbb{P}(G|A) = \mathbb{P}(G|M, A)\mathbb{P}(M|A) + \mathbb{P}(G|M^c, A)\mathbb{P}(M^c|A).$$

But what is $\mathbb{P}(G|M, A)$? It is the probability that the system outputs "Mask" given that a student has gone through Entrance A and the student is actually wearing a mask. Note that actually which entrance is not important here, because as long as the student is wearing a mask, the probability is 0.95. In other words, $\mathbb{P}(G|M, A) = \mathbb{P}(G|M)$ and similarly, $\mathbb{P}(G|M^c, A) = \mathbb{P}(G|M^c)$. Therefore,

$$\begin{aligned} \mathbb{P}(G|A) &= \mathbb{P}(G|M)\mathbb{P}(M|A) + \mathbb{P}(G|M^c)\mathbb{P}(M^c|A) \\ &= 0.95 \times 0.7 + 0.08 \times (1 - 0.7) \\ &= 0.689. \end{aligned}$$

Now, we apply Bayes' rule,

$$\begin{aligned} \mathbb{P}(A|G) &= \frac{\mathbb{P}(G|A)\mathbb{P}(A)}{\mathbb{P}(G)} \\ &= \frac{0.689 \times 0.6}{0.6368} \\ &\approx 0.64918. \end{aligned}$$

Question 3

After a particular movie screening ends, a lot of people leave the cinema and may choose to spend some additional time in the mall before departing.

We randomly select two members of the audience from this group and denote them by A and B. Let X and Y represent the lengths of time (in hours) that A and B, respectively, stay in the shopping mall after the movie ends. To simplify the analysis, we assume that the stay durations are discretized and that both X and Y take values in the finite set $\{1, 2, 3\}$, corresponding to short, medium, and long stays, respectively.

The joint PMF of (X, Y) is given by

$$p_{X,Y}(x,y) = \begin{cases} c(4 - |x - y|), & x, y \in \{1, 2, 3\}, \\ 0, & \text{otherwise,} \end{cases}$$

where $c > 0$ is a normalization constant.

1. Find c .
2. Find the marginal PMFs $p_X(x)$ and $p_Y(y)$. Are X and Y independent? Justify your answer.
3. The total time that the two audience members spend in the shopping mall is denoted by the random variable $T = X + Y$.
 - (a) Find the PMF of T.
 - (b) Find $\mathbb{E}(T)$ and $\text{var}(T)$.
4. The imbalance in the stay durations of the two audience members is measured by the random variable $D = |X - Y|$. Find $\mathbb{E}(D)$ and $\text{var}(D)$.

1. The normalization constant c must satisfy

$$\sum_{x=1}^3 \sum_{y=1}^3 c(4 - |x - y|) = 1.$$

Count the number of pairs (x, y) for each value of $|x - y|$:

- $|x - y| = 0$ (3 pairs): each contributes $4c$.
- $|x - y| = 1$ (4 pairs): each contributes $3c$.
- $|x - y| = 2$ (2 pairs): each contributes $2c$.

Thus

$$c(3 \cdot 4 + 4 \cdot 3 + 2 \cdot 2) = c(12 + 12 + 4) = 28c = 1$$

so

$$c = \frac{1}{28}.$$

2. Marginal PMF of X :

$$p_X(1) = \sum_{y=1}^3 p_{X,Y}(1, y) = \frac{4 + 3 + 2}{28} = \frac{9}{28},$$

$$p_X(2) = \sum_{y=1}^3 p_{X,Y}(2, y) = \frac{3 + 4 + 3}{28} = \frac{10}{28},$$

$$p_X(3) = \sum_{y=1}^3 p_{X,Y}(3, y) = \frac{2 + 3 + 4}{28} = \frac{9}{28}.$$

By symmetry, $p_Y(y) = p_X(y)$ for $y = 1, 2, 3$.

Check independence: $p_{X,Y}(1, 1) = \frac{4}{28}$ while $p_X(1)p_Y(1) = \frac{9}{28} \cdot \frac{9}{28} = \frac{81}{784} \neq \frac{4}{28}$. Hence X and Y are not independent.

