



359













Teacher : Prof. Emmanuel Abbe
MATH-232 Probability and Statistics - Final Exam
20 June 2023
Duration : 180 minutes

Additional

SCIPER: 0

Do not turn the page before the start of the exam. This document is double-sided, has 24 pages, the last ones possibly blank. Do not unstaple.

- Place your student card on your table.
- **No other paper materials** are allowed to be used during the exam.
- Using a **calculator** or any electronic device is **not permitted** during the exam.
- A cheat sheet is provided on the last pages of this booklet.
- For the **multiple choice** questions, we give :
 - +3 points if your answer is correct,
 - 0 points if you give no answer or more than one,
 - 1 points if your answer is incorrect.
- For the **true/false** questions, we give :
 - +2 points if your answer is correct,
 - 0 points if you give no answer or more than one,
 - 2 points if your answer is incorrect.
- Use a **black or dark blue ballpen** and clearly erase with **correction fluid** if necessary.
- If a question is wrong, the teacher may decide to nullify it.
- The multiple choice questions are shuffled and hence are not in the order of difficulty.

Respectez les consignes suivantes Observe this guidelines Beachten Sie bitte die unten stehenden Richtlinien		
choisir une réponse select an answer Antwort auswählen	ne PAS choisir une réponse NOT select an answer NICHT Antwort auswählen	Corriger une réponse Correct an answer Antwort korrigieren
  		 
ce qu'il ne faut PAS faire what should NOT be done was man NICHT tun sollte		
     		

**First part: multiple choice questions**

For each question, mark the box corresponding to the correct answer. Each question has **exactly one** correct answer. No justifications are needed for this part.

Question 1 Suppose that random variable U follows the uniform distribution on $[0, 1]$, i.e., $U \sim \text{Uniform}(0, 1)$. Define $X = -\lambda \ln U$. What is $\mathbb{P}(X > x)$? ($\exp(z) = e^z$)

- $\frac{1}{\lambda} \exp(-x/\lambda)$
 $\exp(-x/\lambda)$
 $\lambda \exp(-\lambda x)$
 $\exp(-\lambda x)$

Question 2 Let X_1, X_2, \dots be a sequence of independent Poisson random variables such that $X_n \sim \text{Poisson}(n\lambda)$ where $\lambda > 0$ is a constant. Consider the sequence given by $Y_n = \frac{X_n}{n}$ and the following claims:

- (a) Y_n converges to λ in distribution.
(b) Y_n converges to λ in probability.
(c) Y_n converges to λ in mean square.

How many of the claims above are actually valid? In other words, how many modes of convergence (among in distribution, in probability and in mean square) hold for $Y_n \rightarrow \lambda$?

- 3
 1
 2
 0

Question 3 Consider three bits $b_1, b_2, b_3 \in \{0, 1\}$ that are sent over a noisy channel that flips each bit independently with probability $p < \frac{1}{2}$. Assume that one transmits two possible sequences on this channel with equal probability: either 000 or 111. Therefore, we define the null and alternative hypotheses as H_0 for '000 is transmitted' and H_1 for '111 is transmitted'. What is the optimal average error probability of a hypothesis test (again, each hypothesis is selected with probability 1/2)?

- $p^2(3 - 2p)$
 $p^2(1 - p)$
 p^3
 p^2

Question 4 Consider a random variable θ taken uniformly over $[0, 2\pi]$, i.e.,

$$f_\theta(\theta) = \begin{cases} \frac{1}{2\pi} & 0 \leq \theta \leq 2\pi \\ 0 & \text{otherwise} \end{cases}.$$

Note that this is equivalent with sampling a point from the unit circle randomly where θ is the angle. We define (X, Y) as the coordinates of the random point from the circle, i.e., $X = \cos \theta$ and $Y = \sin \theta$. What is $\mathbb{E}[X^2]$?

(Hint: This question is only about X but one may still use Y to come up with the answer.)

- $\frac{1}{2}$
 $\frac{3}{4}$
 $\frac{2}{3}$
 $\frac{1}{3}$



Question 5 Suppose that an individual has Covid with probability 0.01. Further, consider a Covid test that has 3% false positive and 0% false negative probability, where under H_0 'the person does not have Covid' and under H_1 'the person has Covid'. If someone tests positive, what is the probability that the person does have Covid?

- $\frac{1}{3.97}$
 $\frac{0.99}{4.96}$
 $\frac{1}{4.96}$
 $\frac{0.97}{0.99}$

Question 6 How many different ways are there to split 8 students into 4 groups of size 2?

- 2520
 105
 1680
 840

Question 7 Let $X_1, X_3, X_5, X_7, \dots$ be a sequence of i.i.d. Poisson(2) random variables and define $X_{2i} = X_{2i-1}$ for $i \in \mathbb{N}$. What does $\frac{1}{n}(X_1^2 + X_2^2 + \dots + X_n^2)$ converge to (in probability)?

- 2
 8
 4
 6

**Second part: true/false questions**

For each question, mark the box TRUE if the statement is **always true** and the box FALSE if it is **not always true** (i.e., it is sometimes false).

Question 8 Consider random variables $X, Y \sim \mathcal{N}_2((0 \ 0)^T, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix})$ having a multinomial Gaussian distribution. For a matrix $A \in \mathbb{R}^{2 \times 2}$ consider the linear transformation of $\begin{pmatrix} X' \\ Y' \end{pmatrix} = A \begin{pmatrix} X \\ Y \end{pmatrix}$.

Claim: There is no other matrix other than $A = I$ (identity matrix) such that X', Y' are independent and each of them are standard normal variables, i.e., $X', Y' \sim \mathcal{N}(0, 1)$.

TRUE FALSE

Question 9 A fair dice is thrown twice independently. Let D_1, D_2 be the values of the dice in the first and second throw.

Claim: the events $E_1 : D_1 = 2$ and $E_2 : D_1 + D_2 = 7$ are independent.

TRUE FALSE

Final 2023 Solutions

Prob Stats Team

July 19, 2023

P10

Solution: There are overall $N!$ distinct orderings for N questions. The probability that one student has a specific ordering is $\frac{1}{N!}$ because the order is picked uniformly at random. Fix the order of the student. Then, the left neighbor has the same order as her/him is $\mathbb{P}[\text{the left is the same}] = \frac{1}{N!}$, and the right neighbor has the same order as her/him is $\mathbb{P}[\text{the right is the same}] = \frac{1}{N!}$. Since the questions are shuffled independently for each student, the probability that at least one of her/his two neighbors has the same ordering is

$$1 - (1 - \mathbb{P}[\text{the left is the same}])(1 - \mathbb{P}[\text{the right is the same}]) \quad (1)$$

$$= 1 - \left(1 - \frac{1}{N!}\right) \left(1 - \frac{1}{N!}\right) = \frac{2}{N!} - \frac{1}{(N!)^2} \quad (2)$$

P11

Correlation

Solution: We can know $\text{cov}(X, Y) = 4$, $\text{var}(X) = \text{var}(Y) = 5$. Therefore, we have

$$\text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\{\text{var}(X)\text{var}(Y)\}^{1/2}} = 4/5 = 0.8 \quad (3)$$

Standard normal distribution

Solution: Suppose the mean vector is μ and the the covariance matrix is Ω , we can infer

$$\begin{pmatrix} X' \\ Y' \end{pmatrix} = A \begin{pmatrix} X \\ Y \end{pmatrix} + b \sim \mathcal{N}_2(b + A\mu, A\Omega A^T) \quad (4)$$

Therefore, if we want X' , Y' are independent and each standard normal distributed, we should have

$$b + A\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, A\Omega A^T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (5)$$

The answers satisfying the equations above are not unique. Here we present two methods:

Method 1: Simply note that

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}^2 = MM^T, \quad (6)$$

where $M = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. So it suffices to take

$$A = M^{-1} = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, \quad (7)$$

to have $A\Omega A^T = AMM^T A^T = I$. Finally, we set

$$b = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - A\mu = \begin{pmatrix} -2/3 \\ 1/3 \end{pmatrix}. \quad (8)$$

Method 2: Since the covariance matrix Ω is symmetric and positive semi-definite, the spectral theorem tells us that we may write $\Omega = CDC^T$, where $D = \text{diag}(d_1, d_2)$ contains the eigenvalues of Ω , with $d_1 \geq d_2$. C is an orthogonal matrix, whose columns C_1 and C_2 are the eigenvectors corresponding to the respective eigenvalues.

We can thus compute the eigenvalues of Ω are $d_1 = 9$ and $d_2 = 1$, with the corresponding eigenvectors $C_1 = [1/\sqrt{2}, 1/\sqrt{2}]$ and $C_2 = [-1/\sqrt{2}, 1/\sqrt{2}]$. Ω can be rewritten as:

$$\Omega = CDC^T = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 9 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \quad (9)$$

D can be further rewritten as:

$$D = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \quad (10)$$

Since C is orthogonal, $CC^T = C^T C = I$. Thus if

$$A = \begin{pmatrix} 1/3 & 0 \\ 0 & 1 \end{pmatrix} C^T = \begin{pmatrix} \sqrt{2}/6 & \sqrt{2}/6 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix} \quad (11)$$

we can make $A\Omega A^T = I$. Accordingly, we can compute

$$b = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - A\mu = \begin{pmatrix} -\sqrt{2}/6 \\ \sqrt{2}/2 \end{pmatrix} \quad (12)$$

P13

Solution: Let $\lambda > 0$. The goal is to show that the sequence of cumulative distribution functions corresponding to the sequence of random variables $\{Y_n\}_{n \geq 1}$ converges pointwise to the cumulative distribution function of the exponential distribution $\text{Exponential}(\lambda)$. It is also correct to show the convergence of the moment generating functions.

Let $n \in \mathbb{N}$, $n \geq 1$ and $y \geq 0$. Recall that X_n is drawn from a geometric distribution with success probability $\frac{\lambda}{n}$. Thus, we have

$$\mathbb{P}[Y_n \leq y] = \mathbb{P}[X_n \leq n \cdot y] = 1 - \mathbb{P}[X_n > n \cdot y] = 1 - \mathbb{P}[X_n > \lfloor n \cdot y \rfloor] = 1 - \left(1 - \frac{\lambda}{n}\right)^{\lfloor n \cdot y \rfloor}.$$

By noting that $\lim_{n \rightarrow +\infty} \left(1 - \frac{\lambda}{n}\right)^{\lfloor n \cdot y \rfloor} = e^{-\lambda y}$ (which can be justified by the squeeze theorem after bounding the floor operator), we deduce that

$$\lim_{n \rightarrow +\infty} \mathbb{P}[Y_n \leq y] = 1 - e^{-\lambda y}.$$

Observing that the right-hand side of the equation above is the cumulative distribution function of $\text{Exponential}(\lambda)$ at y concludes the proof.

