STAT 302: Assignment 2

Summer 2026, Due May 27th, 11:59pm on Gradescope.

Last modified: May 28, 2026 21:26

Question 1 [12 points]

A University is developing an AI-driven plagiarism detector. The detector returns positive (says a student plagiarized) 92% of the time for students who did plagiarize.

Assume 97% of the student body did not plagiarize on an assignment. The probability of testing negative for plagiarism across all students is 85.6%.

(a) [**1.5 points**] Let C be the event that a student plagiarized. Let T be the event that the detector returns positive (says a student plagiarized). Write the quantities that are given in this question as probabilities, in terms of these events.

(b) [**1.5 points**] What is the probability of testing negative, given that you did not plagiarize? Show all steps of this calculation.

(c) [**2 points**]A student is flagged by the detector. What is the probability that they actually plagiarized?

(d) [**3 points**]The University is considering a two-stage screening process. Students who test positive have their assessments run through the detector again. The results of the second test used to flag whether or not the student plagiarized. What is the probability that, even with this strategy, the detector is incorrect? Assume each the two tests are conditionally independent, given the plagiarism status. Use the notation T_1 and T_2 to denote the respective tests.

e) [**3 points**] Using the same two-stage screening process as in (d), what is the probability that a student didn't plagiarize, given they were flagged?

f) [**1 point**]Based on your answers for (d) and (e), would you be comfortable if your course implemented this two-stage screening process? Answer in 1 - 2 sentences.

Solution

(a) [1.5 points]

- $\mathbb{P}(T \mid C) = 0.92$ [0.5]
- $\mathbb{P}(C^c) = 0.97$ [0.5 points]
- $\mathbb{P}(T^c) = 0.856$ [0.5 points]

(b) [1.5 points]

We want to solve for $P(T \mid C)$.

$$\begin{aligned}
\mathbb{P}(T^C) &= [\mathbb{P}(T^c \mid C)\mathbb{P}(C) + \mathbb{P}(T^c \mid C^c)\mathbb{P}(C^c)] \\
0.856 &= [0.08 * 0.03 + \mathbb{P}(T^c \mid C^c)0.97] \\
\rightarrow \mathbb{P}(T^c \mid C^c) &= \frac{0.856 - 0.0024}{0.97} = 0.88
\end{aligned}$$

- 0.5 mark for identifying using the law of total probability
- 0.5 mark for correctly filling in the probabilities
- 0.5 for correct final answer

(c) [2 points]

Want to solve for

$$\begin{aligned}
\mathbb{P}(C \mid T) &= \frac{\mathbb{P}(T|C)\mathbb{P}(C)}{\mathbb{P}(T)} \\
&= \frac{0.92 \times 0.03}{0.144} \\
&= \frac{0.0276}{0.144} \\
&\approx 0.192
\end{aligned}$$

- 1 mark for formula
- 1 mark for plug and chug answer

(d) [3 points]

There are a *three* ways the detector can be incorrect. The first is that the student tests negative (only) once, but the student did plagiarize. The next two involve the first test showing up positive. A student can test on the first test, get retested, and test positive again, but the student did not plagiarize. Further, a student can test positive, then negative, but actually did cheat. We will add these probabilities up.

$$\begin{aligned}\mathbb{P}(\text{incorrect test}) = \mathbb{P}((&\text{Student did plagiarize and test 1 is negative,}\\ &\text{OR student did plagiarize and test 1 is positive but test 2 is negative, OR,}\\ &\text{student did not plagiarize and test 1 is positive and test 2 is positive. }))\\ &= \mathbb{P}(T_1^c \mid C)\mathbb{P}(C) + \mathbb{P}(T_1, T_2^c \mid C)\mathbb{P}(C) + \mathbb{P}(T_1, T_2 \mid C^c)P(C^c)\\ &= 0.08(0.03) + \mathbb{P}(T_1, T_2^c \mid C)(0.03) + \mathbb{P}(T_1, T_2 \mid C^c)(0.97))\end{aligned}$$

$\mathbb{P}(T_1, T_2 \mid C^c) = \mathbb{P}(T_1 \mid C^c) \times \mathbb{P}(T_2 \mid C^c) = 0.12^2 = 0.0144$

$\mathbb{P}(T_1, T_2^c \mid C) = \mathbb{P}(T_1 \mid C) \times \mathbb{P}(T_2^c \mid C) = (0.92)(1 - 0.92) = 0.074$

We know $\mathbb{P}(T^c|C^c) = 0.88$, so $\mathbb{P}(T|C^c) = 1 - 0.88 = 0.12$ for any test