3. (a) $T = X + Y$ takes values in $\{2, 3, 4, 5, 6\}$.

$$p_T(2) = p_{X,Y}(1, 1) = \frac{4}{28},$$

$$p_T(3) = p_{X,Y}(1, 2) + p_{X,Y}(2, 1) = \frac{3}{28} + \frac{3}{28} = \frac{6}{28},$$

$$p_T(4) = p_{X,Y}(1, 3) + p_{X,Y}(2, 2) + p_{X,Y}(3, 1) = \frac{2}{28} + \frac{4}{28} + \frac{2}{28} = \frac{8}{28},$$

$$p_T(5) = p_{X,Y}(2, 3) + p_{X,Y}(3, 2) = \frac{3}{28} + \frac{3}{28} = \frac{6}{28},$$

$$p_T(6) = p_{X,Y}(3, 3) = \frac{4}{28}.$$

- (b)

$$\mathbb{E}(T) = \sum_t t p_T(t) = \frac{2 \cdot 4 + 3 \cdot 6 + 4 \cdot 8 + 5 \cdot 6 + 6 \cdot 4}{28} = \frac{112}{28} = 4.$$

$$\mathbb{E}(T^2) = \frac{4 \cdot 4 + 9 \cdot 6 + 16 \cdot 8 + 25 \cdot 6 + 36 \cdot 4}{28} = \frac{492}{28} = \frac{123}{7}.$$

$$\text{var}(T) = \mathbb{E}(T^2) - (\mathbb{E}(T))^2 = \frac{123}{7} - 16 = \frac{11}{7}.$$

Hence $\mathbb{E}(T) = 4$ and $\text{var}(T) = \frac{11}{7}$.

4. $D = |X - Y|$ takes values in $\{0, 1, 2\}$.

$$p_D(0) = \sum_{|x-y|=0} p_{X,Y}(x, y) = \frac{4+4+4}{28} = \frac{12}{28},$$

$$p_D(1) = \sum_{|x-y|=1} p_{X,Y}(x, y) = \frac{3+3+3+3}{28} = \frac{12}{28},$$

$$p_D(2) = \sum_{|x-y|=2} p_{X,Y}(x, y) = \frac{2+2}{28} = \frac{4}{28}.$$

Then

$$\mathbb{E}(D) = 0 \cdot \frac{12}{28} + 1 \cdot \frac{12}{28} + 2 \cdot \frac{4}{28} = \frac{20}{28} = \frac{5}{7}.$$

$$\mathbb{E}(D^2) = 0^2 \cdot \frac{12}{28} + 1^2 \cdot \frac{12}{28} + 4 \cdot \frac{4}{28} = 1.$$

$$\text{var}(D) = \mathbb{E}(D^2) - (\mathbb{E}(D))^2 = 1 - \left(\frac{5}{7}\right)^2 = \frac{24}{49}.$$

Hence $\mathbb{E}(D) = \frac{5}{7}$ and $\text{var}(D) = \frac{24}{49}$.

Question 4

A computer server is monitored daily. On any given day, the server crashes with probability $p = 0.05$, independently of other days. Let X be the number of days until the first crash occurs (including the day of the crash). For example, if the server crashes on Day 1, then $X = 1$.

1. Identify the distribution of X and write down its probability mass function (PMF), $p_X(k)$.
2. Find the probability that the server runs without crashing for at least 10 days. (i.e., the first crash happens on day 11 or later, $\mathbb{P}(X > 10)$).
3. Suppose the server has already survived 10 days without a crash. What is the probability that it will survive at least another 10 days? (i.e., Find $\mathbb{P}(X > 20 | X > 10)$).
4. Calculate the expected number of days until the first crash, $\mathbb{E}(X)$.

1. X follows a geometric distribution with parameter $p = 0.05$. The PMF is

$$p_X(k) = (1 - p)^{k-1} p = (0.95)^{k-1} (0.05), \quad k = 1, 2, 3, \dots$$

2. The event $X > 10$ means that the first 10 days were all failures (no crash).

$$\mathbb{P}(X > 10) = (1 - p)^{10} = 0.95^{10} \approx 0.5987$$

Alternatively, using the sum of the PMF, we have

$$\mathbb{P}(X > 10) = \sum_{k=11}^{\infty} (1 - p)^{k-1} p = (1 - p)^{10}$$

3. Using the definition of conditional probability,

$$\begin{aligned} \mathbb{P}(X > 20 | X > 10) &= \frac{\mathbb{P}(X > 20 \cap X > 10)}{\mathbb{P}(X > 10)} \\ &= \frac{\mathbb{P}(X > 20)}{\mathbb{P}(X > 10)} \\ &= \frac{(1 - p)^{20}}{(1 - p)^{10}} \\ &= (1 - p)^{10} \\ &\approx 0.5987 \end{aligned}$$

This illustrates the memoryless property of the geometric distribution. The probability of surviving another 10 days is the same as surviving the initial 10 days, regardless of the history.

