University of Washington
Department of Mathematics

Practice Midterm Solutions

MATH 394: Probability I

Summer 2026

Instructor: Arman Jahangiri

Recommended time: 60 minutes
Total: 60 points

Problem 1. 10 points

A university student ID consists of three uppercase English letters followed by four digits. Letters and digits may repeat. The first letter cannot be `X`, and at least one of the four digits must be even.

How many such student IDs are possible? Justify your counting.

Solution

There are 25 choices for the first letter because it cannot be `X`. Each of the next two letters has 26 choices.

For the four digits, count all digit strings and subtract those containing only odd digits. There are 10^4 total digit strings and 5^4 strings in which all four digits are odd. Therefore the number of digit strings containing at least one even digit is

$$10^4 - 5^4.$$

By the multiplication principle, the required number is

$$25 \cdot 26^2 \big(10^4 - 5^4\big).$$

Since $10^4 - 5^4 = 9375$, this equals

$$25 \cdot 676 \cdot 9375 = 158{,}437{,}500.$$

Problem 2. **10 points**

A warehouse receives items from three factories. Factory A supplies 40% of the items and has a defect rate of 2%. Factory B supplies 35% of the items and has a defect rate of 5%. Factory C supplies 25% of the items and has a defect rate of 8%.

An item is selected uniformly at random from the warehouse and is found to be defective.

(a) Find the probability that the item came from Factory C. **(6 points)**

(b) Find the probability that the item did not come from Factory A. **(4 points)**

Solution

Let D be the event that the selected item is defective. The given probabilities are

$$P(A) = 0.40, \qquad P(B) = 0.35, \qquad P(C) = 0.25,$$

and

$$P(D \mid A) = 0.02, \qquad P(D \mid B) = 0.05, \qquad P(D \mid C) = 0.08.$$

By the law of total probability,

$$\begin{aligned} P(D) &= P(D \mid A)P(A) + P(D \mid B)P(B) + P(D \mid C)P(C) \\ &= (0.02)(0.40) + (0.05)(0.35) + (0.08)(0.25) \\ &= 0.008 + 0.0175 + 0.020 = 0.0455. \end{aligned}$$

(a) By Bayes' rule,

$$P(C \mid D) = \frac{P(D \mid C)P(C)}{P(D)} = \frac{(0.08)(0.25)}{0.0455} = \frac{40}{91}.$$

(b) Since the factories are mutually exclusive,

$$\begin{aligned} P(A^c \mid D) &= 1 - P(A \mid D) \\ &= 1 - \frac{P(D \mid A)P(A)}{P(D)} \\ &= 1 - \frac{0.008}{0.0455} = 1 - \frac{16}{91} = \frac{75}{91}. \end{aligned}$$

Problem 3. **10 points**

Let A, B, and C be events.

(a) Prove that

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

(6 points)

(b) Deduce that

$$P(A \cup B \cup C) \leq P(A) + P(B) + P(C).$$

(2 points)

(c) Give a simple necessary and sufficient condition for equality in part (b). **(2 points)**

Solution

(a) Apply the two-event addition rule to $(A \cup B)$ and C:

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

Now

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

and, by distributivity,

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C).$$

Applying the addition rule again,

$$\begin{aligned} P\big((A \cup B) \cap C\big) &= P\big((A \cap C) \cup (B \cap C)\big) \\ &= P(A \cap C) + P(B \cap C) - P(A \cap B \cap C). \end{aligned}$$

Substitution gives

$$\begin{aligned} P(A \cup B \cup C) = P(A) &+ P(B) + P(C) \\ &- P(A \cap B) - P(A \cap C) - P(B \cap C) \\ &+ P(A \cap B \cap C). \end{aligned}$$

(b) Either the proof provided in the lecture (much better), or this boring proof I have written below. (I feel like I used a tank for killing a mosquito.)
Pointwise, $\mathbf{1}_{A \cup B \cup C} \leq \mathbf{1}_A + \mathbf{1}_B + \mathbf{1}_C$. Reason: If $1_{A \cup B \cup C} = 1$, then at least one of the indicators on the right hand side is 1. So the left hand side is always less than or equal to the right hand side. Taking expectations of both sides, we obtain

$$\mathbb{E}[\mathbf{1}_{A \cup B \cup C}] \leq \mathbb{E}[\mathbf{1}_A + \mathbf{1}_B + \mathbf{1}_C]\,.$$

Using the fact that $\mathbb{E}[\mathbf{1}_E] = P(E)$ for any event E, together with the linearity of expectation, gives

$$P(A \cup B \cup C) \leq P(A) + P(B) + P(C).$$

(c) Can we shown using the same method of indicator functions, but I like the following better, which utilizes part (a):

