M378K Introduction to Mathematical Statistics
Fall 2024
University of Texas at Austin
Practice for In-Term Exam I
Instructor: Milica Čudina

Notes: This is a closed book and closed notes exam. The maximal score on the real exam will be 100 points.
There are many ways in which any single problem can be solved. The solutions herein are just one possible way to tackle the given problems.
Time: 50 minutes

All written work handed in by the student is considered to be
their own work, prepared without unauthorized assistance.

The University Code of Conduct

"The core values of The University of Texas at Austin are learning, discovery, freedom, leadership, individual opportunity, and responsibility. Each member of the university is expected to uphold these values through integrity, honesty, trust, fairness, and respect toward peers and community. As a student of The University of Texas at Austin, I shall abide by the core values of the University and uphold academic integrity."

"I agree that I have complied with the UT Honor Code during my completion of this exam."

Signature:

1.1. DEFINITIONS.

Problem 1.1. Write down the definition of the **cumulative distribution function** of a random variable Y.

Solution:

$$F_Y(x) = \mathbb{P}[Y \leq x] \quad \text{for } x \in \mathbb{R}.$$

Problem 1.2. Let Y be a continuous random variable with the probability density function denoted by f_Y. Let g be a function taking real values such that $g(Y)$ is well defined. How is $\mathbb{E}[g(Y)]$ evaluated using f_Y, if it exists?

Solution: We have that

$$\mathbb{E}[g(Y)] = \int_{-\infty}^{\infty} g(y) f_Y(y)\, dy$$

if the above integral is absolutely convergent.

1.2. TRUE/FALSE QUESTIONS.

Problem 1.3. Let $Y \sim b(n, p)$. Then, $\mathbb{E}[Y] \geq \text{Var}[Y]$. *True or false? Why?*

Solution: **TRUE**
We have

$$\mathbb{E}[Y] = np \geq np(1-p) = \text{Var}[Y]$$

1.3. Free-response problems.

Please, explain carefully all your statements and assumptions. Numerical results or single-word answers without an explanation (even if they're correct) are worth 0 points.

Problem 1.4. A die is rolled 5 times; let the obtained numbers be given by $Y_1, \dots, Y_5$. Use counting to compute the probability that

(1) all $Y_1, \dots, Y_5$ are even?
(2) at most 4 of $Y_1, \dots, Y_5$ are odd?
(3) the values of $Y_1, \dots, Y_5$ are all different from each other?

Solution: There are There are 6^5 different (equally likely) outcomes of 5 rolls of a die. We need to find the number of those 5-tuples of rolls that correspond to the situation described in each question, and simply divide by 6^5.

(1) The number of 5-tuples of rolls where each outcome is even is 3^5, because each roll can come up an even number in three ways, namely 2, 4 or 6. Therefore, The answer is

$$\tfrac{3^5}{6^5} = \left(\tfrac{1}{2}\right)^5 = \tfrac{1}{32}.$$

(2) We have seen above that the number of 5-tuples of rolls where all values are even is 3^5. Therefore, the number of 5-tuples where at most 4 are even is $6^5 - 3^5$, and the required probability is

$$\tfrac{6^5-3^5}{6^5} = \tfrac{31}{32}.$$

(3) Exactly $6 \times 5 \times 4 \times 3 \times 2 = 6!$ 5-tuples have all numbers different. Therefore, the required probability is $\frac{6!}{6^5} = \frac{5}{54}$.

Problem 1.5. The random vector (X, Y) is jointly continuous with the joint probability density function given by

$$f_{(X,Y)}(x, y) = \begin{cases} (1/8)xe^{-(x+y)/2}, & x > 0, y > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Are random variables X and Y independent? Justify your answer; answers without a correct justification will be awarded zero points.

Solution: The joint p.d.f. can be rewritten as

$$f_{(X,Y)}(x, y) = \frac{1}{4}xe^{-x/2} \times \frac{1}{2}e^{-y/2} = f_X(x)f_Y(y)\,.$$

So, the criterion for independence of jointly continuous random variables is satisfied. We conclude that X and Y are independent.

Problem 1.6. *Source: "Probability" by Jim Pitman.*
Consider a set-up in which a transmitter is transmitting either a 0 or a 1 and the receiver indicates that it received either a 0 or a 1. Denote the events that $i = 0, 1$ was transmitted by T_i, and the events that $i = 0, 1$ was indicated as received by R_i.
It is possible to have transmission errors. In fact, you are given the following data on accuracy and the frequency of transmitted signals:

$$\mathbb{P}[R_0 \,|\, T_0] = 0.99,\ \mathbb{P}[R_1 \,|\, T_1] = 0.98,$$

and

$$\mathbb{P}[T_0] = 0.75.$$

(a) Given that the receiver indicated 1, what is the probability that there was an error in the transmission?
(b) What is the overall probability that there was an error in transmission?