P14

part a

solution

suppose X_i is the random variable indicating if a person voted in favor of the proposal. Then the problem is asking for

$$\mathbb{P}\left[\sum_{i=1}^{100} X_i \geq 70\right]$$

To use CLT, we should remove the mean and divide by the square root of variance to get a standard normal random variable Z . The mean of the $\sum_{i=1}^{100} X_i$ is 50 and the variance is $100\sigma_{X_i} = 25$

$$\mathbb{P}\left[\sum_{i=1}^{100} X_i \geq 70\right] = \mathbb{P}\left[\sum_{i=1}^{100} X_i - 50 \geq 20\right] = \mathbb{P}\left[\frac{\sum_{i=1}^{100} X_i - 50}{5} \geq 4\right] \approx 1 - \Phi(4) \approx 3 \times 10^{-5}$$

part b

solution

In $\mathbb{P}\left[\frac{1}{n} \sum_{i=1}^n X_i \geq p + \epsilon\right]$ we can remove the mean and use Chebyshev or Markov inequality.

$$\mathbb{P}\left[\frac{1}{n} \sum_{i=1}^n X_i \geq p + \epsilon\right] = \mathbb{P}\left[\frac{1}{n} \sum_{i=1}^n X_i - p \geq \epsilon\right] \leq \mathbb{P}\left[\left(\frac{1}{n} \sum_{i=1}^n X_i - p\right)^2 \geq \epsilon^2\right] \leq \frac{\mathbb{E}\left[\left(\frac{1}{n} \sum_{i=1}^n X_i - p\right)^2\right]}{\epsilon^2}$$

Note that $\mathbb{E}\left[\left(\frac{1}{n} \sum_{i=1}^n X_i - p\right)^2\right]$ is just the variance of the sample average, which is $\frac{1}{n}\sigma_{X_i} = \frac{1}{n}p(1-p)$ and for $p \in [0, 1]$, $p(1-p) \leq 0.25$. substituting this back to the right-hand side, we prove the result.

part c

solution

The desiderata can be translated to $\mathbb{P}\left[\left|\frac{1}{n} \sum_{i=1}^n X_i - p\right| \leq 0.01\right] \leq 0.95$. Hence,

$$\mathbb{P}\left[\left|\frac{1}{n} \sum_{i=1}^n X_i - p\right| \leq 0.01\right] = \mathbb{P}\left[\frac{\left|\frac{1}{n} \sum_{i=1}^n X_i - p\right|}{\sqrt{\frac{1}{n}p(1-p)}} \leq \frac{0.01}{\sqrt{\frac{1}{n}p(1-p)}}\right] = \mathbb{P}\left[\frac{\left|\frac{1}{n} \sum_{i=1}^n X_i - p\right|}{\sqrt{\frac{1}{n}p(1-p)}} \leq \frac{0.01\sqrt{n}}{0.5}\right]$$

where $\frac{1}{n}p(1-p)$ is the variance of the mean, and we have used the inequality $p(1-p) \leq 0.25$ again. Using the CLT approximation like part a, we have:

$$\mathbb{P}\left[\left|\frac{1}{n} \sum_{i=1}^n X_i - p\right| \leq 0.01\right] \approx \mathbb{P}\left[|Z| \leq 0.02\sqrt{n}\right]$$

Now knowing that $\mathbb{P}[Z_{0.025} < Z < Z_{0.975}] = 0.95$, we demand $0.02\sqrt{n} \geq Z_{0.975} = 1.96$ meaning $n \geq 9604$.

P15

Method of moments

Note that in this case, we have two unknown parameters, α, β ; thus, we compare the first and second moments. For the first moment, we have

$$\bar{y} := \frac{y_1 + y_2 + \dots + y_n}{n} = \frac{\alpha + \beta}{2}, \quad (13)$$

and for the second moment, we have

$$\frac{(y_1 - \bar{y})^2 + (y_2 - \bar{y})^2 + \dots + (y_n - \bar{y})^2}{n} = \frac{(\beta - \alpha)^2}{12} \quad (14)$$

We can easily solve the system by observing

$$\frac{\beta - \alpha}{2} = \sqrt{\frac{3}{n}((y_1 - \bar{y})^2 + \dots + (y_n - \bar{y})^2)}, \quad (15)$$

and hence

$$\alpha = \bar{y} - \sqrt{\frac{3}{n}((y_1 - \bar{y})^2 + \dots + (y_n - \bar{y})^2)} \quad (16)$$

$$\beta = \bar{y} + \sqrt{\frac{3}{n}((y_1 - \bar{y})^2 + \dots + (y_n - \bar{y})^2)}. \quad (17)$$

Maximum likelihood

Note that

$$f(y; \alpha, \beta) = \begin{cases} \frac{1}{\beta - \alpha} & \alpha \leq y \leq \beta \\ 0 & \text{otherwise} \end{cases}. \quad (18)$$

Thus we have

$$f(y_1, y_2, \dots, y_n; \alpha, \beta) = \begin{cases} \frac{1}{(\beta - \alpha)^n} & \alpha \leq y_1, \dots, y_n \leq \beta \\ 0 & \text{otherwise} \end{cases}. \quad (19)$$

Therefore, in order to maximize the likelihood, we want to indeed minimize $\beta - \alpha$ given the condition $\alpha \leq y_1, \dots, y_n \leq \beta$. Thus the answer is $\beta = \max\{y_1, \dots, y_n\}$ and $\alpha = \min\{y_1, \dots, y_n\}$.



Third part, open questions

Answer in the empty space below. Your answer should be carefully justified, and all the steps of your argument should be discussed in details. Unjustified answers will not get the full point even if correct; while, partial reasoning may receive some points. Leave the check-boxes empty, they are used for the grading.

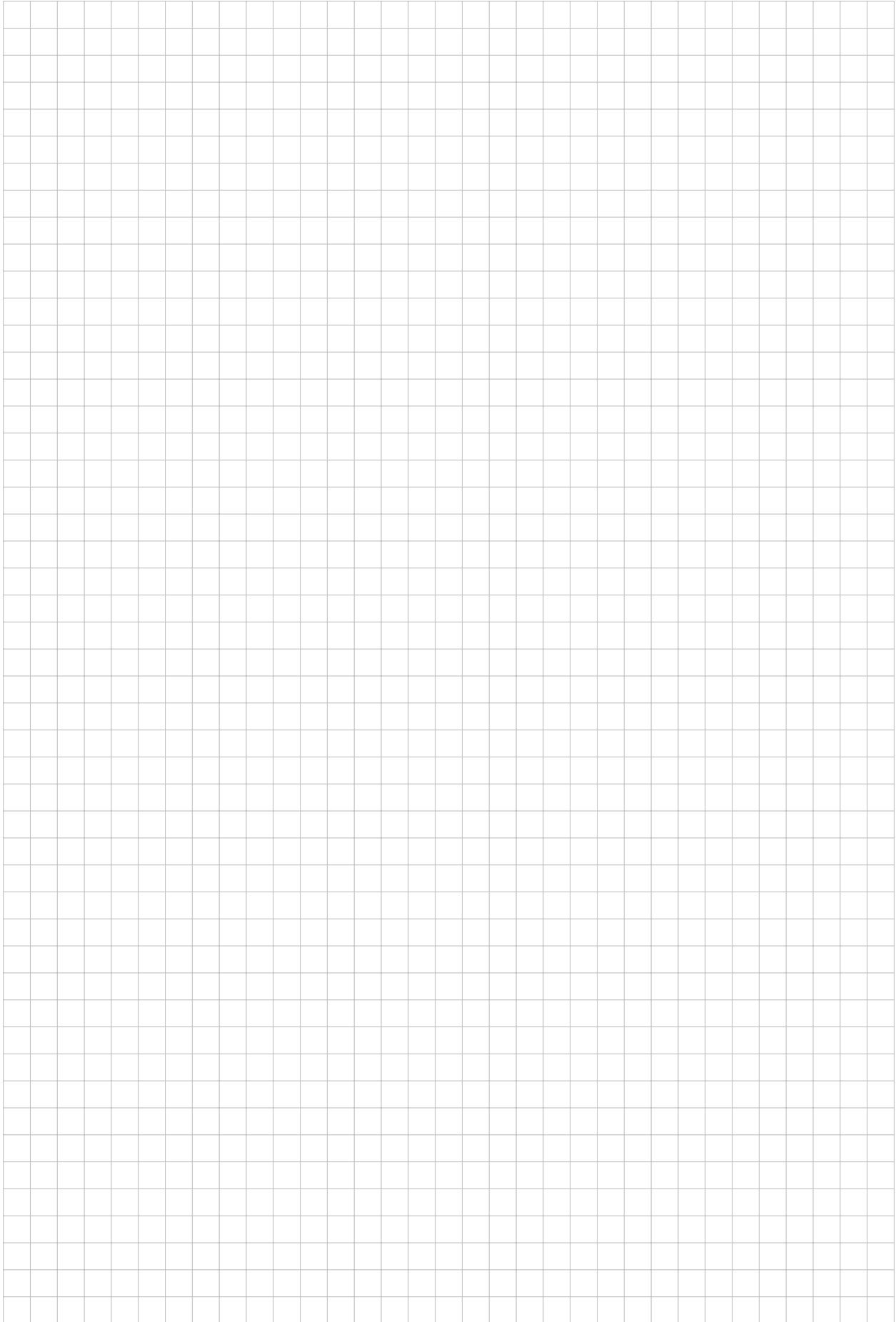
Question 10: *This question is worth 5 points.*

<input type="checkbox"/>	0	<input type="checkbox"/>	.5	<input type="checkbox"/>	1	<input type="checkbox"/>	.5	<input type="checkbox"/>	2	<input type="checkbox"/>	.5	<input type="checkbox"/>	3	<input type="checkbox"/>	.5	<input type="checkbox"/>	4	<input type="checkbox"/>	.5	<input checked="" type="checkbox"/>	5
--------------------------	---	--------------------------	----	--------------------------	---	--------------------------	----	--------------------------	---	--------------------------	----	--------------------------	---	--------------------------	----	--------------------------	---	--------------------------	----	-------------------------------------	---

An exam has N questions which are shuffled (i.e., the order of the questions is picked uniformly at random) for each student independently. Students are taking the exam next to each other in a room. Consider a student who has 2 neighbors. What is the probability that she/he has the same question ordering as at least one of her/his two neighbors?



