Important Distributions

The nine most important distributions we have discussed so far are listed below, each with its PMF/PDF. Each of these distributions is important because it has a natural, useful story, so understanding these stories (and recognizing equivalent stories) is crucial. It is also important to know how these distributions are related to each other. For example, Bern(p) is the same as Bin($1, p$), and Bin(n, p) is approximately Pois(λ) if n is large, p is small, and $\lambda = np$ is moderate.

Name	**Param.**	**PMF or PDF**
Bernoulli	p	$P(X=1)=p,\ P(X=0)=1-p$
Binomial	n, p	$\binom{n}{k}p^k(1-p)^{n-k}$, for $k \in \{0,1,\dots,n\}$
Geometric	p	$(1-p)^k p$, for $k \in \{0,1,2,\dots\}$
Negative Binomial	r, p	$\binom{r+n-1}{r-1}p^r(1-p)^n$, $n \in \{0,1,2,\dots\}$
Hypergeometric	w, b, n	$\frac{\binom{w}{k}\binom{b}{n-k}}{\binom{w+b}{n}}$, for $k \in \{0,1,\dots,n\}$
Poisson	λ	$\frac{e^{-\lambda}\lambda^k}{k!}$, for $k \in \{0,1,2,\dots\}$
Exponential	λ	$\lambda e^{-\lambda x}$ for $x > 0$
Uniform	$a < b$	$\frac{1}{b-a}$, for $x \in (a,b)$
Normal	μ, σ^2	$\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Some Useful Formulas

De Morgan's Laws

$$(A_1 \cup A_2 \cup \cdots \cup A_n)^c = A_1^c \cap A_2^c \cap \cdots \cap A_n^c$$
$$(A_1 \cap A_2 \cap \cdots \cap A_n)^c = A_1^c \cup A_2^c \cup \cdots \cup A_n^c$$

Complements

$$P(A^c) = 1 - P(A)$$

Unions

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} P(A_i), \text{ if the } A_i \text{ are disjoint}$$

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) \le \sum_{i=1}^{n} P(A_i)$$

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{k=1}^{n} \left((-1)^{k+1} \sum_{i_1 < i_2 < \cdots < i_k} P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) \right)$$

(Inclusion-Exclusion)

Intersections

$$P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$$

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1, A_2) \cdots P(A_n|A_1, \dots, A_{n-1})$$

Law of Total Probability

If $E_1, E_2, \dots, E_n$ are a partition of the sample space S (i.e., they are disjoint and their union is all of S) and $P(E_j) \neq 0$ for all j, then

$$P(B) = \sum_{j=1}^{n} P(B|E_j)P(E_j)$$

PROBABILITY AND STATISTICS (MATH-235)

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Final exam - questions

Date: 27th of January, 2025
Time: 9:15–12:15

Name: ________________________

SCIPER: ________________________

INSTRUCTIONS TO CANDIDATES

- To obtain the maximum number of points you should be clear about your reasoning and present your arguments explicitly. You have **3 hours** to complete the exam.
- All that can be used for this exam is a pen and the cheat sheet that **we provide for you**. No external notes, books, summaries, formula collections or calculators are allowed. All questions should be answered.
- The finest enumerated item in each question will be marked on a scale of $0-2$ points, indicating an incorrect, partially correct and completely correct answer respectively (half-points are not given). **The exam has 8 questions with a total of 84 points**.
- **Write the answer to every question in the other booklet (final exam - answers)**. Scrap paper will be provided for rough work, but only answers written in the booklet will be marked.
- At the end of the exam, you will have to return everything : the booklet with the questions, the booklet with your answers, and the scrap paper.
-
- When the exercise directs you to "**Simplify!**" provide the simplest fraction, a decimal, or the simplest form possible using the function or parameter in terms of which your answer must be given. If you are explicitly instructed with "**Do not simplify!**", it is acceptable to leave your answer as a non-simplified fraction, integral, sum, or another more complex form, depending on the context.
- Please write your name and SCIPER number on every sheet you use.
- You can write your answers in English or French, but please commit to the same language throughout the exam.