Solution:

(1) We need $\mathbb{P}[T_0|R_1]$. By the Bayes formula,

$$\begin{aligned}\mathbb{P}[T_0|R_1] &= \frac{\mathbb{P}[R_1|T_0]\mathbb{P}[T_0]}{\mathbb{P}[R_1|T_0]\mathbb{P}[T_0] + \mathbb{P}[R_1|T_1]\mathbb{P}[T_1]} \\ &= \frac{(1-0.99) \times 0.75}{(1-0.99) \times 0.75 + 0.98 \times 0.25} \\ &= \frac{3}{101}\,.\end{aligned}$$

(2) An error will happen if $T_0 \cap R_1$ or $T_1 \cap R_0$ occur, i.e.,

$$\begin{aligned}\mathbb{P}[\text{error}] &= \mathbb{P}[T_0 \cap R_1] + \mathbb{P}[T_1 \cap R_0] \\ &= \mathbb{P}[R_1|T_0] \times \mathbb{P}[T_0] + \mathbb{P}[R_0|T_1] \times \mathbb{P}[T_1] \\ &= (1 - \mathbb{P}[R_0|T_0]) \times \mathbb{P}[T_0] \\ &\quad + (1 - \mathbb{P}[R_1|T_1]) \times (1 - \mathbb{P}[T_0]) \\ &= 0.01 \times 0.75 + 0.02 \times 0.25 \\ &= \frac{1}{80}\,.\end{aligned}$$

Problem 1.7. A fair coin is tossed 3 times. Let the random variable X stand for the number of heads (H) in the *first* two of the three coin tosses, and let Y stand for the number of tails (T) in the *last* two of the three coin tosses.

- Write down the table for the joint probability (mass) function of the random pair (X, Y).
- Find the marginal distribution of Y.
- Determine the conditional distribution of X, given $Y = 1$.
- Find the distribution of the random variable $Z = X + Y$.

Solution:

•

	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	
2	$\frac{1}{8}$	$\frac{1}{8}$	0	$\frac{1}{4}$
1	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$
0	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$
Y \ X	**0**	**1**	**2**	

•

k	**0**	**1**	**2**
$\mathbb{P}[Y = k]$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

•

k	**0**	**1**	**2**
$\mathbb{P}[X = k \mid Y = 1]$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

•

k	**0**	**1**	**2**	**3**	**4**
$\mathbb{P}[Z = k]$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{8}$

Problem 1.8. *Source: Sample P Exam, Problem #483.*
A doctor tests 100 patients for two diseases, **A** and **B**. Each patient has probability p of having disease **A** and probability p of having disease **B**, with $0 \leq p \leq \frac{1}{2}$.

For each patient, the event of having disease **A** and the event of having disease **B** are independent. The test outcomes for different patients are mutually independent.

The variance of the number of patients who have disease **A** is 9.

Calculate the variance of the number of patients who have at least one of the two diseases.

Solution: Let Y_A be the number of people who have disease **A** and let Y_B be the number of people who have disease **B**. We are given that they are both $b(100, p)$.

For disease **A**, we are also given

$$\mathrm{Var}[Y_A] = 100p(1-p) = 9 \quad \Rightarrow \quad p = 0.1$$

since we know that $0 \leq p \leq 0.5$. So, the probability that a single patient has neither of the two diseases is $0.9 \cdot 0.9 = 0.81$. Hence, the probability that a single patient has at lease one of the two diseases is $1 - 0.81 = 0.19$. The variance of the number of people who have at least one disease is

$$100(0.81)(0.19) = 15.39.$$

Problem 1.9. *Source: Sample P exam, Problem #29.*
An actuary has discovered that policyholders are three times as likely to file two claims as to file four claims. The number of claims filed has a Poisson distribution.

Calculate the variance of the number of claims filed.

Solution: Let N denote the number of claims. We are given that $N \sim P(\lambda)$. Also,

$$p_2 = 3p_4 \quad \Rightarrow \quad e^{-\lambda}\frac{\lambda^2}{2!} = 3e^{-\lambda}\frac{\lambda^4}{4!} \quad \Rightarrow \quad \lambda^2 = 4 \quad \Rightarrow \quad \lambda = 2.$$

We now know that $\mathrm{Var}[N] = \lambda = 2$.

Problem 1.10. *Source: Sample P exam, Problem #442.*
Let Y be a random variable that is uniform on $[a, b]$. The probability that Y is greater than 8 is 0.60. The probability that Y is greater than 11 is 0.20.

Calculate the variance of Y.

Solution: We are given the two probabilities that imply that the probability that Y lands between 8 and 11 is 0.4. So,

$$\frac{11-8}{b-a} = \frac{2}{5} \quad \Rightarrow \quad b - a = \frac{15}{2}.$$

Finally, the variance is

$$\frac{\left(\frac{15}{2}\right)^2}{12} = \frac{225}{48}.$$

Problem 1.11. A random variable Y has the normal distribution with mean 6 Its $0.8-$quantile is 8. What is its standard deviation?