+359/6/43+





Question 11: *This question is worth 7 points.*

<input type="checkbox"/>	0	<input type="checkbox"/>	.5	<input type="checkbox"/>	1	<input type="checkbox"/>	.5	<input type="checkbox"/>	2	<input type="checkbox"/>	.5	<input type="checkbox"/>	3	<input type="checkbox"/>	.5
<input type="checkbox"/>	4	<input type="checkbox"/>	.5	<input type="checkbox"/>	5	<input type="checkbox"/>	.5	<input type="checkbox"/>	6	<input type="checkbox"/>	.5	<input checked="" type="checkbox"/>	7		

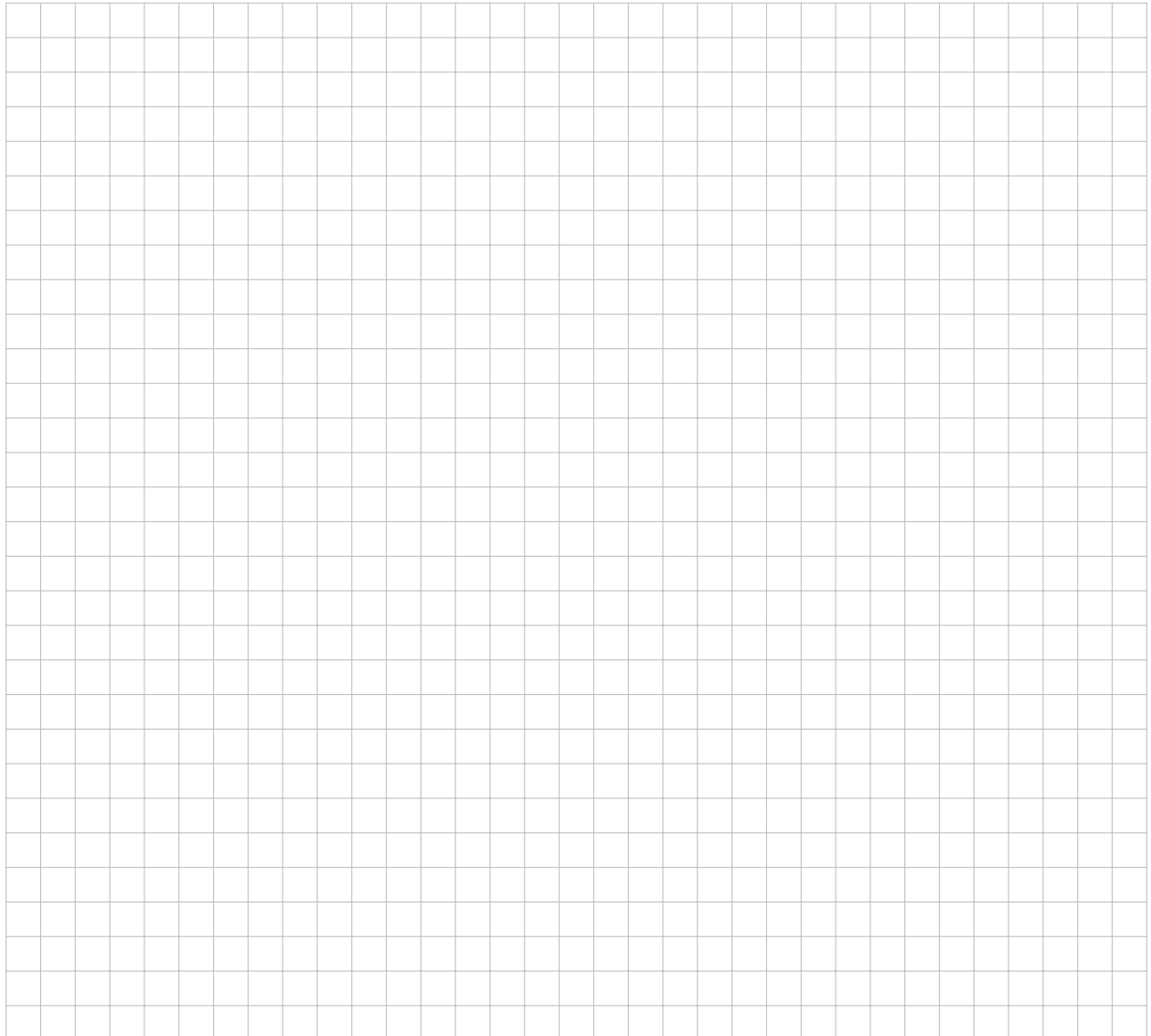
Consider random variables X, Y following the joint normal distribution

$$X, Y \sim \mathcal{N}_2\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}\right).$$

- (a) (2 points) What is the correlation between X, Y ?
- (b) (5 points) For a matrix $A \in \mathbb{R}^{2 \times 2}$ and a vector $b \in \mathbb{R}^2$, let

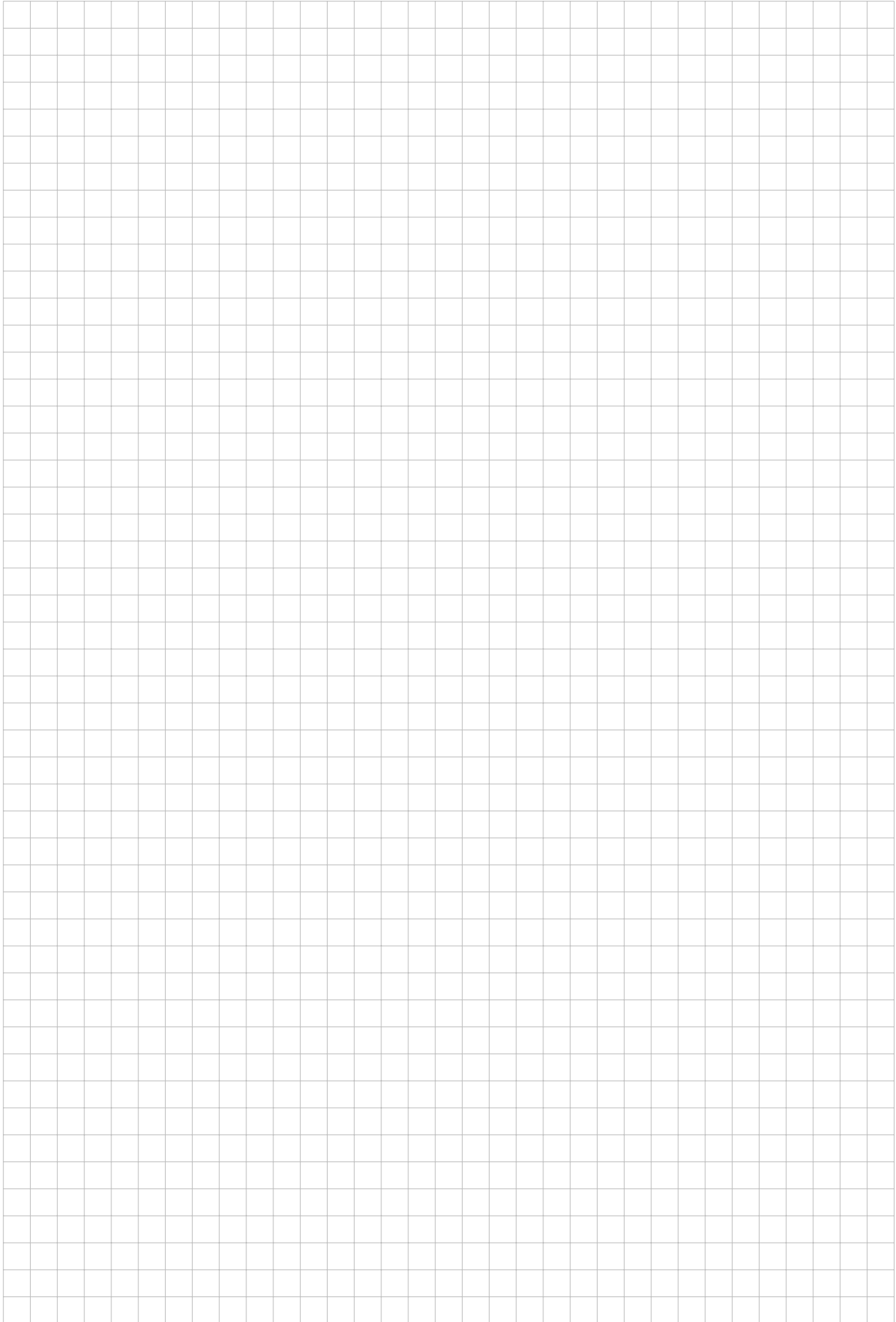
$$\begin{pmatrix} X' \\ Y' \end{pmatrix} = A \begin{pmatrix} X \\ Y \end{pmatrix} + b.$$

Find A, b such that X', Y' are independent and each standard normal distributed, i.e., $X' \sim \mathcal{N}(0, 1)$ and $Y' \sim \mathcal{N}(0, 1)$.