Mark question 1 (TOT: 14 points):

Mark question 2 (TOT: 8 points):

Mark question 3 (TOT: 8 points):

Mark question 4 (TOT: 12 points):

Mark question 5 (TOT: 14 points):

Mark question 6 (TOT: 6 points):

Mark question 7 (TOT: 8 points):

Mark question 8 (TOT: 14 points):

Variable definitions

All the random variables have finite mean and variance.

Question 1.

Write the most appropriate of $\leq$, $\geq$, $=$ or ? in the blank for each part (where "?" means that no relation holds in general.). No justification is needed for full points. In this exercise, you do not get partial points, for each subquestion you can get 2 or 0 points.

In this exercise, suppose A and B are independent events, while the positive random variables X and Y are not necessarily independent, however, identically distributed.

(a) You toss two fair coins, Coin C_1 and C_2 twice and four times, respectively.
$P(\text{At least 1 Head out of the 2 tosses with } C_1)$ ____
$P(\text{At least 2 Heads out of the 4 tosses with } C_2)$

(b) $P((A \cup B)^c)$ ____ $P(A^c)P(B^c)$

(c) $P(A|B)$ ____ $P(B|B)$

(d) $E((X+Y)^2)$ ____ $2E(X^2)$

(e) $E(E(\log(Y)|X))$ ____ $\log(E(Y))$

(f) $E\left(\frac{X}{Y}\right)$ ____ 1

(g) $P(X > 3)$ ____ $\frac{1}{20}E(X^3)$

Solution 1.

(a) You toss two fair coins, Coin C_1 and C_2 twice and four times, respectively.
$P(\text{At least 1 Head out of the 2 tosses with } C_1)$ $\underline{\;\geq\;}$
$P(\text{At least 2 Heads out of the 4 tosses with } C_2)$
Follows from

$$\frac{\binom{2}{1}+\binom{2}{2}}{2^2} = \frac{3}{4} \geq \frac{11}{16} = \frac{\binom{4}{2}+\binom{4}{3}+\binom{4}{4}}{2^4}.$$

(b) $P((A \cup B)^c)$ $\underline{\;=\;}$ $P(A^c)P(B^c)$

$$P((A \cup B)^c) = P(A^c \cap B^c) = P(A^c)P(B^c),$$

by de Morgan's law and independence.

(c) $P(A|B)$ $\underline{\;\leq\;}$ $P(B|B)$

$$P(B|B) = 1 \geq P(A|B).$$

(d) $E((X+Y)^2)$ $\underline{\;\geq\;}$ $2E(X^2)$

$$E((X+Y)^2) = E(X^2) + 2E(XY) + E(Y^2) = 2E(X^2) + 2E(XY) \geq 2E(X^2),$$

by using that X and Y have the same distribution and the positivity of the random variables, i.e. that $E(XY) > 0$.

(e) $E(E(\log(Y)|X))$ $\underline{\ \leq\ }$ $\log(E(Y))$

$$E(E(\log(Y)|X)) = E(\log(\mathrm{Y})) \leq \log(\mathrm{E(Y)}),$$

by Adam's law (tower property/law of total expectation) and Jensen's inequality.

(f) $E\left(\frac{X}{Y}\right)$ $\underline{\ \geq\ }$ 1

$$E(X/Y) \geq 1/E(Y/X)$$

by Jensen's inequality, and using that X and Y are positive. Since X and Y are identically distributed, $E(X/Y) = E(Y/X)$, so $E(X/Y)^2 \geq 1$, thus from the positivity of X/Y, the inequality follows.

(g) $P(X > 3)$ $\underline{\ \leq\ }$ $\frac{1}{20}E(X^3)$

$$P(X > 3) = P(X^3 > 27) \leq \frac{E(X^3)}{27} \leq \frac{E(X^3)}{20},$$

by Markov's inequality for a positive random variable.

Question 2.