4. The expectation of a geometric random variable is $1/p$.

$$\mathbb{E}(X) = \frac{1}{p} = \frac{1}{0.05} = 20 \text{ days.}$$

Question 5

A cloud computing platform consists of two servers, Server A and Server B, that process a stream of incoming computation jobs. The jobs arrive one by one, and each arriving job is immediately assigned to exactly one of the two servers. Each job is assigned to Server A with probability p , where $\frac{1}{2} < p < 1$, and to Server B with probability $q = 1 - p$. Job assignments are independent across different jobs. Let $S_A(k)$ and $S_B(k)$ denote the total numbers of jobs assigned to Server A and Server B, respectively, after a total of k jobs have arrived.

1. Define the throughput difference as

$$D_k = S_A(k) - S_B(k),$$

which records the net number of additional jobs completed by Server A relative to Server B after k scheduling intervals.

- (a) Express D_k in terms of $S_A(k)$ and k .
- (b) List all possible values of D_2 and compute their probabilities.
- (c) Show that when k is even,

$$\mathbb{P}(D_k = 0) = \binom{k}{k/2} p^{k/2} q^{k/2}.$$

2. To distinguish clear performance advantages from small fluctuations, we focus on events where one server completes at least two more jobs than the other. For any integer $k \geq 2$, define

$$P_k = \mathbb{P}(D_k \geq 2),$$

$$Q_k = \mathbb{P}(D_k \leq -2).$$

- (a) For a general integer $k \geq 2$, express P_k and Q_k as sums of probabilities corresponding to different possible numbers of jobs assigned to Server A. (You may find it useful to make use of the ceiling function $\lceil \cdot \rceil$, which rounds a number up to the smallest integer greater than or equal to it, and the floor function $\lfloor \cdot \rfloor$, which rounds a number down to the greatest integer less than or equal to it.)
- (b) Show that for any $k \geq 2$,

$$P_k > Q_k \quad \text{whenever} \quad p > q.$$

1. (a) Since exactly one job is assigned in each scheduling interval, after k intervals the total number of completed jobs satisfies

$$S_A(k) + S_B(k) = k.$$

Hence, $D_k \triangleq S_A(k) - S_B(k) = S_A(k) - (k - S_A(k)) = 2S_A(k) - k$.

- (b) For $k = 2$, that means there are two jobs: they can both be sent to A, or one to each, or both to B. Therefore, $D_2 \in \{-2, 0, 2\}$.

With $S_A(k) \sim \text{Bin}(k, p)$, the probabilities are:

$$\mathbb{P}(D_2 = 2) = \mathbb{P}(S_A(2) = 2) = p^2,$$

$$\mathbb{P}(D_2 = 0) = \mathbb{P}(S_A(2) = 1) = 2pq,$$

$$\mathbb{P}(D_2 = -2) = \mathbb{P}(S_A(2) = 0) = q^2.$$

- (c) When k is even, then $D_k = 0$ implies $S_A(k) = \frac{k}{2}$. Thus, with $S_A(k) \sim \text{Bin}(k, p)$,

$$\mathbb{P}(D_k = 0) = \mathbb{P}\left(S_A(k) = \frac{k}{2}\right) = \binom{k}{k/2} p^{k/2} q^{k/2}.$$

2. (a) Substituting the conclusion from 1(a), the event

$$D_k \triangleq S_A(k) - S_B(k) \geq 2$$

is equivalent to

$$\begin{aligned} 2S_A(k) - k &\geq 2, \\ S_A(k) &\geq \frac{k+2}{2}. \end{aligned}$$

Since $S_A(k) \sim \text{Bin}(k, p)$ and has to be integer-valued, we obtain

$$P_k = \mathbb{P}\left(S_A(k) \geq \frac{k+2}{2}\right) = \sum_{i=\lceil (k+2)/2 \rceil}^k \binom{k}{i} p^i (1-p)^{k-i},$$

where $\lceil \cdot \rceil$ is the ceiling function to round a number up to the smallest integer greater than or equal to it.

Similarly,

$$\begin{aligned} S_B(k) - S_A(k) &\geq 2, \\ k - 2S_A(k) &\geq 2, \\ S_A(k) &\leq \frac{k-2}{2}, \end{aligned}$$

and therefore

$$Q_k = \mathbb{P}\left(S_A(k) \leq \frac{k-2}{2}\right) = \sum_{i=0}^{\lfloor (k-2)/2 \rfloor} \binom{k}{i} p^i (1-p)^{k-i},$$

where $\lfloor \cdot \rfloor$ is the floor function to round a number down to the greatest integer less than or equal to it.

- (b) When $p > q$, $\Leftrightarrow p > 1-p$, $\Leftrightarrow p > \frac{1}{2}$.

For each i , the probability mass $\binom{k}{i} p^i (1-p)^{k-i}$ is increasing in i .

When $p > q$, the event defining P_k consists of larger values of i than the event defining Q_k . Therefore,

$$P_k > Q_k \quad \text{for all } k \geq 2 \text{ whenever } p > q.$$