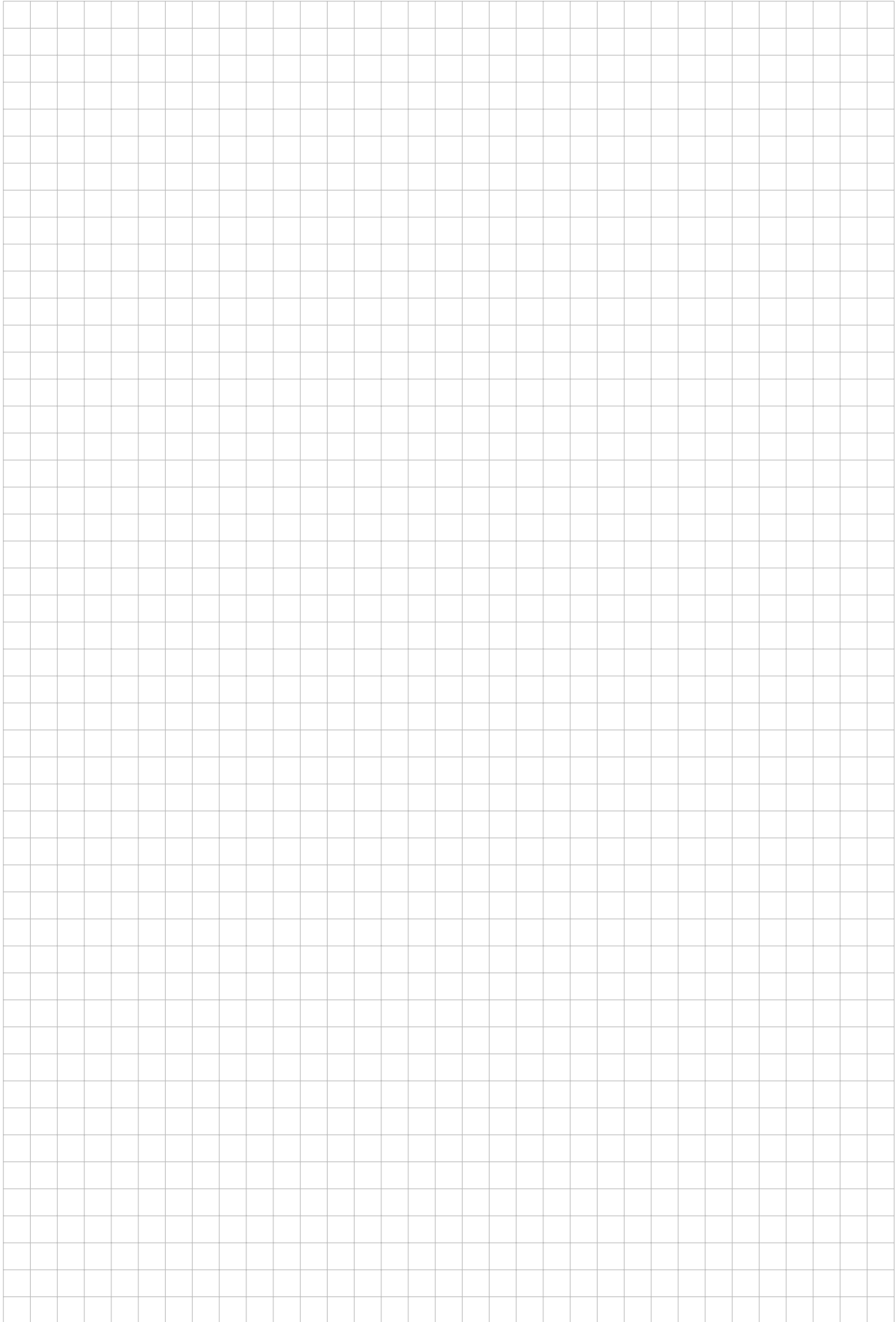


+359/8/41+





+359/9/40+





Question 13: *This question is worth 6 points.*

0 .5 1 .5 2 .5 3 .5 4 .5 5 .5 6

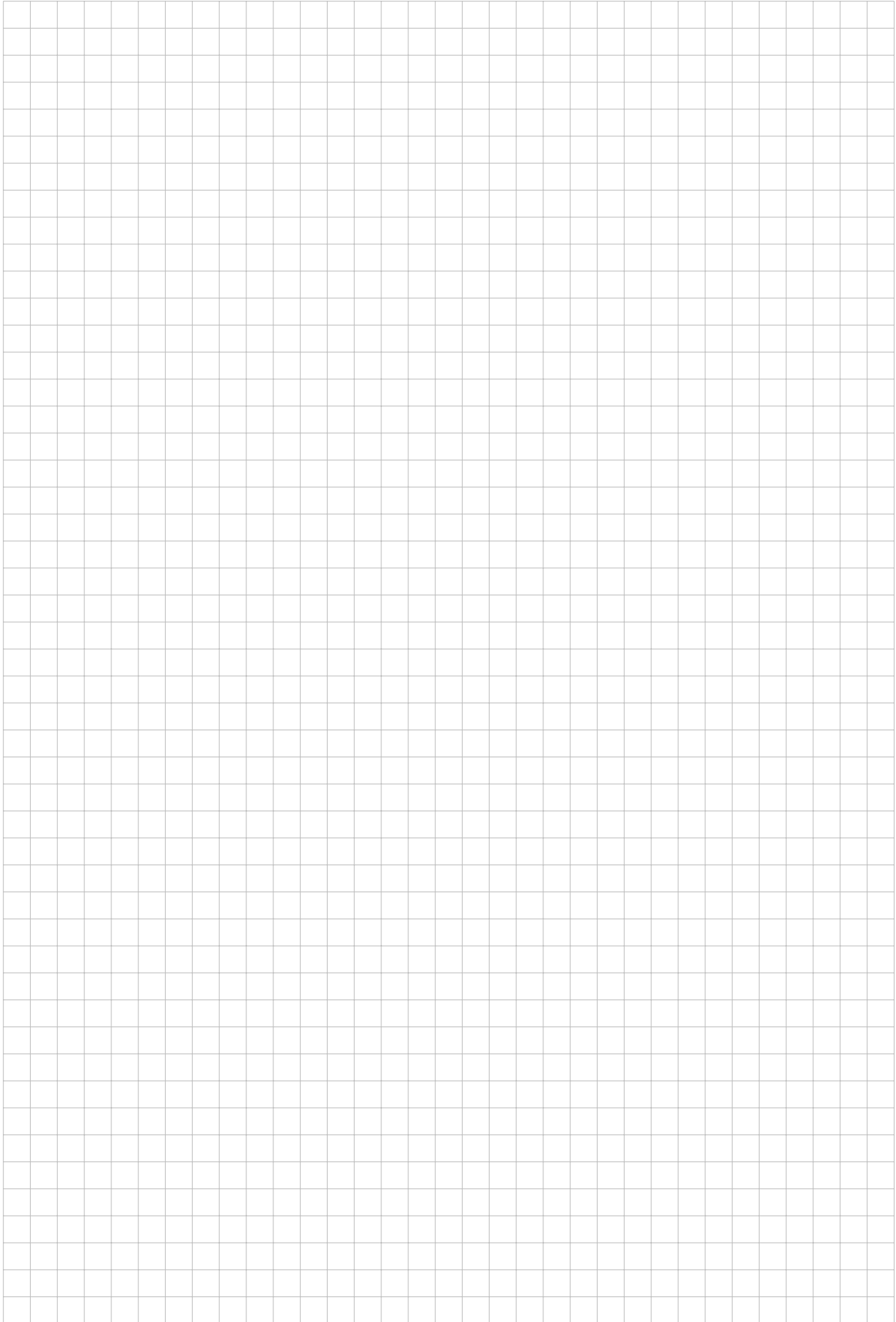
Let X_1, X_2, \dots be a sequence of independent geometric random variables such that $X_n \sim \text{Geometric}(\frac{\lambda}{n})$ where $\lambda > 0$ is a constant. Prove that the sequence given by $Y_n = \frac{X_n}{n}$ converges in distribution to $\text{Exponential}(\lambda)$.

You may want to use the following formula for this question: $\lim_{n \rightarrow \infty} (1 - \frac{\lambda}{n})^n = e^{-\lambda}$, although there may be various ways to solve the problem.





+359/11/38+



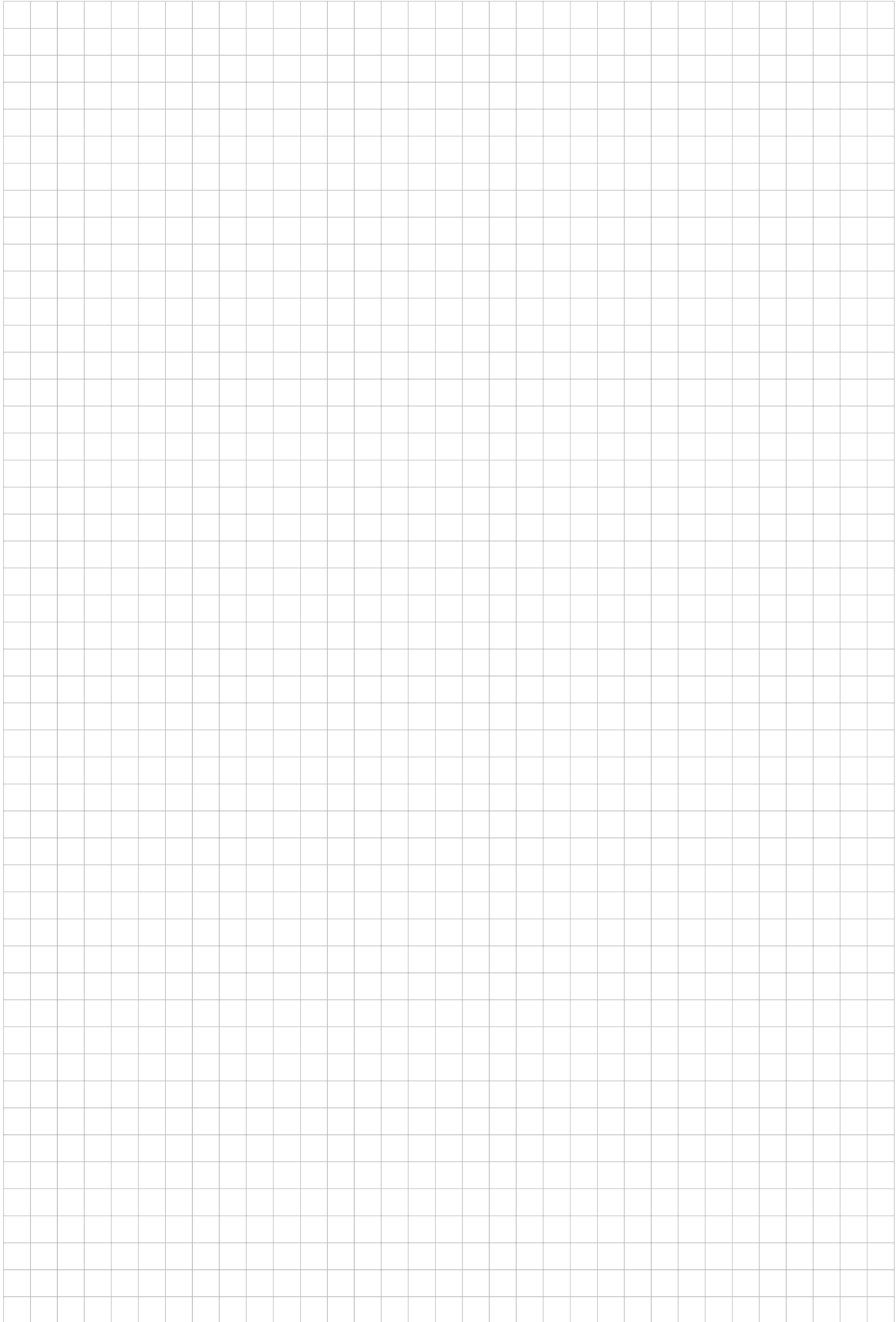


Question 14: *This question is worth 9 points.*

<input type="checkbox"/>	0	<input type="checkbox"/>	.5	<input type="checkbox"/>	1	<input type="checkbox"/>	.5	<input type="checkbox"/>	2	<input type="checkbox"/>	.5	<input type="checkbox"/>	3	<input type="checkbox"/>	.5	<input type="checkbox"/>	4	<input type="checkbox"/>	.5
<input type="checkbox"/>	5	<input type="checkbox"/>	.5	<input type="checkbox"/>	6	<input type="checkbox"/>	.5	<input type="checkbox"/>	7	<input type="checkbox"/>	.5	<input type="checkbox"/>	8	<input type="checkbox"/>	.5	<input checked="" type="checkbox"/>	9		