Therefore,

$$\begin{aligned}\mathbb{P}(\text{incorrect test}) &= 0.08(0.03) + \mathbb{P}(T_1, T_2^c \mid C)(0.03) + \mathbb{P}(T_1, T_2 \mid C^c)(0.97))\\ &= 0.08(0.03) + 0.074(0.03) + 0.0144(0.97)\\ &= 0.0186\end{aligned}$$

- 1 point for identifying the three ways to be incorrect
- 0.5 point for $\mathbb{P}(T_1, T_2 \mid C^c) = 0.12 * 0.12$
- 0.5 point for $\mathbb{P}(T_1, T_2^c \mid C) = (0.92)(1 - 0.92)$
- 1 point for final answer

e) [3 points]

Want to solve for $\mathbb{P}(C^c \mid T_1, T_2)$. We can solve this using Bayes' Rule:

$$\begin{aligned}\mathbb{P}(C^c \mid T_1, T_2) &= \frac{\mathbb{P}(T_1, T_2|C^c)\mathbb{P}(C^c)}{P(T_1, T_2)}\\ &= \frac{\mathbb{P}(T_1, T_2|C^c)\mathbb{P}(C^c)}{\mathbb{P}(T_1, T_2|C)\mathbb{P}(C) + \mathbb{P}(T_1, T_2|C^c)\mathbb{P}(C^c)}\\ &= \frac{0.0144(0.97)}{(0.92)^2(0.03) + (0.0144)(0.97)}\\ &\approx 0.355\end{aligned}$$

- 1 point for numerator
- 1 point for each term in the denominator
- 1 point for final answer

(f) [1 point]

No. Any short answer will receive credit.

Question 2 [9 points]

A student union runs a raffle to raise money for campus events. Each ticket costs **$3** to enter. There are **500 tickets** sold in total, and the prizes are awarded as follows:

Prize	Number of Winners
$150 cash	1
$30 gift card	4
$3 (money back)	20

(a) **[1 point]** Let Y be the dollar value of the prize. Define (mathematically) an appropriate random variable X to represent a student's **net gain** from buying one ticket.

(b) **[2 points]** Describe each event in X in words. Write the set Ω.

(c) **[3 points]** Write out the probability mass function (PMF) of X for any $X \in \mathbb{R}$.

(d) **[2 points]** What is the probability that that the student will have a positive net gain?

Solution

(a) [1 point]

Let X = the net gain (in dollars) to a student who buys one ticket, $X = Y - 3$ (value minus ticket cost)

(b) [2 points]

There are four possible outcomes:

- X = 147 (student wins /$150 cash)
- X = 27 (student wins /$30 gift card)
- X = 0 (student wins free ticket)
- X = -3 (student wins nothing)

$\Omega = \{147, 27, 0, -3\}$

- 0.25 for each correct element
- 0.25 for each description
- 1 for writing out Ω as a set.

(c) [3 points]

$$\mathbb{P}(X = x) = \begin{cases} \frac{1}{500} & x = 147 \\ \frac{4}{500} & x = 27 \\ \frac{20}{500} & x = 0 \\ \frac{475}{500} & x = -3 \\ 0 & \text{otherwise} \end{cases}$$

* 0.5 points for each probability where $x \neq 0$ * 1 point for listing "0 otherwise"

(d) [2 points]

$$\begin{aligned} \mathbb{P}(X > 0) &= P(X = 27 \cap X = 147) \\ &= \mathbb{P}(X = 27) + \mathbb{P}(X = 147) \\ &= \frac{4}{500} + \frac{1}{500} \\ &= 0.01 \end{aligned}$$

- 1 point for answer
- 1 point for showing work

Question 3 [11 points]

(a) **[5 points]** Let $X \sim Binom(12, \theta)$. What value of θ maximizes $\mathbb{P}(X = 11)$?

(b) **[2 points]** A mystery novel book club has 20 members. 8 of them have secretly already read this month's selection, and know the ending. The organizer randomly calls on 6 members during the meeting to share their predictions how the novel ends. What is the probability that no more than 2 of the 6 people know how the story ends?

(c) **[4 points]** A barista makes 12 vanilla lattes during a rush. For each drink, she independently has a 30% chance of forgetting to add the vanilla syrup. The shift manager will only intervene if 3 or more drinks are missing the syrup — otherwise he assumes it's just customer preference. A regular customer arrives and orders one of the 12 drinks at random after they've all been made. What is the probability that the manager does not intervene, but the regular customer's drink is still missing the syrup?