(c) From part (a), equality in part (b) holds if and only if

$$P(A \cap B) + P(A \cap C) + P(B \cap C) = P(A \cap B \cap C).$$

We show that this is equivalent to

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = 0.$$

First, suppose equality holds in part (b). Then

$$P(A \cap B) + P(A \cap C) + P(B \cap C) = P(A \cap B \cap C).$$

Since

$$A \cap B \cap C \subseteq A \cap B, \qquad A \cap B \cap C \subseteq A \cap C, \qquad A \cap B \cap C \subseteq B \cap C,$$

we have

$$P(A \cap B) + P(A \cap C) + P(B \cap C) \geq 3P(A \cap B \cap C).$$

Therefore,

$$3P(A \cap B \cap C) \leq P(A \cap B \cap C),$$

which implies

$$P(A \cap B \cap C) = 0.$$

Consequently,

$$P(A \cap B) + P(A \cap C) + P(B \cap C) = 0.$$

Because all three terms are nonnegative, each one must be zero:

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = 0.$$

Conversely, suppose

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = 0.$$

Since $A \cap B \cap C \subseteq A \cap B$, we also have

$$P(A \cap B \cap C) = 0.$$

Thus, by the formula in part (a),

$$P(A \cup B \cup C) = P(A) + P(B) + P(C).$$

Hence, equality holds in part (b) if and only if A, B, and C are pairwise disjoint up to events of probability zero.

Problem 4. **15 points**

The cumulative distribution function of a continuous random variable X is

$$F_X(x) = \begin{cases} 0, & x < 0, \\ \dfrac{x^2}{4}, & 0 \le x < 1, \\ \dfrac{2x-1}{4}, & 1 \le x < 2, \\ 1 - \dfrac{(3-x)^2}{4}, & 2 \le x < 3, \\ 1, & x \ge 3. \end{cases}$$

(a) Find the probability density function $f_X(x)$. **(4 points)**

(b) Prove that $f_X(x)$ found in part (a) is a PDF. **(3 points)**

(c) Find $P\left(\frac{1}{2} < X \le \frac{5}{2}\right)$. **(3 points)**

(d) Find $\mathbb{E}[X]$. **(5 points)**

Solution

(a)
Since X is continuous, its density is obtained by differentiating the CDF at every point at which the derivative exists:

$$f_X(x) = F'_X(x).$$

Therefore,

$$f_X(x) = \begin{cases} 0, & x < 0, \\ \dfrac{x}{2}, & 0 < x < 1, \\ \dfrac{1}{2}, & 1 < x < 2, \\ \dfrac{3-x}{2}, & 2 < x < 3, \\ 0, & x > 3. \end{cases}$$

The values of a density at the individual points $x = 0, 1, 2, 3$ do not affect any probabilities. Thus, for example, we may define

$$f_X(x) = \begin{cases} \dfrac{x}{2}, & 0 \le x < 1, \\ \dfrac{1}{2}, & 1 \le x < 2, \\ \dfrac{3-x}{2}, & 2 \le x < 3, \\ 0, & \text{otherwise.} \end{cases}$$

(b)
To prove that f_X is a PDF, we must verify that

$$f_X(x) \ge 0 \qquad \text{for every } x,$$

and

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1.$$

First, $f_X(x) \geq 0$ on every interval in its definition. In particular, $x/2 \geq 0$ for $0 \leq x < 1$, $1/2 > 0$, and $(3-x)/2 \geq 0$ for $2 \leq x < 3$.
Next,

$$\begin{aligned}
\int_{-\infty}^{\infty} f_X(x)\,dx &= \int_0^1 \frac{x}{2}\,dx + \int_1^2 \frac{1}{2}\,dx + \int_2^3 \frac{3-x}{2}\,dx \\
&= \left[\frac{x^2}{4}\right]_0^1 + \left[\frac{x}{2}\right]_1^2 + \left[\frac{3x}{2} - \frac{x^2}{4}\right]_2^3 \\
&= \frac{1}{4} + \frac{1}{2} + \frac{1}{4} \\
&= 1.
\end{aligned}$$

Therefore, f_X is a valid probability density function.

(c)

Using the CDF,

$$P\left(\frac{1}{2} < X \leq \frac{5}{2}\right) = F_X\left(\frac{5}{2}\right) - F_X\left(\frac{1}{2}\right).$$

We have

$$\begin{aligned}
F_X\left(\frac{5}{2}\right) &= 1 - \frac{\left(3 - \frac{5}{2}\right)^2}{4} \\
&= 1 - \frac{\left(\frac{1}{2}\right)^2}{4} \\
&= 1 - \frac{1}{16} = \frac{15}{16},
\end{aligned}$$

and

$$F_X\left(\frac{1}{2}\right) = \frac{\left(\frac{1}{2}\right)^2}{4} = \frac{1}{16}.$$

Therefore,

$$P\left(\frac{1}{2} < X \leq \frac{5}{2}\right) = \frac{15}{16} - \frac{1}{16} = \frac{7}{8}.$$

(d)

By the definition of expectation for a continuous random variable,

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\,dx.$$

Thus,

$$\begin{aligned}
\mathbb{E}[X] &= \int_0^1 x\left(\frac{x}{2}\right)\,dx + \int_1^2 x\left(\frac{1}{2}\right)\,dx + \int_2^3 x\left(\frac{3-x}{2}\right)\,dx \\
&= \frac{1}{2}\int_0^1 x^2\,dx + \frac{1}{2}\int_1^2 x\,dx + \frac{1}{2}\int_2^3 (3x - x^2)\,dx.
\end{aligned}$$