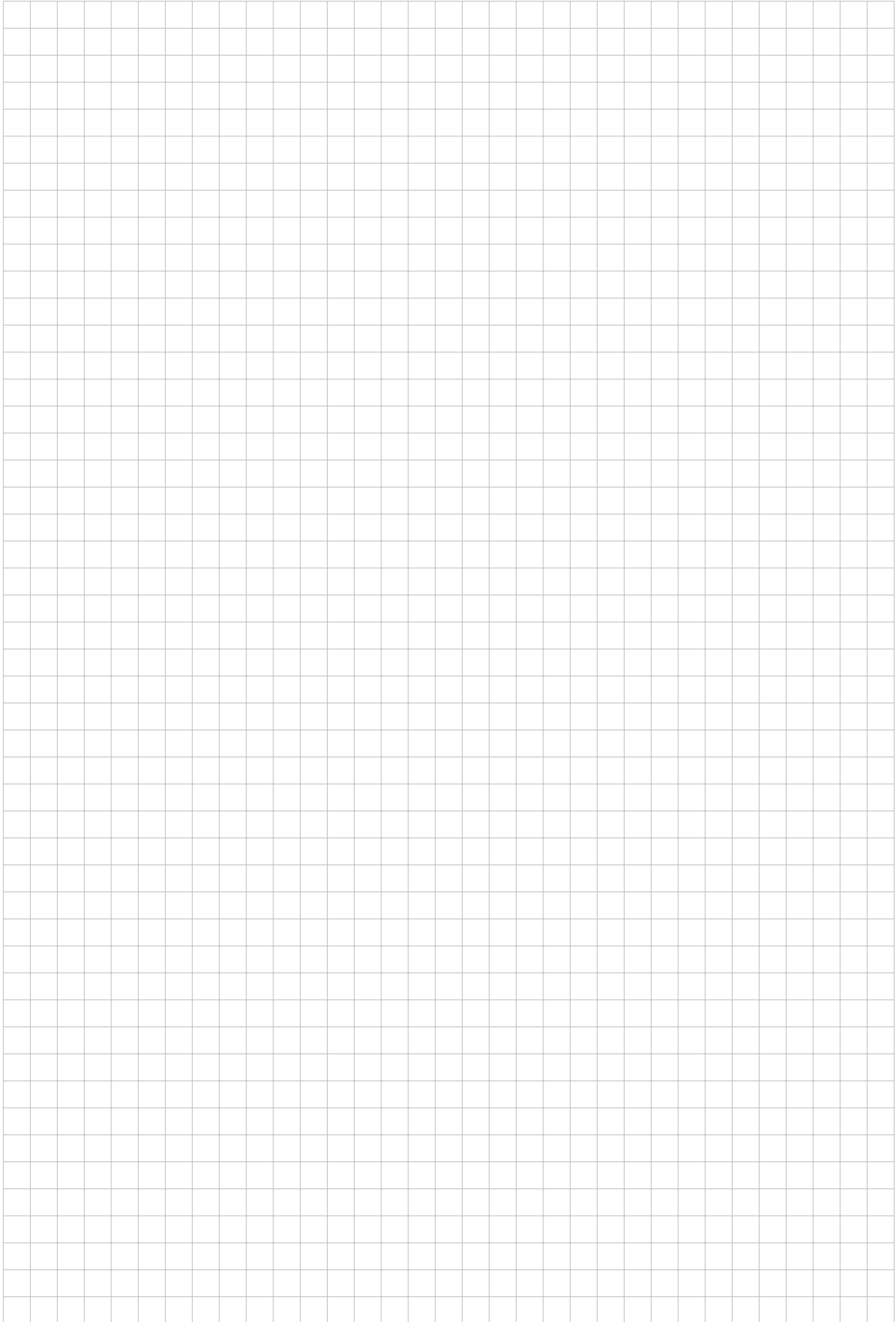
A referendum is going to take place in Switzerland on a specific proposal. Assume that each citizen is in favor of the proposal with probability p independently of other citizens. Answer the questions below. Note that these questions are independent of each other, so you do not need to solve them in order.

- (a) (3 points) Assume $p = 0.5$. Using the Central Limit Theorem, estimate the probability that in a group of 100 citizens, at least 70 support the referendum (you should use the table provided in the cheat sheet).
- (b) (3 points) For any p , consider a group of n citizens. Define \bar{X}_n as the fraction of the people who support the referendum in the group. Prove that $\mathbb{P}(\bar{X}_n \geq p + \epsilon) \leq \frac{\epsilon^2}{4n}$.
- (c) (3 points) Assume that a polling institute wants to estimate p . To do this, the institute takes n citizens and computes \bar{X}_n , the fraction of people in favor of the referendum within this group. What is the smallest group size n in order to have a 95% confidence that \bar{X}_n is within 0.01 of the true value of p ?





