

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

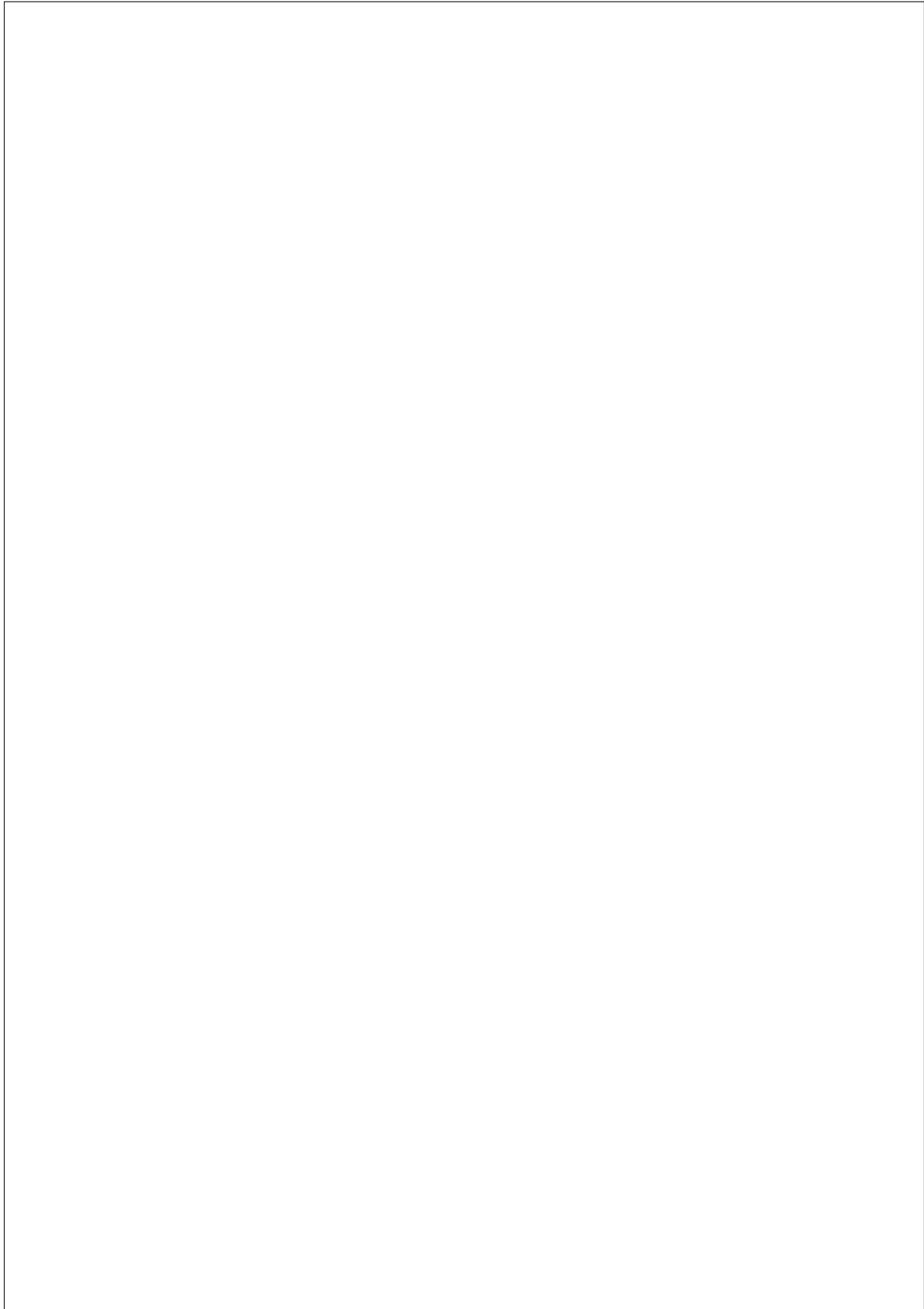
In class, we solved this problem using two different methods:

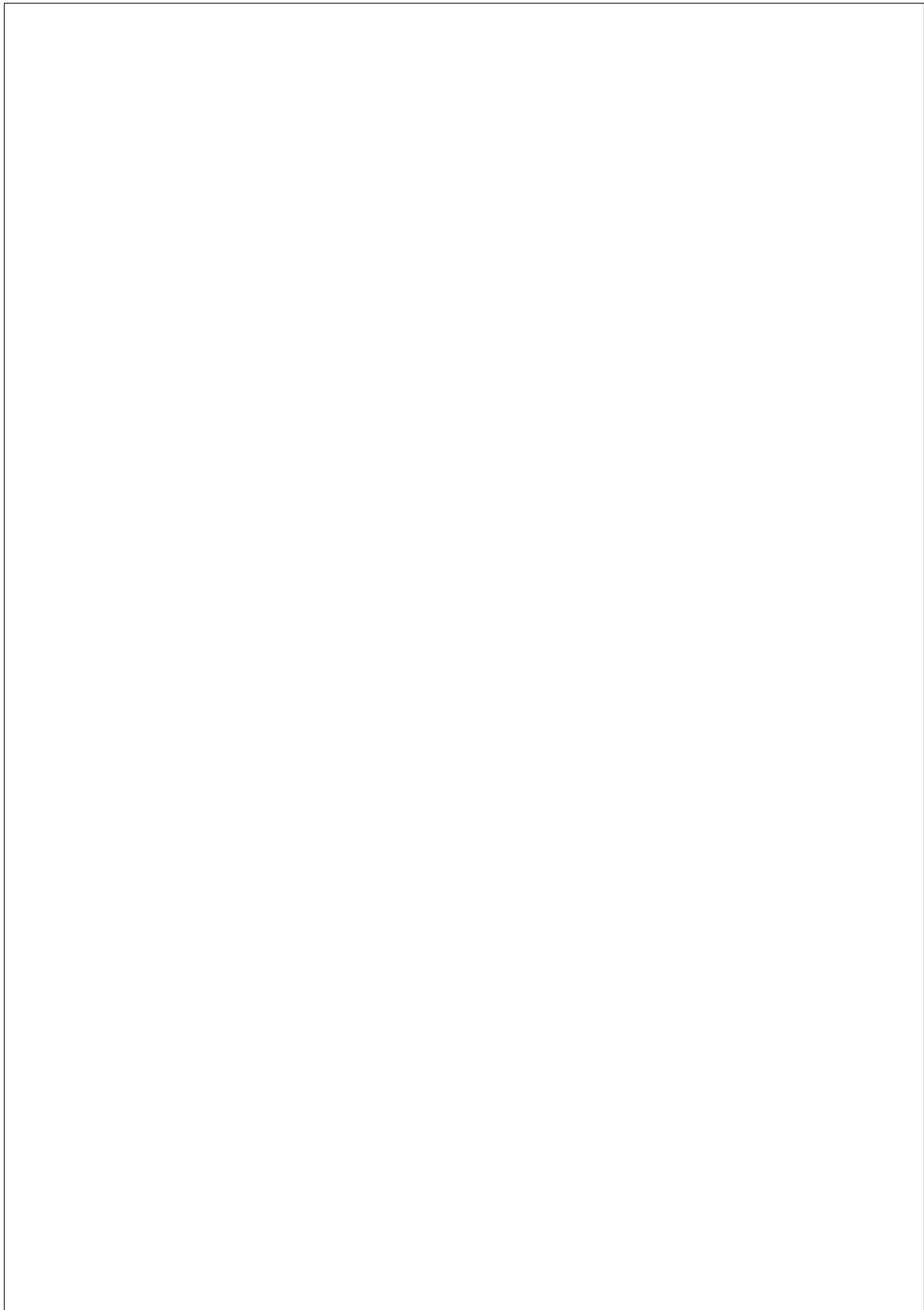
We divide a class of 16 students — 4 male, 12 female — into 4 groups of 4 students each. What is the probability that each group has exactly one male student?