A company designed two tests to identify SARS-CoV-infected individuals, Test A and Test B. An individual who has the disease (event D) or is disease-free (event D^c) is tested. Suppose that 5% of the population has the disease. The tests can either be positive (events T_A and T_B) or negative (events T_A^c and T_B^c). Suppose that Test A has a sensitivity and specificity of 90% and Test B has a sensitivity and specificity of 50%, that is

$$P(T_A|D) = P(T_A^c|D^c) = 0.9$$
$$P(T_B|D) = P(T_B^c|D^c) = 0.5.$$

Given whether an individual is infected or not, the test results are conditionally independent.

(a) An individual sampled at random is tested with Test A and returns as positive. Conditional on the test result, what is the probability that this individual has the disease? Do not simplify!

(b) A second individual independent of the first one is tested with Test B and returns as negative. Conditional on the test result, what is the probability that the second individual is disease-free? Simplify!

(c) Would multiple negative results of Test B increase the conditional probability of the disease-free event, given these results? Justify!

(d) Are T_A and T_B (unconditionally) independent? It is sufficient to justify with words, a numerical computation is not necessarily required.

Solution 2.

(a) By Bayes' rule and the law of total probability

$$P(D|T_A) = \frac{P(T_A|D)P(D)}{P(T_A|D)P(D) + P(T_A|D^c)P(D^c)} = \frac{0.9 \cdot 0.05}{0.9 \cdot 0.05 + 0.1 \cdot 0.95}.$$

If they simplify (which is not required), they should get $\frac{9}{28}$.

(b) By Bayes' rule and the law of total probability

$$P(D^c|T_B) = \frac{P(T_B|D^c)P(D^c)}{P(T_B|D)P(D) + P(T_B|D^c)P(D^c)} = \frac{0.5 \cdot 0.95}{0.5 \cdot 0.05 + 0.5 \cdot 0.95} = \frac{0.95}{1} = 0.95.$$

The fraction form should also receive full marks: $\frac{19}{20}$.

(c) Either they can calculate multiple conditional probabilities, or they can argue that by part (b) it is confirmed that the probability of no infection given the test result is the same as the unconditional probability, i.e. $P(D^c) = P(D^c|T_B)$, thus the outcome and the test are independent, so multiple test results would contain no information.

Alternatively, they can argue that Test B is as good as random guessing and is therefore independent of the outcome, meaning no additional information is gained through multiple testing.

(d) Since T_B corresponds to random guessing, T_B and T_A are independent. Any reasonable explanation arguing for independence should receive full marks.

Technically, they should mention that this holds due to the conditional independence that is stated in the pretext of the exercise ($P(T_A, T_B|D) = P(T_A|D)P(T_B|D)$), but even if they miss this detail, but their argument is coherent, give full marks.

Question 3.

Let X and Y have joint probability density function (PDF) $f_{X,Y}(x,y) = c(x+y)$, for $0 < x < y < 1$, and $f_{X,Y}(x,y) = 0$ otherwise.

(a) Find c to make this a valid joint PDF.

(b) Find the marginal PDFs of X and Y.

(c) Are X and Y independent?

(d) Find the conditional PDF of Y given $X = x$.

Solution 3.

(a) A PDF must be non-negative and should integrate to 1 over the support of it. To satisfy the first condition $c > 0$ must hold, for the second, solve the integral

$$\int_0^1 \int_0^y c(x+y)\,\mathrm{d}x\mathrm{d}y = c\int_0^1 \left[\frac{x^2}{2} + xy\right]_{x=0}^{y} \mathrm{d}y = c\int_0^1 \frac{3y^2}{2}\,\mathrm{d}y = c\left[\frac{y^3}{2}\right]_{y=0}^{1} = \frac{c}{2}.$$

We have to find c such that $\frac{c}{2} = 1$ holds, so $c = 2$.