+359/14/35+



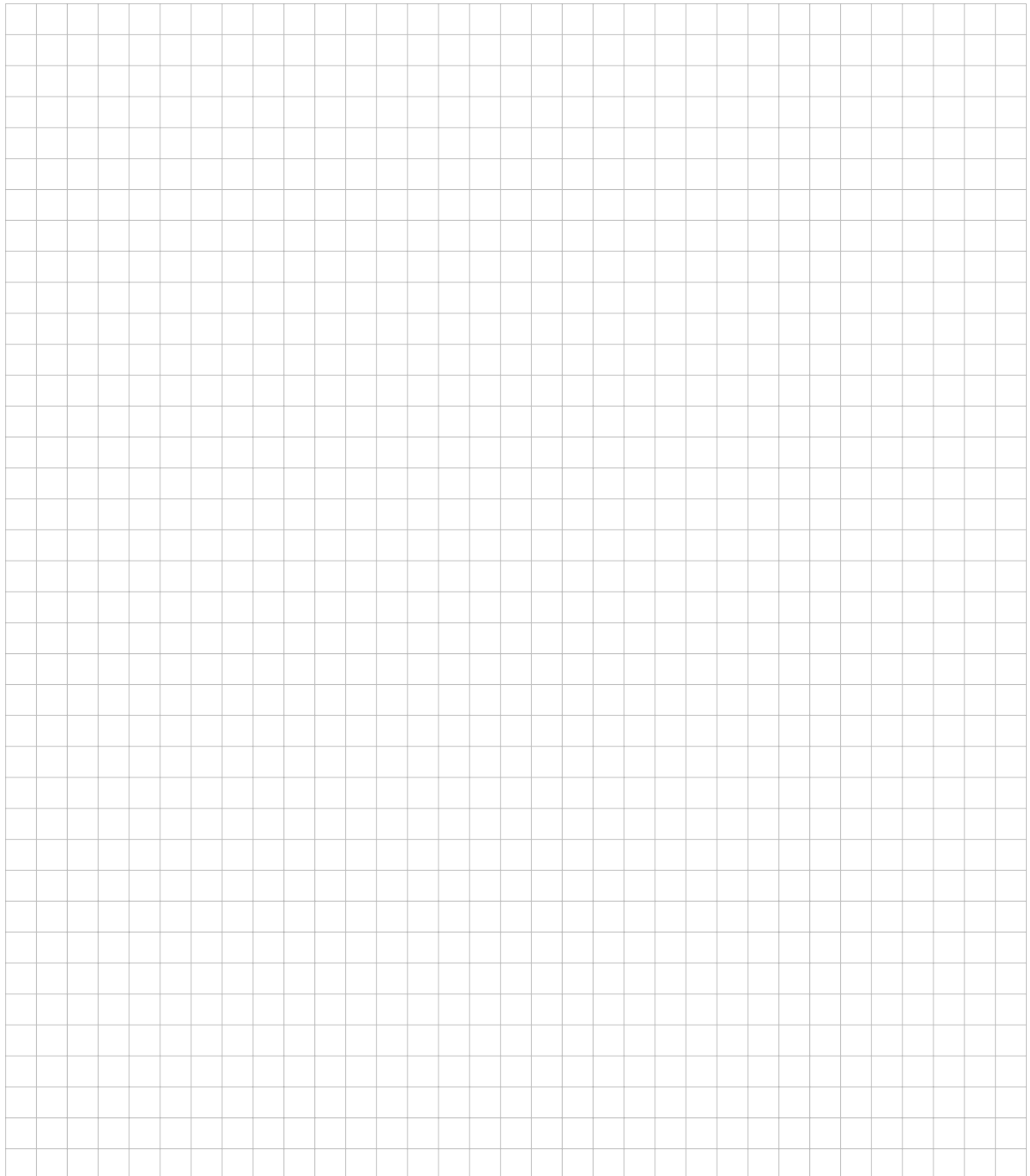


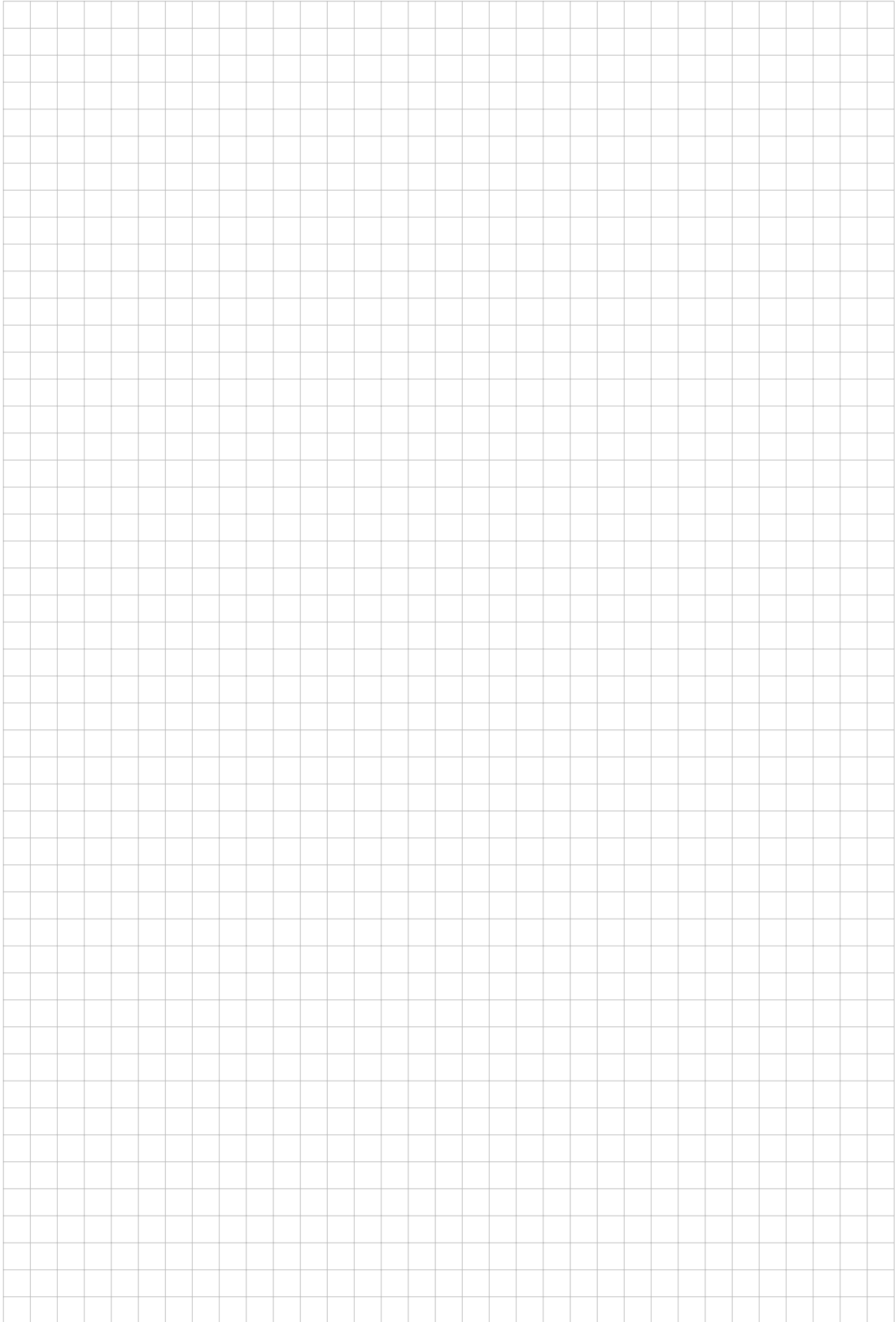
Question 15: *This question is worth 8 points.*

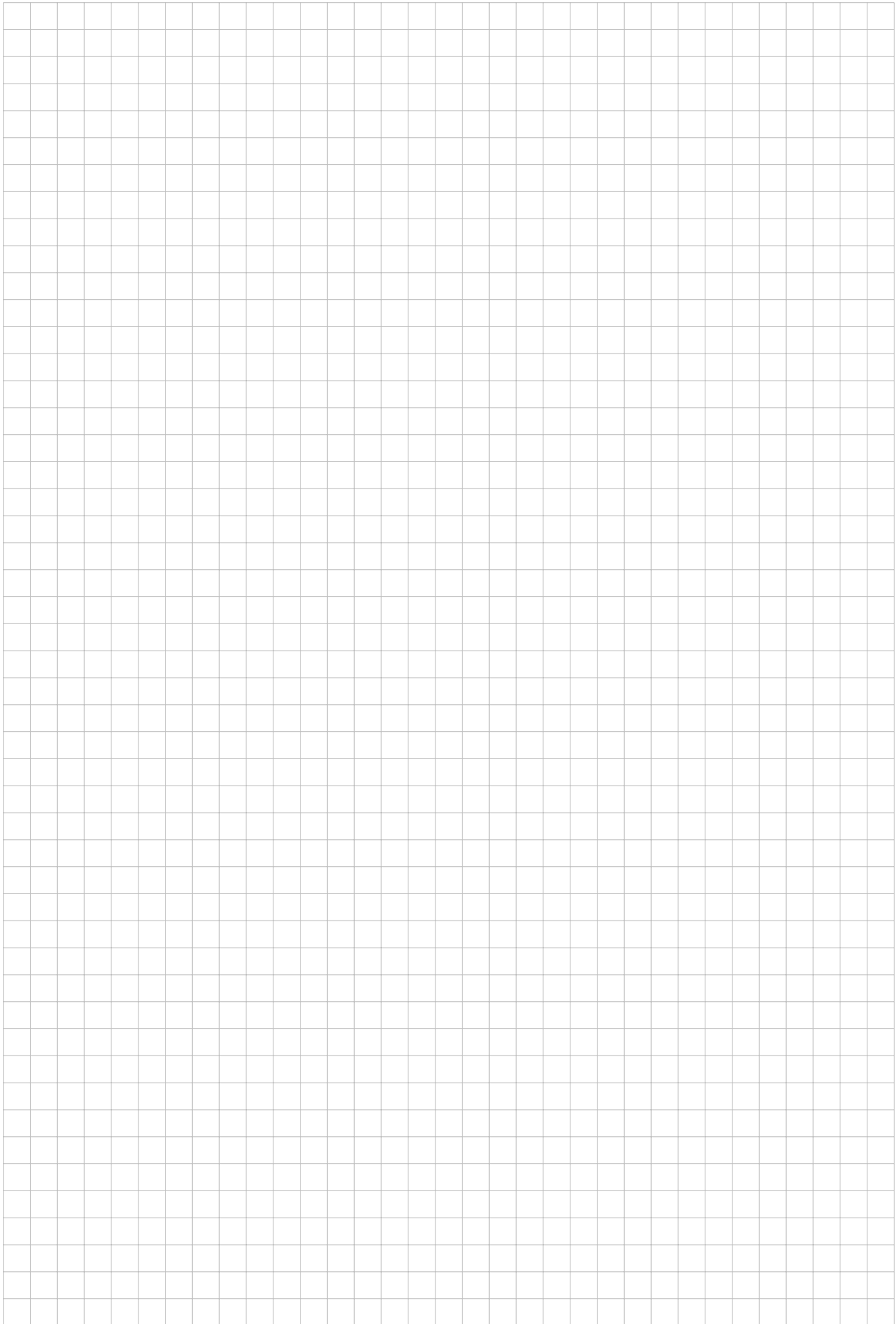
<input type="checkbox"/>	0	<input type="checkbox"/>	.5	<input type="checkbox"/>	1	<input type="checkbox"/>	.5	<input type="checkbox"/>	2	<input type="checkbox"/>	.5	<input type="checkbox"/>	3	<input type="checkbox"/>	.5	<input type="checkbox"/>	4
<input type="checkbox"/>	.5	<input type="checkbox"/>	5	<input type="checkbox"/>	.5	<input type="checkbox"/>	6	<input type="checkbox"/>	.5	<input type="checkbox"/>	7	<input type="checkbox"/>	.5	<input checked="" type="checkbox"/>	8		

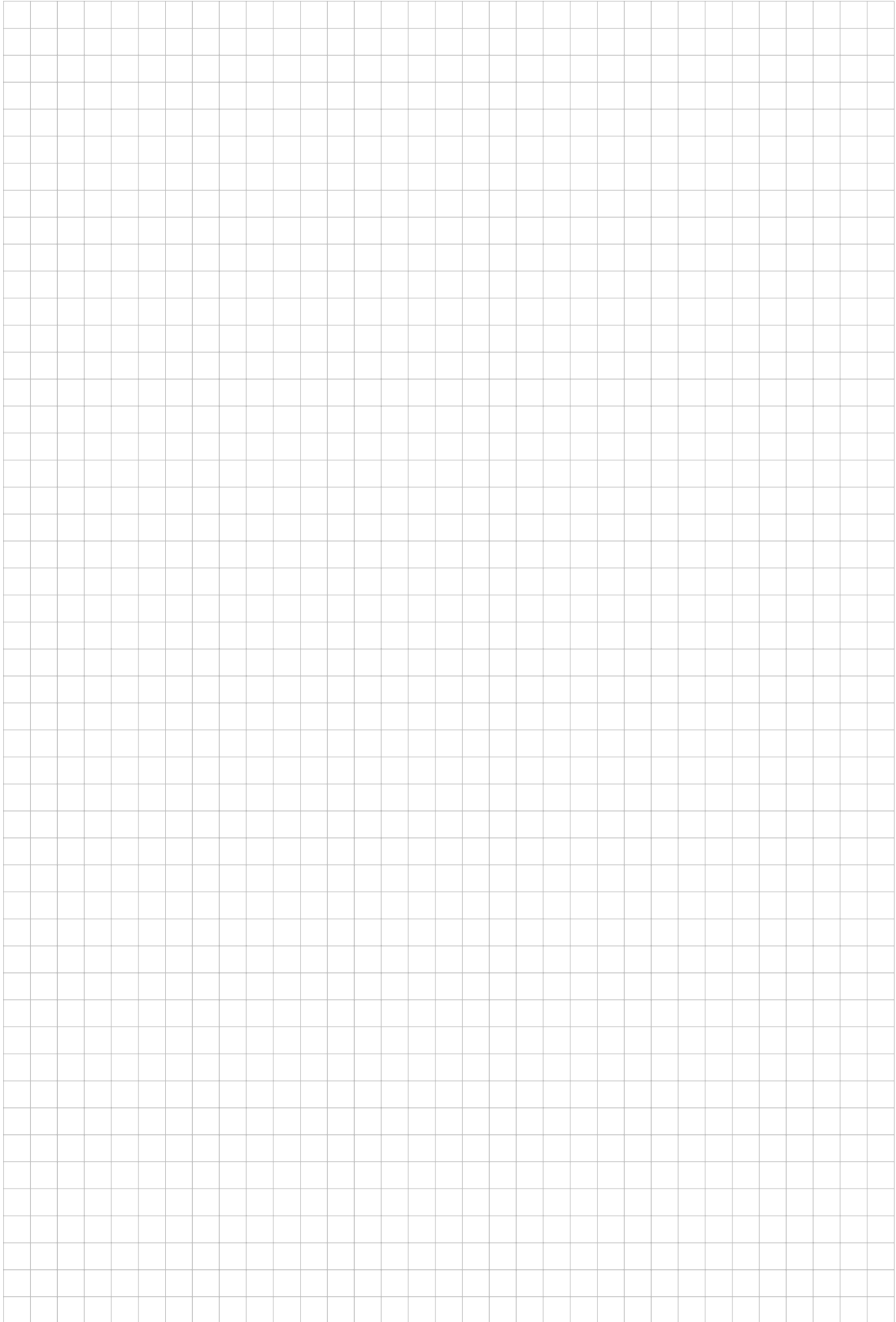
Let y_1, \dots, y_n be samples drawn from the uniform distribution $U(\alpha, \beta)$ where α and β are unknown real-valued parameters.