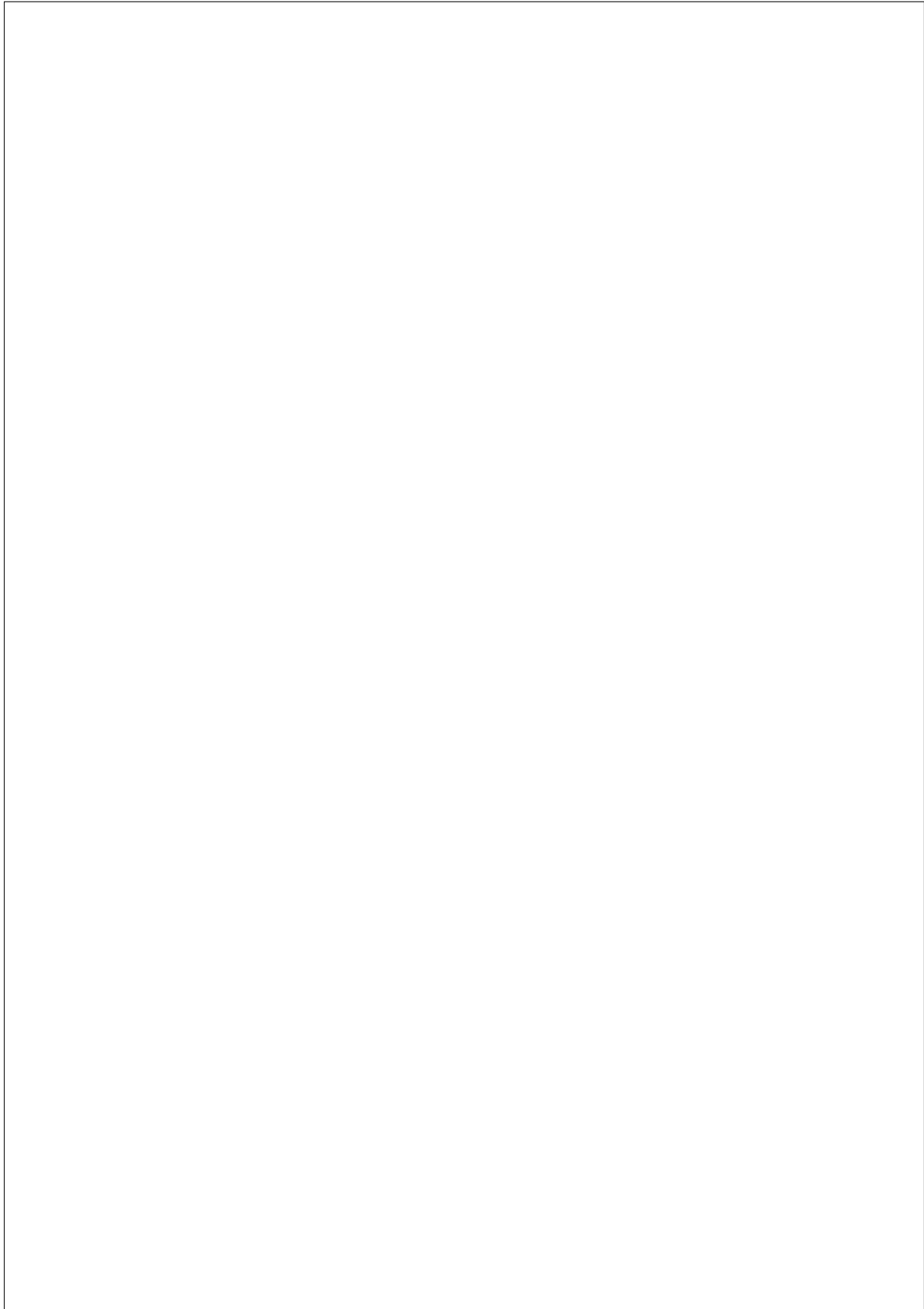
Here, we would like to try a third method: *computer simulation*.

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.)







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.