Evaluating the three integrals,

$$\frac{1}{2}\int_0^1 x^2\,dx = \frac{1}{2}\left[\frac{x^3}{3}\right]_0^1 = \frac{1}{6},$$

$$\frac{1}{2}\int_1^2 x\,dx = \frac{1}{2}\left[\frac{x^2}{2}\right]_1^2 = \frac{3}{4},$$

and

$$\begin{aligned}\frac{1}{2}\int_2^3 (3x - x^2)\,dx &= \frac{1}{2}\left[\frac{3x^2}{2} - \frac{x^3}{3}\right]_2^3 \\ &= \frac{7}{12}.\end{aligned}$$

Therefore,

$$\begin{aligned}\mathbb{E}[X] &= \frac{1}{6} + \frac{3}{4} + \frac{7}{12} \\ &= \frac{2}{12} + \frac{9}{12} + \frac{7}{12} \\ &= \frac{3}{2}.\end{aligned}$$

‘‘‘

Problem 5. **15 points**

A city operates a large bike-share system. Assume that the outcomes of different rides are independent.

(a) On a typical day, 8000 rides are completed. Each ride results in a serious mechanical failure with probability 0.0005.

Let X be the number of serious mechanical failures during the day.

Determine whether a Poisson approximation or a Normal approximation to the distribution of X is more appropriate. Verify the relevant conditions and state the approximating distribution, including its parameter. **(3 points)**

(b) Use the approximation from part (a) to estimate the probability that at least 3 serious mechanical failures occur during the day. **(3 points)**

(c) At one busy station, 300 bikes are rented during the afternoon. Each bike is returned late with probability 0.35.

Let Y be the number of bikes returned late.

Determine whether a Poisson approximation or a Normal approximation to the distribution of Y is more appropriate. Verify the relevant conditions and state the approximating distribution, including its mean and variance. **(3 points)**

(d) Using the approximation from part (d), estimate

$$P(95 \leq Y \leq 115).$$

(3 points)

Solution

(a)
Each of the 8000 rides independently results in a serious mechanical failure with probability 0.0005. Therefore,

$$X \sim \text{Bin}(8000, 0.0005).$$

For the Normal approximation, we would need both

$$np \qquad \text{and} \qquad n(1-p)$$

to be sufficiently large. However,

$$np = 8000(0.0005) = 4,$$

which is less than 10. Thus, the usual Normal approximation is not appropriate.
For a Poisson approximation, n should be large and p should be small. Here,

$$n = 8000$$

is large and

$$p = 0.0005$$

is very small. Therefore, the Poisson approximation is appropriate.
Its parameter is

$$\lambda = np = 8000(0.0005) = 4.$$

Thus,

$$X \approx W, \qquad W \sim \text{Pois}(4).$$

(b)
We want

$$P(X \geq 3).$$

Using the Poisson approximation,

$$P(X \geq 3) \approx P(W \geq 3).$$

It is easier to use the complement:

$$P(W \geq 3) = 1 - P(W \leq 2).$$

Therefore,

$$\begin{aligned} P(W \geq 3) &= 1 - [P(W = 0) + P(W = 1) + P(W = 2)] \\ &= 1 - \left[e^{-4} + e^{-4}\frac{4}{1!} + e^{-4}\frac{4^2}{2!}\right] \\ &= 1 - e^{-4}(1 + 4 + 8) \\ &= 1 - 13e^{-4} \\ &\approx 0.7619. \end{aligned}$$

(c)
Each of the 300 bikes is independently returned late with probability 0.35. Therefore,

$$Y \sim \text{Bin}(300, 0.35).$$

A Poisson approximation is not appropriate because

$$p = 0.35$$

is not small.
For the Normal approximation, we compute

$$np = 300(0.35) = 105$$

and

$$n(1 - p) = 300(0.65) = 195.$$

Since both quantities are greater than 10, the Normal approximation is appropriate.
The mean is

$$\mu = np = 105,$$

and the variance is

$$\sigma^2 = np(1 - p) = 300(0.35)(0.65) = 68.25.$$

Thus,

$$Y \approx V, \qquad V \sim N(105, 68.25),$$

where

$$\sigma = \sqrt{68.25} \approx 8.26.$$

(d)

We want to approximate

$$P(95 \le Y \le 115).$$

Using the continuity correction,

$$P(95 \le Y \le 115) \approx P(94.5 < V < 115.5).$$

Standardizing,

$$\begin{aligned} P(94.5 < V < 115.5) &= P\left(\frac{94.5 - 105}{\sqrt{68.25}} < Z < \frac{115.5 - 105}{\sqrt{68.25}}\right) \\ &\approx P(-1.27 < Z < 1.27). \end{aligned}$$

Using the standard Normal table,

$$\Phi(1.27) = 0.8980.$$

By symmetry,

$$\Phi(-1.27) = 1 - \Phi(1.27) = 0.1020.$$

Therefore,

$$\begin{aligned} P(95 \le Y \le 115) &\approx \Phi(1.27) - \Phi(-1.27) \\ &= 0.8980 - 0.1020 \\ &= 0.7960. \end{aligned}$$