(b) Integrating out the X and Y respectively, we get

$$f_X(x) = \int_x^1 2(x+y)\,\mathrm{d}y = 2\left[xy + \frac{y^2}{2}\right]_{y=x}^{1} = 2\left(x + \frac{1}{2} - \frac{3x^2}{2}\right) = -3x^2 + 2x + 1,$$

$$f_Y(y) = \int_0^y 2(x+y)\,\mathrm{d}x = 2\left[\frac{x^2}{2} + xy\right]_{x=0}^{y} = 2\left(\frac{3y^2}{2}\right) = 3y^2.$$

(c) No, they are not independent as

$$f_X(x) \cdot f_Y(y) = (-3x^2 + 2x + 1) \cdot 3y^2 \neq 2(x+y) = f_{X,Y}(x,y).$$

(d) $f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{2(x+y)}{-3x^2+2x+1}$

Question 4.

Suppose a factory produces a random number of gadgets, N, where $N \sim Pois(\lambda)$. Each gadget is independently operational with probability p and is faulty with probability $q = 1-p$. Let X be the number of operational gadgets and Y the number of faulty gadgets, so $X+Y=N$.

(a) Find the conditional distributions $X|N=1$ and $Y|N=1$.

(b) Find the conditional distributions $X|N=n$ and $Y|N=n$.

(c) Is $P(X=x, Y=y|N=x+y)$ equal to $P(X=x|N=x+y)$? Justify!

(d) What is the joint PMF of X and Y? **Hint:** $e^z = e^{zc+z(1-c)}$

(e) What are the marginal PMFs of X and Y?

(f) Are X and Y independent? Are X and Y conditionally independent given $N=n$? Justify!

Solution 4.

(a) Given that $N=1$ gadget was produced, it is operational with probability p and faulty with probability q, therefore $X=1$ with probability p and $X=0$ with probability $1-p$, while $P(Y=1)=q$ and $P(Y=0)=1-q$, thus

$$X|N=1 \sim \text{Bernoulli}(p)$$
$$Y|N=1 \sim \text{Bernoulli}(q)$$

(b) If $N=n$ gadgets were produced, that means that n independent trials are performed, with independent success probabilities p and q, corresponding to X and Y respectively. So

$$X|N=n \sim \text{Bin}(n,p)$$
$$Y|N=n \sim \text{Bin}(n,q)$$

(c) By the definition of N, it follows that $Y=N-X$. Writing out the conditional probability, we get

$$P(X=x, Y=y|N=x+y) = \frac{P((X=x)\cap(Y=y)\cap(N=x+y)}{P(N=x+y)}$$
$$= \frac{P((X=x)\cap(N-X=y)\cap(N=x+y)}{P(N=x+y)} = \frac{P((X=x)\cap(N=x+y)}{P(N=x+y)}$$
$$= P(X=x|N=x+y),$$

where we used that $\{(X=x)\cap(N=x+y)\} \subseteq \{N-X=y\}$

Alternatively, they can argue directly that conditional on $N=x+y$ the events $X=x$ and $Y=y$ are the same.

(d) Note that the two events $\{(X = x) \cap (Y = y)\}$ and $\{(X = x) \cap (X + Y = x + y)\}$ are equal. Using this and the definition of the conditional probability, we can write

$$P(X = x, Y = y) = P(X = x, N = x + y) = P(X = x|N = x + y)P(N = x + y)$$
$$= \left(\binom{x+y}{x} p^x (1-p)^{x+y-x}\right)\left(\frac{e^{-\lambda}\lambda^{x+y}}{(x+y)!}\right) = \frac{(x+y)!}{x!y!} p^x q^y \frac{e^{-p\lambda}e^{-q\lambda}\lambda^x\lambda^y}{(x+y)!}$$
$$= \frac{e^{-p\lambda}(p\lambda)^x}{x!}\frac{e^{-q\lambda}(q\lambda)^y}{y!},$$

where we used that $N \sim \text{Pois}(\lambda)$ and that from part (b) $X|N = x+y \sim \text{Bin}(x+y, p)$.