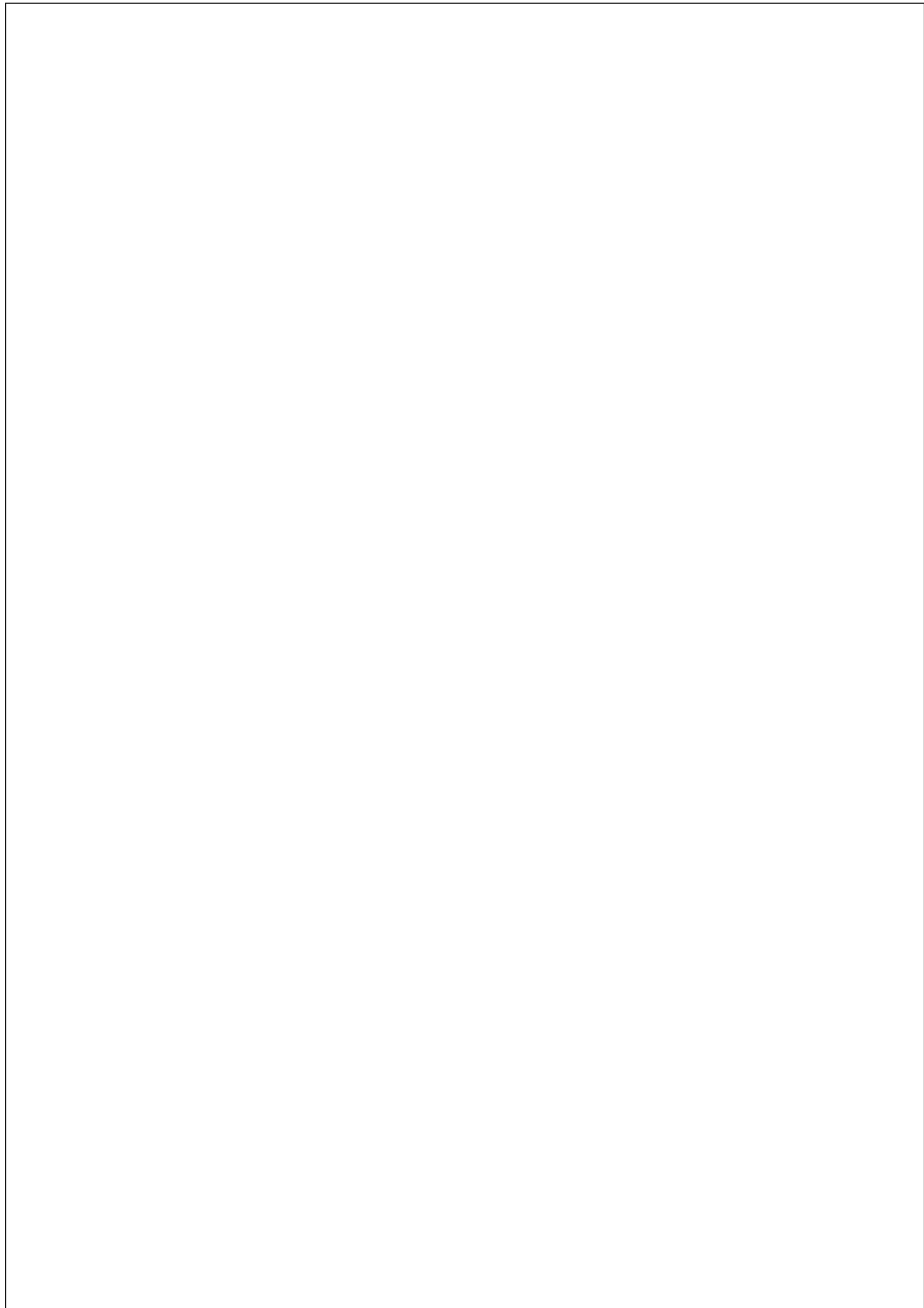
At each entrance, a mask-detection system is used, and it outputs either “Mask” or “No Mask”. Unfortunately, the detector is imperfect:

- 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 $\mathbb{P}(A | G)$.



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.