- (a) (4 points) Estimate α, β based on the method of moments.
- (b) (4 points) Estimate α, β using the maximum likelihood method.











Basic formulas and definitions

- Properties of binomial coefficients

(a) Pascal's triangle $\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}$.

(b) Vandermonde's formula $\sum_{j=0}^r \binom{m}{j} \binom{n}{r-j} = \binom{m+n}{r}$.

(c) Negative binomial series $(1-x)^{-n} = \sum_{i=0}^{\infty} \binom{n+i-1}{i} x^i, |x| < 1$.

(d) $\lim_{n \rightarrow \infty} n^{-r} \binom{n}{r} = \frac{1}{r!}$, where $r \in \mathbb{N}$ is fixed.

- Inclusion-exclusion formula:

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{r=1}^n (-1)^{r+1} \sum_{1 \leq i_1 < i_2 < \dots < i_r \leq n} \mathbb{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_r})$$

- Let $g(X, Y)$ be a function of a random vector (X, Y) . Its conditional expectation given $X = x$ is

$$\mathbb{E}[g(X, Y)|X = x] = \begin{cases} \sum_y g(x, y) f_{Y|X}(y|x) & \text{discrete case} \\ \int g(x, y) f_{Y|X}(y|x) dy & \text{continuous case} \end{cases}$$

on the condition that $f_X(x) > 0$ and $\mathbb{E}[|g(X, Y)||X = x|] < \infty$.

- For random variables X and $Y = g(X)$, where g is a monotone increasing or decreasing function with differentiable inverse g^{-1} , we have

$$f_Y(y) = \left| \frac{dg^{-1}(y)}{dy} \right| f_X(g^{-1}(y)).$$

- For a random variable X , the moment generating function is defined as $M_X(t) = \mathbb{E}[e^{tX}]$ for $t \in \mathbb{R}$ such that $M_X(t) < \infty$. Similarly, for a random vector $X_{p \times 1} = (X_1, X_2, \dots, X_p)^T$, we have $M_X(t) = \mathbb{E}[e^{t^T X}]$ for $t \in \mathbb{R}^p$ such that $M_X(t) < \infty$.
- The random vector $X \sim \mathcal{N}_p(\mu, \Omega)$ has a density function on \mathbb{R}^p if and only if Ω is positive definite, i.e., Ω has rank p . If so, the density function is

$$f(x; \mu, \Omega) = \frac{1}{(2\pi)^{p/2} |\Omega|^{1/2}} \exp\left(-\frac{1}{2}(x - \mu)^T \Omega^{-1} (x - \mu)\right).$$

If not, X is a linear combination of variables that have a density function on \mathbb{R}^m , where $m < p$ is the rank of Ω .

- Let $X \sim \mathcal{N}_p(\mu_{p \times 1}, \Omega_{p \times p})$, where $|\Omega| > 0$, and let $\mathcal{A}, \mathcal{B} \subset \{1, \dots, p\}$ with $|\mathcal{A}| = q < p, |\mathcal{B}| = r < p$ and $\mathcal{A} \cap \mathcal{B} = \emptyset$. Let $\mu_{\mathcal{A}}, \Omega_{\mathcal{A}}$ and $\Omega_{\mathcal{A}\mathcal{B}}$ be respectively the $q \times 1$ subvector of μ , $q \times q$ and $q \times r$ submatrices of Ω conformable with $\mathcal{A}, \mathcal{A} \times \mathcal{A}$ and $\mathcal{A} \times \mathcal{B}$. Then:

- the marginal distribution of $X_{\mathcal{A}}$ is normal, $X_{\mathcal{A}} \sim \mathcal{N}_q(\mu_{\mathcal{A}}, \Omega_{\mathcal{A}})$;
- the conditional distribution of $X_{\mathcal{A}}$ given $X_{\mathcal{B}} = x_{\mathcal{B}}$ is normal, $X_{\mathcal{A}}|X_{\mathcal{B}} = x_{\mathcal{B}} \sim \mathcal{N}_q(\mu_{\mathcal{A}} + \Omega_{\mathcal{A}\mathcal{B}} \Omega_{\mathcal{B}}^{-1} (x_{\mathcal{B}} - \mu_{\mathcal{B}}), \Omega_{\mathcal{A}} - \Omega_{\mathcal{A}\mathcal{B}} \Omega_{\mathcal{B}}^{-1} \Omega_{\mathcal{B}\mathcal{A}})$.

- Let $Y = g(X) \in \mathbb{R}^n$, where $X \in \mathbb{R}^n$ is a continuous variable and

$$(X_1, \dots, X_n) \rightarrow (Y_1 = g_1(X_1, \dots, X_n), \dots, Y_n = g_n(X_1, \dots, X_n)),$$

where g_i 's are continuously differentiable. If the inverse transformation $h_i = g_i^{-1}$ exist, and we have Jacobian $J(x_1, \dots, x_n) \in \mathbb{R}^{n \times n}$ such that $J_{ij} = \frac{\partial g_i}{\partial x_j}$ such that $|J(x_1, \dots, x_n)| > 0$ if $f_{X_1, \dots, X_n}(x_1, \dots, x_n) > 0$ then

$$f_{Y_1, \dots, Y_n}(y_1, \dots, y_n) = f_{X_1, \dots, X_n}(x_1, \dots, x_n) |J(x_1, \dots, x_n)|^{-1}$$

evaluated at $x_1 = h_1(y_1, \dots, y_n), \dots, x_n = h_n(y_1, \dots, y_n)$.



- Convergence of random variables: we consider the following definitions for convergence of random variables X_1, X_2, \dots
 - X_n converges to X in mean square, $X_n \xrightarrow{2} X$ if $\lim_{n \rightarrow \infty} \mathbb{E}[(X_n - X)^2] = 0$, where $\mathbb{E}[X^2], \mathbb{E}[X_n^2] < \infty$.
 - X_n converges to X in probability, $X_n \xrightarrow{P} X$ if for all $\varepsilon > 0$, $\lim_{n \rightarrow \infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0$.
 - X_n converges to X in distribution, $X_n \xrightarrow{D} X$ if $\lim_{n \rightarrow \infty} F_n(x) = F(x)$, at each point where $F(x)$ is continuous where F represents the cumulative distribution function.
- Continuity theorem. Let $\{X_n\}, X$ be random variables with cumulative distribution functions $\{F_n\}, F$, whose MGFs $M_n(t), M(t)$ exist for $0 \leq |t| < b$. If there exists a $0 < a < b$ such that $M_n(t) \rightarrow M(t)$ for $|t| \leq a$ when $n \rightarrow \infty$, then $X_n \xrightarrow{D} X$.
- Combination of convergent sequences including Slutsky's Lemma. Let x_0, y_0 be constants, $X, Y, \{X_n\}, \{Y_n\}$ random variables, and h a function continuous at x_0 . Then
 - $X_n \xrightarrow{D} x_0 \implies X_n \xrightarrow{P} x_0$,
 - $X_n \xrightarrow{P} x_0 \implies h(X_n) \xrightarrow{P} h(x_0)$,
 - $X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{P} y_0 \implies X_n + Y_n \xrightarrow{D} X + y_0, X_n Y_n \xrightarrow{D} X y_0$.
- Some inequalities:
 - Markov's inequality: $\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$ assuming that X only takes non-negative values and $a > 0$.
 - Chebyshev's inequality: $\mathbb{P}(|X - \mathbb{E}[X]| \geq a) \leq \frac{\text{var}(X)}{a^2}$
 - Jensen's inequality: $\mathbb{E}[g(X)] \geq g(\mathbb{E}[X])$, where g is a convex function
 - Hoeffding's inequality: Let Z_1, \dots, Z_n be independent random variables with $\mathbb{E}[Z_i] = 0$ and $a_i \leq Z_i \leq b_i$. For $\varepsilon > 0$ and any $t > 0$, we have $\mathbb{P}(\sum_{i=1}^n Z_i \geq \varepsilon) \leq e^{-t\varepsilon} \prod_{i=1}^n e^{t^2(b_i - a_i)^2/8}$. Particularly, for i.i.d. $X_1, \dots, X_n \sim \text{Bernoulli}(p)$ and $\varepsilon > 0$, we have $\mathbb{P}(|\bar{X} - p| \geq \varepsilon) \leq 2e^{-2n\varepsilon^2}$, where $\bar{X} = (X_1 + \dots + X_n)/n$.
 - Cauchy-Schwarz inequality: For random variables X, Y we have $|\mathbb{E}[XY]| \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[Y^2]}$ assuming $\mathbb{E}[X^2], \mathbb{E}[Y^2] < \infty$. As a special case $\text{cov}(X, Y)^2 \leq \text{var}(X)\text{var}(Y)$ (assuming variances are defined).
- For an estimator $\hat{\theta}$ of θ we have the bias-variance decomposition $\text{MSE}(\hat{\theta}) = \mathbb{E}[(\hat{\theta} - \theta)^2] = \text{var}(\hat{\theta}) + b(\theta)^2$
- When we decide between the hypotheses, we can make two sorts of error:
 - Type I error (false positive): H_0 is true, but we wrongly reject it (and choose H_1);
 - Type II error (false negative): H_1 is true, but we wrongly accept H_0 .

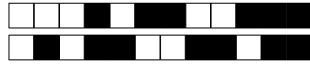
	Decision	
	Accept H_0	Reject H_0
State of nature H_0 true	Correct choice (true negative)	Type I error (false positive)
State of nature H_1 true	Type II error (false negative)	Correct choice (true positive)

Further, we call

- the false positive probability the *size* α of the test, and
 - the true positive probability the *power* β of the test.
- Pearson statistic (or chi-square statistic). Let O_1, \dots, O_k be the number of observations of a random sample of size $n = n_1 + \dots + n_k$ falling into the categories $1, \dots, k$, whose expected numbers are E_1, \dots, E_k , where $E_i > 0$. Then the Pearson statistic (or chi-square statistic) is $T = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$. If the joint distribution of O_1, \dots, O_k is multinomial with denominator n and probabilities $p_1 = \frac{E_1}{n}, \dots, p_k = \frac{E_k}{n}$, then $T \sim \chi_{k-1}^2$, the approximation being good if $k^{-1} \sum E_i \geq 5$.