(e) Note that from part (d), the joint PMF of X and Y is a product of two Poisson PMF-s with parameters $p\lambda$ and $q\lambda$ respectively. Thus the marginal PMF-s are

$$P(X = x) = \frac{e^{-p\lambda}(p\lambda)^x}{x!}$$
$$P(Y = y) = \frac{e^{-q\lambda}(q\lambda)^y}{y!}.$$

Alternatively they can write that from part (d) it follows that $X \sim \text{Pois}(p\lambda)$ and $Y \sim \text{Pois}(q\lambda)$

(f) From part (e) it follows that X and Y are independent, as $P(X = x, Y = y) = P(X = x)P(Y = y)$. Since $\{X = x|N = x + y\}$ implies $\{Y = y|N = x + y\}$ and vice versa, X and Y are **not** conditionally independent given $N = n$. Alternatively they can argue using parts (b) and (c) to show that the product of the conditional marginals is not equal to the conditional joint.

Question 5.

Let $X \sim \text{Unif}(-1, 1)$.

(a) Define $Y = X^2 - \frac{1}{3}$.

(i) Find the PDF of Y.

(ii) Find $E(Y)$.

(iii) Find $\text{Cov}(X, Y)$.

(iv) Are X and Y independent? Justify!

(b) Define $W = \frac{X+1}{2}$.

(i) What is the distribution of W?

(ii) As a function of W create Z with CDF $F(z) = e^z/(1 + e^z)$.

(iii) Find $E(Z)$.

Solution 5.

(a) (i) First we have to find the CDF. Since $X^2 - \frac{1}{3}$ takes values in $[-\frac{1}{3}, \frac{2}{3}]$, $F_Y(y) = 0$ for $y \leq -\frac{1}{3}$ and $F_Y(y) = 1$ for $y \geq \frac{2}{3}$. For $y \in [-\frac{1}{3}, \frac{2}{3}]$

$$F_Y(y) = P(Y \leq y) = P\left(X^2 - \frac{1}{3} \leq y\right) = P\left(X^2 \leq y + \frac{1}{3}\right)$$
$$= P\left(-\sqrt{y + \frac{1}{3}} \leq X \leq \sqrt{y + \frac{1}{3}}\right) = \sqrt{y + \frac{1}{3}}.$$

Then the PDF of Y is $f_Y(y) = F_Y'(y)$, so

$$f_Y(y) = \begin{cases} \frac{1}{2\sqrt{y+\frac{1}{3}}} & \text{if } y \in [-\frac{1}{3}, \frac{2}{3}] \\ 0 & \text{otherwise} \end{cases}$$

(ii) By the "Law of the Unconscious Statistician"

$$E(Y) = E\left(X^2 - \frac{1}{3}\right) = \int_{-1}^{1} x^2 \frac{1}{2} \, \mathrm{d}x - \frac{1}{3} = \left[\frac{x^3}{6}\right]_{x=-1}^{1} - \frac{1}{3} = \frac{1}{6} + \frac{1}{6} - \frac{1}{3} = 0.$$

Alternatively, for full marks they can calculate $E(Y)$ from the PDF, or $E(X^2)$ from the variance of a Uniform random variable.

(iii) $\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = E(XY)$, since by part (a)(ii) $E(Y) = 0$ (in fact $E(X) = 0$ as well).
$E(XY) = E\left(X\left(X^2 - \frac{1}{3}\right)\right) = E(X^3) - \frac{1}{3}E(X)$. At this point they can realize that x^3 and x are odd functions and X has PDF that is symmetric around 0, thus the expectations and consequently the covariance is equal to 0.

Alternatively

$$E(X^3) = \int_{-1}^{1} x^3 \frac{1}{2}\,\mathrm{d}x = \left[\frac{x^4}{8}\right]_{x=-1}^{1} = \frac{1}{8} - \frac{1}{8} = 0$$

and

$$E(X) = \int_{-1}^{1} x \frac{1}{2}\,\mathrm{d}x = \left[\frac{x^2}{4}\right]_{x=-1}^{1} = \frac{1}{4} - \frac{1}{4} = 0,$$

thus $\mathrm{Cov}(X, Y) = E(XY) = 0$.