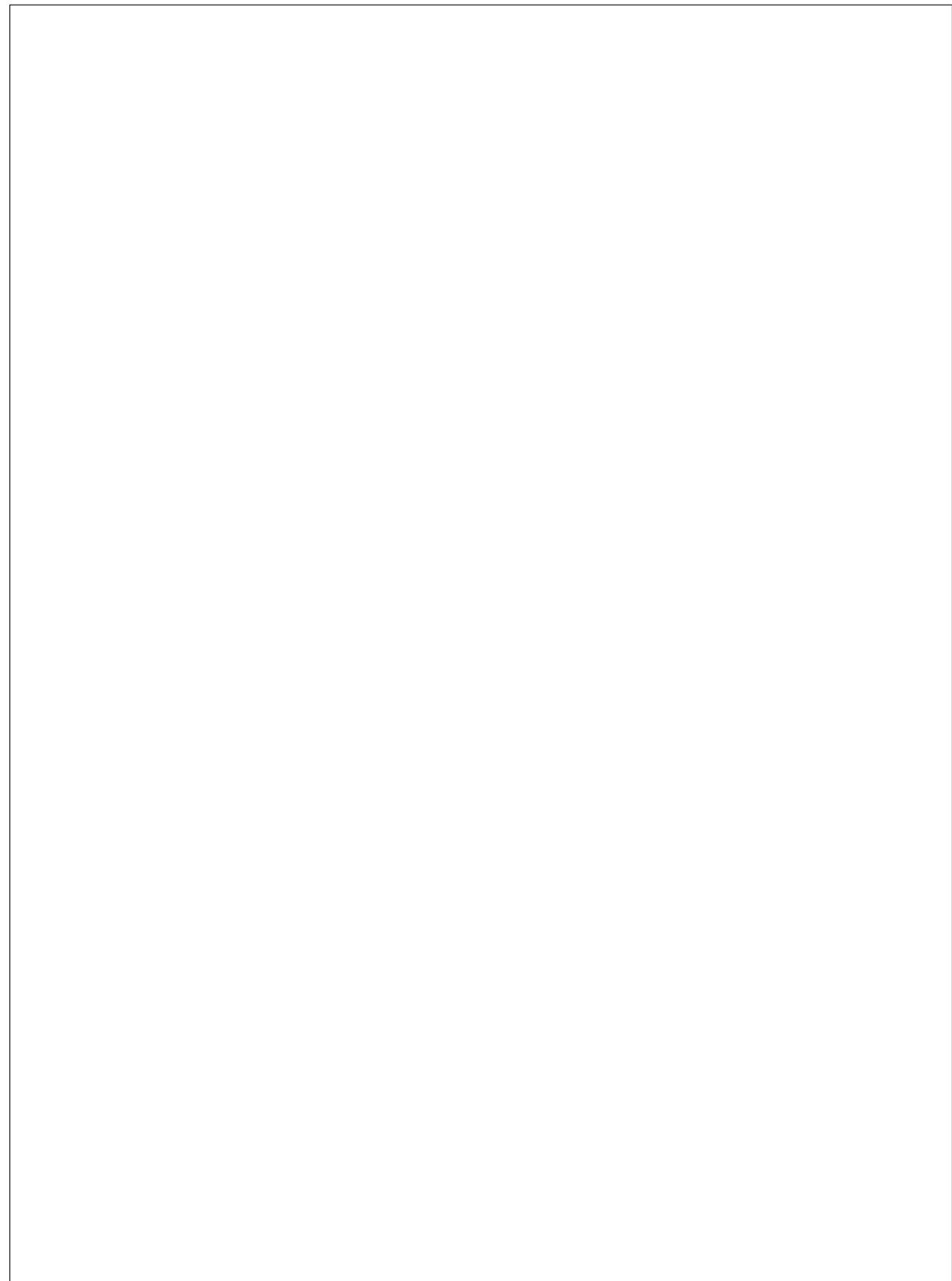
We randomly select two members of the audience from this group and denote them by A and B. Let X and Y represent the lengths of time (in hours) that A and B, respectively, stay in the shopping mall after the movie ends. To simplify the analysis, we assume that the stay durations are discretized and that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