Solutions

(a) [5 points]

$P(X = 11) = \binom{12}{11}\theta^{11}(1-\theta)^1 = 12\theta^{11}(1-\theta)$

To find the inflection point(s), we can take the derivative and set it equal to zero.

$$\begin{aligned}\frac{d}{d\theta}12\theta^{11}(1-\theta) &= \left[\frac{d}{d\theta}12\theta^{11}\right](1-\theta) + 12\theta^{11}\left[\frac{d}{d\theta}(1-\theta)\right]\\ &= (12)(11)\theta^{10}(1-\theta) + 12\theta^{11}(-1)\end{aligned}$$

Then,

$$\begin{aligned}&(12)(11)\theta^{10}(1-\theta) + 12\theta^{11}(-1) = 0\\ \implies &11\theta^{10}(1-\theta) - 12\theta^{11} = 0\\ \implies &11\theta^{10}(1-\theta) = 12\theta^{11}\end{aligned}$$

This will be zero when $\theta = 0$. We can also consider the case of $\theta \neq 0$:

$$\begin{aligned}&\implies 11(1-\theta) = \theta \quad \text{if } \theta \neq 0 \text{ can divide by } \theta^{10}\\ &\implies 11 - 11\theta = \theta\\ &\implies 11 = 12\theta\\ &\implies \theta = 11/12\end{aligned}$$

We need to check that this is a maximum using second derivative test:

$$\begin{aligned}\frac{d^2}{d\theta^2}12\theta^{11}(1-\theta) &= \frac{d}{d\theta}\left[(12)(11)\theta^{10}(1-\theta) - 12\theta^{11}\right]\\ &= \frac{d}{d\theta}\left[132\theta^{10} - 132\theta^{11} - 12\theta^{11}\right]\\ &= 132(10)\theta^9 - 132(11)\theta^{10} - 12(11)\theta^{10}\end{aligned}$$

We can plug the critical points into here and see if it is negative. If it is, then it's a maximum:

The second derivative evaluated at $\theta = 0$ is 0, indicating it is a saddle point or local max/min, (not a maximum).

The second derivative evaluated at $\theta = 11/12$ is:

$$132(10)(11/12)^9 - 132(11)(11/12)^{10} - 12(11)(11/12)^{10} \approx -60.32.$$

which indicates that this point is a maximum.

Grading:

- 1 point for setting derivative to zero
- 0.5 for correct derivative
- 0.5 for finding critical value of 0
- 0.5 for finding critical value of 11/12
- 0.5 point for taking second derivative

- 0.5 for evaluating second derivative at $\theta = 11/12$
- 0.5 for explicitly saying this is a maximum
- 0.5 for evaluating second derivative at $\theta = 0$
- 0.5 for explicitly saying this is NOT a maximum

(b) [2 points]

This distribution is the Hypergeometric distribution with $N = 20$, $K = 8$, $n = 6$.

$$\begin{aligned}
\mathbb{P}(X \leq 2) &= \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) \quad \text{[1 point]} \\
&= \frac{\binom{8}{0}\binom{12}{6}}{\binom{20}{6}} + \frac{\binom{8}{1}\binom{12}{5}}{\binom{20}{6}} \frac{\binom{8}{2}\binom{12}{4}}{\binom{20}{6}} \\
&= 0.0238 + 0.1634 + 0.3581 \\
&= 0.5453
\end{aligned}$$

- 0.5 points for correctly identifying distribution
- 0.5 points for correctly identifying parameters
- 1 point for correct final answer

(c) [4 points]

To solve this, we need two things to be true: at most two drinks are missing syrup AND the customer's drink is one of those that's missing the syrup. These events are not independent.

Let $X \sim Binom(n = 12, \theta = 0.3)$ be the number of drinks missing syrup.

The probability that the customer chooses a drink with missing syrup is the probability that the drink is missing syrup (which depends on the number of drinks with missing syrup) times the probability of choosing a missing syrup drink (which is dependent on the number of drinks with missing syrup).

We can write this as $\mathbb{P}(\text{customer chooses drink with no syrup}) = \mathbb{P}(X = x) * \frac{x}{12}$

Then,

$$\begin{aligned}
\mathbb{P}(\text{manager silent and customer unlucky}) &= \sum_{k=0}^{2} \mathbb{P}(X = k)\frac{k}{12} \\
&= \mathbb{P}(X = 0)\frac{0}{12} + \mathbb{P}(X = 1)\frac{1}{12} + \mathbb{P}(X = 2)\frac{2}{12} \\
&= 0 + \binom{12}{1}(0.3)^1(1 - 0.3)^{12-1}\frac{1}{12} + \binom{12}{2}(0.3)^2(1 - 0.3)^{12-2}\frac{2}{12} \\
&\approx 0.034
\end{aligned}$$

- No marks for answer of $0.30 * \mathbb{P}(X \leq 2)$

- 1 mark for getting the weighted probability equation
- 1 mark for identifying parameters of the binomial (12, 0.3)
- 1 mark for adding X = 1, and 2 probabilities (OK if 0 is omitted)
- 1 mark for final answer