- Neyman–Pearson Lemma. Let $f_0(y), f_1(y)$ be the densities of Y under simple null and alternative hypotheses. Assume the set $\mathcal{Y}_\alpha = \{y \in \Omega : \frac{f_1(y)}{f_0(y)} > t_\alpha\}$ such that $\mathbb{P}_0(Y \in \mathcal{Y}_\alpha) = \alpha$ exists. Then, \mathcal{Y}_α maximises $\mathbb{P}_1(Y \in \mathcal{Y}_\alpha)$ amongst all the \mathcal{Y}' such that $\mathbb{P}_0(Y \in \mathcal{Y}') \leq \alpha$. Thus, to maximise the power of a given threshold, we must base the decision on \mathcal{Y}_α (\mathcal{Y}_α should be the reject region for H_0).

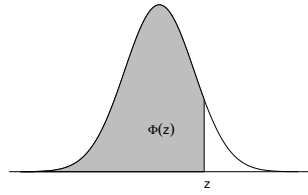


Distributions

Distribution	PMF/PDF	Expected Value	Variance	MGF
Bernoulli Bern(p)	$P(X = 1) = p$ $P(X = 0) = 1 - p$	p	$p(1 - p)$	$1 - p + pe^t$
Binomial Bin(n, p)	$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$ $k \in \{0, 1, 2, \dots, n\}$	np	$np(1 - p)$	$(1 - p + pe^t)^n$
Geometric Geom(p)	$P(X = k) = (1 - p)^{k-1} p$ $k \in \{1, 2, \dots\}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$ $(1 - p)e^t < 1$
Neg. Binom. NegBin(r, p)	$P(X = x) = \binom{x-1}{r-1} p^r (1 - p)^{x-r}$ $x \in \{r, r + 1, r + 2, \dots\}$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$	$(\frac{pe^t}{1-(1-p)e^t})^r$ $(1 - p)e^t < 1$
Hypergeom. HypG(w, b, n)	$P(X = k) = \frac{\binom{w}{k} \binom{b}{n-k}}{\binom{w+b}{n}}$ $k \in \{0, 1, 2, \dots, n\}$	$\mu = \frac{nw}{b+w}$	$(\frac{w+b-n}{w+b-1}) n \frac{\mu}{n} (1 - \frac{\mu}{n})$	messy
Poisson Pois(λ)	$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$ $k \in \{0, 1, 2, \dots\}$	λ	λ	$e^{\lambda(e^t - 1)}$
Uniform U(a, b)	$f(x) = \frac{1}{b-a}$ $x \in [a, b]$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{tb} - e^{ta}}{t(b-a)}$
Exponential exp(λ)	$f(x) = \lambda e^{-\lambda x}$ $x \in (0, \infty)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda - t}, t < \lambda$
Normal $\mathcal{N}(\mu, \sigma^2)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $x \in (-\infty, \infty)$	μ	σ^2	$e^{t\mu + \frac{\sigma^2 t^2}{2}}$
Chi-Square χ_n^2	$\frac{1}{2^{n/2} \Gamma(n/2)} x^{n/2-1} e^{-x/2}$ $x \in (0, \infty)$	n	$2n$	$(1 - 2t)^{-n/2}$ $t < 1/2$



Standard normal distribution $\Phi(z)$

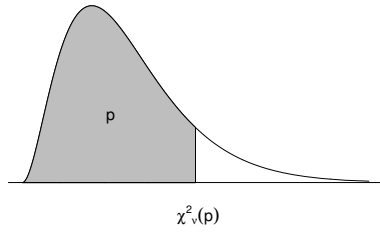


For $z < 0$ we use symmetry: $\mathbb{P}(Z \leq z) = \Phi(z) = 1 - \Phi(-z)$, $z \in \mathbb{R}$.

z	0	1	2	3	4	5	6	7	8	9
0.0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
0.1	.53983	.54380	.54776	.55172	.55567	.55962	.56356	.56750	.57142	.57535
0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
1.0	.84134	.84375	.84614	.84850	.85083	.85314	.85543	.85769	.85993	.86214
1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91309	.91466	.91621	.91774
1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92786	.92922	.93056	.93189
1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
2.0	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361
2.5	.99379	.99396	.99413	.99430	.99446	.99461	.99477	.99492	.99506	.99520
2.6	.99534	.99547	.99560	.99573	.99585	.99598	.99609	.99621	.99632	.99643
2.7	.99653	.99664	.99674	.99683	.99693	.99702	.99711	.99720	.99728	.99736
2.8	.99744	.99752	.99760	.99767	.99774	.99781	.99788	.99795	.99801	.99807
2.9	.99813	.99819	.99825	.99831	.99836	.99841	.99846	.99851	.99856	.99861
3.0	.99865	.99869	.99874	.99878	.99882	.99886	.99889	.99893	.99896	.99900
3.1	.99903	.99906	.99910	.99913	.99916	.99918	.99921	.99924	.99926	.99929
3.2	.99931	.99934	.99936	.99938	.99940	.99942	.99944	.99946	.99948	.99950
3.3	.99952	.99953	.99955	.99957	.99958	.99960	.99961	.99962	.99964	.99965
3.4	.99966	.99968	.99969	.99970	.99971	.99972	.99973	.99974	.99975	.99976
3.5	.99977	.99978	.99978	.99979	.99980	.99981	.99981	.99982	.99983	.99983
3.6	.99984	.99985	.99985	.99986	.99986	.99987	.99987	.99988	.99988	.99989
3.7	.99989	.99990	.99990	.99990	.99991	.99991	.99992	.99992	.99992	.99992
3.8	.99993	.99993	.99993	.99994	.99994	.99994	.99994	.99995	.99995	.99995
3.9	.99995	.99995	.99996	.99996	.99996	.99996	.99996	.99996	.99997	.99997



χ^2_ν distribution



$\chi^2_\nu(p)$: quantiles for the chi-square distribution with ν degrees of freedom.

ν	.005	.01	.025	.05	.10	.25	.50	.75	.90	.95	.975	.99	.995	.999
1	0	.0002	.010	.0039	.0158	.102	.455	1.32	2.71	3.84	5.02	6.63	7.88	10.8
2	.0100	.0201	.0506	.103	.211	.575	1.39	2.77	4.61	5.99	7.38	9.21	10.6	13.8
3	.0717	.115	.216	.352	.584	1.21	2.37	4.11	6.25	7.81	9.35	11.3	12.8	16.3
4	.207	.297	.484	.711	1.06	1.92	3.36	5.39	7.78	9.49	11.1	13.3	14.9	18.5
5	.412	.554	.831	1.15	1.61	2.67	4.35	6.63	9.24	11.1	12.8	15.1	16.7	20.5
6	.676	.872	1.24	1.64	2.20	3.45	5.35	7.84	10.6	12.6	14.4	16.8	18.5	22.5
7	.989	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.0	14.1	16.0	18.5	20.3	24.3
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.2	13.4	15.5	17.5	20.1	22.0	26.1
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.4	14.7	16.9	19.0	21.7	23.6	27.9
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.5	16.0	18.3	20.5	23.2	25.2	29.6
11	2.60	3.05	3.82	4.57	5.58	7.58	10.3	13.7	17.3	19.7	21.9	24.7	26.8	31.3
12	3.07	3.57	4.40	5.23	6.30	8.44	11.3	14.8	18.5	21.0	23.3	26.2	28.3	32.9
13	3.57	4.11	5.01	5.89	7.04	9.30	12.3	16.0	19.8	22.4	24.7	27.7	29.8	34.5
14	4.07	4.66	5.63	6.57	7.79	10.2	13.3	17.1	21.1	23.7	26.1	29.1	31.3	36.1
15	4.60	5.23	6.26	7.26	8.55	11.0	14.3	18.2	22.3	25.0	27.5	30.6	32.8	37.7
16	5.14	5.81	6.91	7.96	9.31	11.9	15.3	19.4	23.5	26.3	28.8	32.0	34.3	39.3
17	5.70	6.41	7.56	8.67	10.1	12.8	16.3	20.5	24.8	27.6	30.2	33.4	35.7	40.8
18	6.26	7.01	8.23	9.39	10.9	13.7	17.3	21.6	26.0	28.9	31.5	34.8	37.2	42.3
19	6.84	7.63	8.91	10.1	11.7	14.6	18.3	22.7	27.2	30.1	32.9	36.2	38.6	43.8
20	7.43	8.26	9.59	10.9	12.4	15.5	19.3	23.8	28.4	31.4	34.2	37.6	40.0	45.3
21	8.03	8.90	10.3	11.6	13.2	16.3	20.3	24.9	29.6	32.7	35.5	38.9	41.4	46.8
22	8.64	9.54	11.0	12.3	14.0	17.2	21.3	26.0	30.8	33.9	36.8	40.3	42.8	48.3
23	9.26	10.2	11.7	13.1	14.8	18.1	22.3	27.1	32.0	35.2	38.1	41.6	44.2	49.7
24	9.89	10.9	12.4	13.8	15.7	19.0	23.3	28.2	33.2	36.4	39.4	43.0	45.6	51.2
25	10.5	11.5	13.1	14.6	16.5	19.9	24.3	29.3	34.4	37.7	40.6	44.3	46.9	52.6
26	11.2	12.2	13.8	15.4	17.3	20.8	25.3	30.4	35.6	38.9	41.9	45.6	48.3	54.1
27	11.8	12.9	14.6	16.2	18.1	21.7	26.3	31.5	36.7	40.1	43.2	47.0	49.6	55.5
28	12.5	13.6	15.3	16.9	18.9	22.7	27.3	32.6	37.9	41.3	44.5	48.3	51.0	56.9
29	13.1	14.3	16.0	17.7	19.8	23.6	28.3	33.7	39.1	42.6	45.7	49.6	52.3	58.3
30	13.8	15.0	16.8	18.5	20.6	24.5	29.3	34.8	40.3	43.8	47.0	50.9	53.7	59.7
40	20.7	22.2	24.4	26.5	29.1	33.7	39.3	45.6	51.8	55.8	59.3	63.7	66.8	73.4
50	28.0	29.7	32.4	34.8	37.7	42.9	49.3	56.3	63.2	67.5	71.4	76.2	79.5	86.7
60	35.5	37.5	40.5	43.2	46.5	52.3	59.3	67.0	74.4	79.1	83.3	88.4	92.0	99.6
70	43.3	45.4	48.8	51.7	55.3	61.7	69.3	77.6	85.5	90.5	95.0	100.	104.	112.