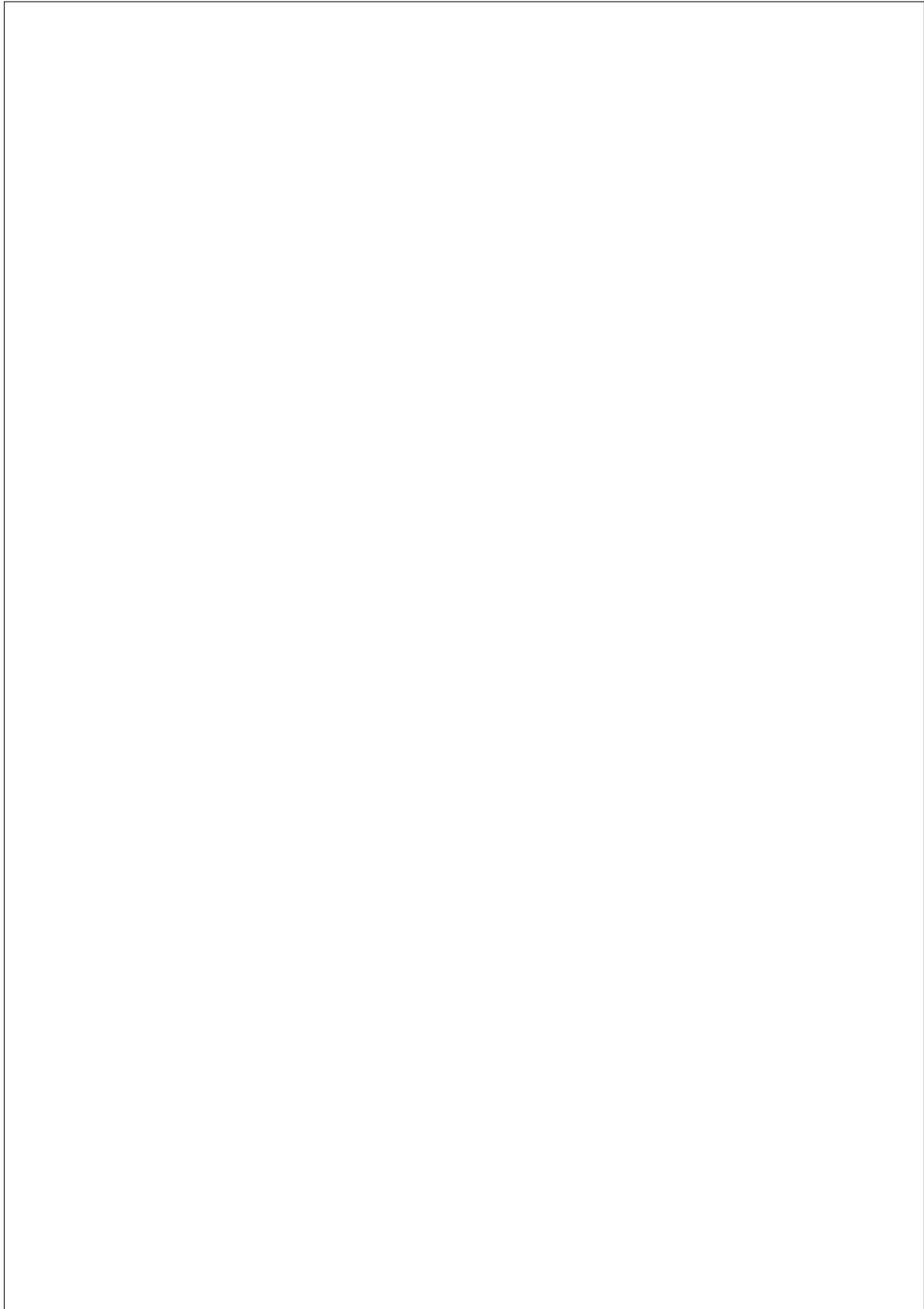
The joint PMF of (X, Y) is given by

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

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 $\text{var}(D)$.





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, $\mathbb{E}(X)$.

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 $k \geq 2$,

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