Solution: Since $Y \sim N(\mu = 6, \sigma)$, we know that Y can be expressed as

$$Y = \mu + \sigma Z$$

where Z is standard normal. We are also given that

$$\mathbb{P}[Y \leq 8] = 0.8.$$

So,

$$\mathbb{P}[6 + Z\sigma \leq 8] = 0.8 \quad \Rightarrow \quad \mathbb{P}[Z\sigma \leq 2] = 0.8 \quad \Rightarrow \quad \mathbb{P}\left[Z \leq \tfrac{2}{\sigma}\right] = 0.8.$$

Using the standard normal tables, we see that

$$\frac{2}{\sigma} = 0.8416 \quad \Rightarrow \quad \frac{1}{\sigma} = 0.4208 \quad \Rightarrow \quad \sigma = \frac{1}{0.4208}\,.$$

Problem 1.12. *Source: Sample P exam, Problem #385.*
A computer manufacturer collects data on how long it takes before its computers fail. The time to fail, in years, follows an exponential distribution. Twenty percent of its computers fail within two years.

The probability a randomly selected computer fails before time t^*, in years, is 0.80.

Calculate t^*.

Solution: Let T be the exponential failure time, i.e., $T \sim E(\tau)$. We are given that

$$\mathbb{P}[T \le 2] = 0.2 \quad \Rightarrow \quad 1 - e^{-\frac{2}{\tau}} = 0.2 \quad \Rightarrow \quad e^{-\frac{2}{\tau}} = 0.8 \quad \Rightarrow \quad \tfrac{1}{\tau} = -\frac{1}{2}\ln(0.8) \quad \Rightarrow \quad \tau = -\frac{2}{\ln(0.8)}.$$

We are looking for t^* such that

$$\mathbb{P}[T \le t^*] = 0.8 \quad \Rightarrow \quad 1 - e^{-\frac{t^*}{\tau}} = 0.8 \quad \Rightarrow \quad e^{-\frac{t^*}{\tau}} = 0.2 \quad \Rightarrow \quad t^* = \frac{2\ln(0.2)}{\ln(0.8)}\,.$$

1.4. MULTIPLE CHOICE QUESTIONS.

Problem 1.13. (5 points) Let X and Y be two independent rolls of a die. Then,

$$\mathbb{P}[X - Y \ge 1 \text{ and } Y - X \le 3]$$

equals

(a) 1/6
(b) 1/3
(c) 1/2
(d) 2/3
(e) none of the above

Solution: The correct answer is **(e)**.
Not that if $X - Y \ge 1$ then $Y - X \le -1 \le 3$. Therefore, the probability above is the same as $\mathbb{P}[X - Y \ge 1]$. We need to count all pairs of (x, y) with $x, y \in \{1, 2, 3, 4, 5, 6\}$ such that $x \ge y + 1$ and add their probabilities (which are all equal to $1/36$). This corresponds to the blue boxed squares in the joint distribution table below:

	1	2	3	4	5	6
1	$\frac{1}{36}$	$\boxed{\frac{1}{36}}$	$\boxed{\frac{1}{36}}$	$\boxed{\frac{1}{36}}$	$\boxed{\frac{1}{36}}$	$\boxed{\frac{1}{36}}$
2	$\frac{1}{36}$	$\frac{1}{36}$	$\boxed{\frac{1}{36}}$	$\boxed{\frac{1}{36}}$	$\boxed{\frac{1}{36}}$	$\boxed{\frac{1}{36}}$
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\boxed{\frac{1}{36}}$	$\boxed{\frac{1}{36}}$	$\boxed{\frac{1}{36}}$
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\boxed{\frac{1}{36}}$	$\boxed{\frac{1}{36}}$
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\boxed{\frac{1}{36}}$
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$

There are 15 boxed squares, so the probability is 15/36.

Problem 1.14. (5 pts) A biased coin for which *Heads* is twice as likely as *Tails* is tossed twice, independently. If the outcomes are two consecutive *Heads*, Bertie gets a \$15 reward; if the outcomes are different, Bertie gets a \$10 reward; if the outcomes are two consecutive *Tails*, Bertie has to pay \$5. What is the expected value of Bertie's winnings?

(a) 75/9

(b) 80/9

(c) 85/9

(d) 95/9

(e) None of the above.

Solution: The correct answer is **(d)**.
Since *Heads* is twice as likely as *Tails*, *Heads* appears with probability 2/3, while *Tails* appears with probability 1/3.

Let X denote the amount Bertie wins. Then, X has the following distribution:

$$X \sim \begin{cases} 15, & \text{with probability } 4/9, \\ 10, & \text{with probability } 4/9, \\ -5, & \text{with probability } 1/9. \end{cases}$$

$$\mathbb{E}[X] = \frac{4}{9}(15) + \frac{4}{9}(10) + \frac{1}{9}(-5) = \frac{95}{9}\,.$$