(iv) Even though $\mathrm{Cov}(X, Y) = 0$, they are not necessarily independent, independence implies 0 covariance, but not the other way around. Since conditioning on the value of X determines the value of Y, they are clearly not independent.

(b) (i)

$$F_W(w) = P(W \le w) = P\left(\frac{X+1}{2} \le w\right) = P(X \le 2w+1)$$
$$= F_X(2w+1) = \frac{2w+1-1}{1-(-1)} = w,$$

thus $W \sim \mathrm{Unif}(0, 1)$.
Alternatively they can argue that if $U \sim \mathrm{Unif}(a, b)$, then $cU + d \sim \mathrm{Unif}(ca + d, cb + d)$. Here $a = -1, b = 1, c = \frac{1}{2}$ and $d = \frac{1}{2}$, thus $ca + d = -\frac{1}{2} + \frac{1}{2} = 0$ and $cb + d = \frac{1}{2} + \frac{1}{2} = 1$, thus $W \sim \mathrm{Unif}(0, 1)$.

(ii) By the universality of the uniform if $Z = F^{-1}(W)$, then Z is distributed according to the CDF F. Inverting $F(z)$ gives

$$F^{-1}(w) = \log\left(\frac{w}{1-w}\right).$$

Thus $Z = \log\left(\frac{W}{1-W}\right)$ has the desired CDF.

(iii) Since $\log(\frac{a}{b}) = \log(a) - \log(b)$ we can rewrite the expectation as

$$E(Z) = E\left(\log\left(\frac{W}{1-W}\right)\right) = E(\log(\mathrm{W})) - E(\log(1-W)).$$

As $W \sim \mathrm{Unif}(0, 1)$ it follows (using the location-scale transformation or otherwise) that $1 - W \sim \mathrm{Unif}(0, 1)$, thus $E(\log(\mathrm{W})) = E(\log(1 - W))$, that means $E(Z) = 0$.
Alternatively they can calculate the PDF of Z and do integration by parts.

Question 6.

Suppose X is a non-negative random variable.

(a) Argue that $X = XI(X > 0)$, where $I(X > 0)$ is the indicator random variable of $X > 0$.

(b) Show that $P(X > 0) \geq \frac{(E(X))^2}{E(X^2)}$.

Hint: Use the Cauchy-Schwarz inequality.

(c) Suppose a student is presented with 5 multiple choice questions with 4 possible answers to each of them. They answer all questions completely at random and independently from each other. Using the inequality above, give an upper bound for the probability that the student made no mistakes. Simplify!

Solution 6.

(a)

$$XI(X > 0) = \begin{cases} 0 \cdot 0 & \text{if } X = 0 \\ X \cdot 1 & \text{if } X > 0 \end{cases} = \begin{cases} 0 & \text{if } X = 0 \\ X & \text{if } X > 0 \end{cases}$$

thus $X = XI(X > 0)$ for all values of X.

(b) By part (a) and the Cauchy-Schwarz inequality

$$E(X) = E(XI(X > 0)) = |E(XI(X > 0))| \leq \sqrt{E((I(X > 0)^2)E(X^2)},$$

where we used that X is non-negative. Note that since the indicator random variable only takes values 0 and 1, $I(X > 0)^2 = I(X > 0)$. Moreover, using the "fundamental bridge" $E(I(X > 0)) = P(X > 0)$.

By taking the square of both sides

$$(E(X))^2 \leq P(X > 0)E(X^2),$$

thus after dividing both sides by $E(X^2)$, we get the desired inequality.

(c) Denote with X the number of questions the student got wrong. Then we want to bound the probability $P(X = 0) = 1 - P(X > 0)$. Since $X \sim \text{Bin}\left(5, \frac{3}{4}\right)$, $E(X) = 5 \cdot \frac{3}{4} = \frac{15}{4}$, so $(E(X))^2 = \frac{225}{16}$. Since $\text{Var}(X) = 5 \cdot \frac{3}{4} \cdot \frac{1}{4} = \frac{15}{16}$, and $\text{Var}(X) = E(X^2) - (E(X)^2)$, it follows that $E(X^2) = \frac{15}{16} + \frac{225}{16} = \frac{240}{16} = 15$. Thus

$$\frac{(E(X)^2)}{E(X^2)} = \frac{\frac{225}{16}}{15} = \frac{15}{16},$$

therefore

$$P(X = 0) = 1 - P(X > 0) \leq 1 - \frac{15}{16} = \frac{1}{16}.$$

Question 7.

Let $Y = 4e^{3X}$, with $X \sim Exp(7)$.

(a) Find the mean of Y. Simplify!

(b) Find the variance of Y. Simplify!

(c) Suppose $Y_1, \ldots, Y_n$ are i.i.d. with the same distribution as Y. What is the approximate distribution of the sample mean $\overline{Y}_n = \frac{\sum_{i=1}^n Y_i}{n}$ when n is large? Simplify, and specify all parameters!

(d) You receive a random sample $y_1, \ldots, y_7$ from the distribution of Y. Suppose that the mean (μ) of Y is unknown; however, the variance (σ^2) is known, as calculated in part (b).
We find that

$$\overline{y}_7 = \frac{\sum_{i=1}^7 y_i}{7} = 6.5 \ .$$

Suppose that 7 is "large" in the sense of part (c).
Construct a 95% confidence interval for μ. Simplify!
Hint: You can take $\Phi^{-1}(0.975) \approx 2$

Solution 7.
Disclaimer: Originally, in the exam, I had a typo in the hint saying $\Phi(0.975) \approx 2$ instead $\boldsymbol{\Phi^{-1}(0.975) \approx 2}$. Even though this was clarified during the exam, if they made a similar typo in part (d), do not penalize that.

(a) By the "Law of the Unconscious Statistician"

$$E(Y) = \int_0^\infty (4e^{3x}) \cdot (7e^{-7x})\,\mathrm{d}x = 28 \int_0^\infty e^{-4x}\,\mathrm{d}x = 28 \left[-\frac{e^{-4x}}{4}\right]_0^\infty = 0 + 7 = 7.$$

(b) By the same law

$$E(Y^2) = \int_0^\infty (16e^{6x}) \cdot (7e^{-7x})\,\mathrm{d}x = 112 \int_0^\infty e^{-x}\,\mathrm{d}x = 112 \left[-e^{-x}\right]_0^\infty = 0 + 112 = 112.$$

Using that $\mathrm{Var}(Y) = E(Y^2) - (E(Y))^2$,

$$\mathrm{Var}(Y) = 112 - 7^2 = 112 - 49 = 63.$$

(c) By the Central Limit Theorem and a location-scale transformation it follows that as $n \to \infty$, the sample mean $\overline{Y}$ converges in distribution to $\mathcal{N}\left(E(Y), \frac{\mathrm{Var(Y)}}{n}\right)$, that is, for large n, $\overline{Y}$ is approximately $\mathcal{N}\left(7, \frac{63}{n}\right)$ distributed.

(d) Using the normal approximation (since 7 is "large"), the confidence interval for the mean μ is calculated as

$$\left[\overline{y}_7 - \Phi^{-1}(0.975)\sqrt{\frac{\sigma^2}{n}}, \overline{y}_7 + \Phi(0.975)^{-1}\sqrt{\frac{\sigma^2}{n}}\right].$$

Using the approximation from the hint, and that $\sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{63}{7}} = \sqrt{9} = 3$, we have $\Phi^{-1}(0.975)\sqrt{\frac{\sigma^2}{n}} = 6$, thus the confidence interval is

$$[0.5, 12.5].$$

Question 8.

A total of 200 teams compete in a marathon relay. Each team consists of two members, and each member runs a half marathon; that is, 400 runners complete the half marathon distance. The runners' completion times are assumed to be independent and identically distributed (iid) with a continuous cumulative distribution function (CDF) F.

(a) Five-digit starting numbers are to be created using the digits $1, 2, 3, 4, 5, 6, 7, 8$. If no digit can be repeated, and the first, third, and fifth digits of the numbers must be odd, how many distinct starting numbers can be formed? Simplify!

(b) The runners with 10 fastest half-marathon times receive a prize. What is the probability that *exactly* one team has both of its members included among the 10 runners with the fastest half-marathon times? Do not simplify!

(c) Outside the 400 marathon relay participants, Alice also completes the half marathon. Her running time is independent and identically distributed according to the same CDF F. For the rest of this question, let A represent the number of runners from the 400 relay participants who finish the half marathon faster than Alice, and let T represent Alice's finishing time.

(i) Find the conditional distribution of A, given that $T = t$.

(ii) Find $E(A|T)$.

(iii) Find $E(A)$. Simplify!

(iv) Find $\text{Var}(A|T)$.

(v) Find $\text{Var}(A)$. Simplify!

Solution 8.

(a) There are 4 odd digits in total, so there are $4 \cdot 3 \cdot 2 = 24$ different ways to put a different odd number to the first, third and fifth places. From the remaining 5 numbers we can choose $5 \cdot 4 = 20$ different pairs for the second and the fourth places, thus $20 \cdot 24 = 480$ distinct starting numbers can be formed.

(b) By symmetry, any order of the 400 runners is equally likely, thus we can use the naive definition of probability. We can select the 10 fastest runners out of the 400 runners $\binom{400}{10}$ different ways. For the favorable outcome we can select the one team from which *both* members are among the 10 fastest $\binom{200}{1} = 200$ different ways. Then from the remaining 199 teams we have to select $10 - 2 = 8$ more teams, from which only one of the members is among the 10 fastest, that can happen $\binom{199}{8}$ different ways. From these 8 teams we have to select one member out of the 2, that can happen 2^8 different ways, thus the probability of interest is

$$\frac{200\binom{199}{8}2^8}{\binom{400}{10}}$$

A different formulation would be

$$\frac{\binom{200}{9}\binom{9}{1}2^8}{\binom{400}{10}}$$

with the story that you select the 9 teams for the top 10, then the team from these 9 that includes both members, and then multiply by 2^8 by the same argument.

(c) (i) The probability that an individual runner finishes before Alice, given that Alice has finished at time t is $F(t)$, by the definition of the CDF and from the fact that each runner has finishing time distributed according to F. Since each runner's finishing time is independent from the others' $A|T = t \sim \text{Bin}(400, F(t))$.

(ii) By part (c) (i), $g(t) = E(A|T = t) = 400 \cdot F(t)$. Thus $g(T) = E(A|T) = 400F(T)$.

(iii) By the universality of the uniform, since T is a random variable with CDF F, $F(T) \sim \text{Unif}(0, 1)$. By Adam's law (tower property/law of total probability) $E(A) = E(E(A|T)) = 400E(F(T)) = 400 \cdot \frac{1}{2} = 200$.

(iv) From part (c) (i) it follows that $\text{Var}(A|T = t) = 400F(t)(1 - F(t))$, thus $\text{Var}(A|T) = 400F(T)(1 - F(T))$

(v) By Eve's law (law of total variance)

$$\text{Var}(A) = E(\text{Var}(A|T)) + \text{Var}(E(A|T)).$$

The first term is

$$E(\text{Var}(A|T)) = E[400F(T)(1 - F(T))] = 400[E(F(T)) - E(F(T)^2)] = 400\left(\frac{1}{2} - \frac{1}{3}\right) = \frac{400}{6},$$

where we used again that $F(T) \sim \text{Unif}(0, 1)$. The second term is

$$\text{Var}(E(A|T)) = \text{Var}(400F(T)) = 400^2\text{Var}(F(T)) = \frac{400^2}{12},$$

as $F(T)$ is standard uniform distributed.
Putting the two together

$$\text{Var}(A) = \frac{400}{6} + \frac{400^2}{12} = \frac{400 \cdot (2 + 400)}{12} = \frac{400 \cdot 402}{12} = 67 \cdot 200 = 13400.$$

(Any of the above forms is acceptable for full points.)