DAEN 427 / ISEN 427 / ISEN 627 - Fall 2026

Exam 1

"An Aggie does not lie, cheat or steal, or tolerate those who do."

By signing below, I state that I will follow the Aggie Code of Honor on this exam.

Name/UIN: _____ Signature: _____

Instructions: This is a closed-book exam. You may use pen or pencil, a basic or scientific calculator, and one regular sticky note (standard 3" x 3" *Post-it* or equivalent). Leave the sticky note attached to your exam when you turn it in. No other materials or devices (smartphones, tables, computers, smart-glasses, etc.) are permitted.

Four questions ... DAEN/ISEN 427: 10 points + 1 extra-credit point; ISEN 627: 11 points total.

Good luck! ツ

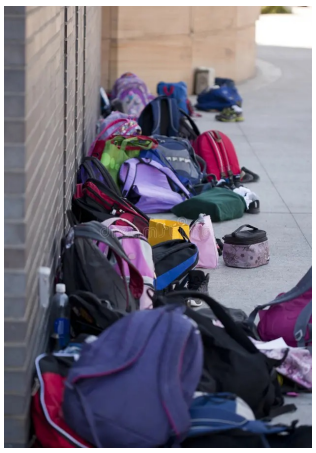
Formulas:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)} \qquad P(L_i \mid R) = \frac{P(R \mid L_i)P(L_i)}{\sum_j P(R \mid L_j)P(L_j)}$$

$$\theta \sim \text{Beta}(\alpha, \beta) \;\Rightarrow\; \theta \mid D \sim \text{Beta}(\alpha + x, \beta + n - x) \quad E(\theta \mid D) = \frac{\alpha'}{\alpha' + \beta'} \quad P(\theta > c \mid D) = 1 - F_{\text{Beta}(\alpha', \beta')}(c)$$

$$\mathcal{R}(a_+) = C_{\text{FP}}(1 - q) \qquad \mathcal{R}(a_-) = C_{\text{FN}} q \qquad \tau = \frac{C_{\text{FP}}}{C_{\text{FP}} + C_{\text{FN}}}$$

Q1. (2 points, 10 min.)



Suppose you are taking a large class with 64 other students. For a class activity, everyone leaves one backpack in a designated storage area before entering the room.

At the end of class, the backpacks are returned one at a time. Your professor reminds everyone: *"A lot of backpacks look alike. Make sure you check carefully before taking one!"*

Good advice: your backpack is a very popular black, medium-sized model used by 15% of students.

From where you are standing, you can see only a backpack's color, size, and general shape. You cannot see its name tag or any other identifying marks until it gets closer.

a. The very first backpack to be returned is black, medium-sized, and looks exactly like your model. What are the chances that it is actually yours? Calculate the posterior probability that the first backpack belongs to you.

b. Fast-forward ... You have now checked the first 63 backpacks and confirmed that none of them is yours. Only two backpacks remain. The 64th backpack appears and, once again, it looks exactly like yours. Now what are the chances that it is yours? Calculate the posterior probability that the 64th backpack belongs to you.

Answer:

Let Y mean "this is my backpack" and M mean "it looks like my backpack." Your own backpack must match, so $P(M \mid Y) = 1$. Assume an unrelated backpack matches with probability $P(M \mid Y^c) = 0.15$. A

visual match is evidence, but it does not identify the owner with certainty.

a. Before seeing the first backpack, each of the 65 backpacks is equally likely to be yours: $P(Y) = 1/65$. A match can arise either because it is yours or because one of the other 64 backpacks happens to look alike. Bayes' rule gives

$$P(Y \mid M) = \frac{(1)(1/65)}{(1)(1/65) + (0.15)(64/65)} = \frac{1}{10.6} \approx 0.0943.$$

Thus the chance is about 9.43%. Equivalently, there is one guaranteed matching candidate (yours) and $64(0.15) = 9.6$ expected matching candidates among the others. Credits for the correct prior, likelihood, and normalization!

b. Confirming that the first 63 backpacks are not yours changes the candidate set: your backpack is one of the two remaining, with prior probability 1/2 for either position. The 64th backpack's visual match has likelihood 1 if it is yours and 0.15 if it belongs to the other student. Therefore

$$P(Y \mid M, 63 \text{ ruled out}) = \frac{(1)(1/2)}{(1)(1/2) + (0.15)(1/2)} = \frac{1}{1.15} \approx 0.8696.$$

Thus the chance is about 86.96%. The higher probability comes from eliminating 63 possible owners, not from the black backpack becoming more distinctive. Again, credits for the correct updated prior and posterior.

Since 15% of 65 is not an integer, these numerical answers treat 0.15 as the probability that another backpack visually matches. Accept a consistent counting if a fixed number of matching backpacks is explicitly assumed.

Q2. (4 points, 10 min.)

You have lost your phone somewhere in your home. Based on past experience, the phone is most likely to be in one of four locations: the Bedroom, Living Room, Kitchen, or Office. You use the *"Ring My Phone"* feature and observe whether you can hear the phone ringing.

a. Draw a Directed Acyclic Graph (DAG) for this problem using the following variables:

- historical tendency to lose the phone in different locations;
- location of the phone (L); and
- whether you hear the phone ringing (R).

Clearly show the direction of the dependencies between the variables.

b. Explain the Bayesian model represented by your DAG. Clearly identify:

- what determines the prior distribution of the location;
- what represents the likelihood of the location given that you hear the phone ringing; and
- what represents the posterior distribution after observing whether the phone can be heard.

c. Suppose the prior probabilities for the phone's location are

$$P(L) = \begin{cases} 0.40 & L = \text{Bedroom}, \\ 0.30 & L = \text{Living Room}, \\ 0.20 & L = \text{Kitchen}, \\ 0.10 & L = \text{Office}. \end{cases}$$

The probability of hearing the phone ring is

$$P(R=1 \mid L)=\begin{cases}0.30 & L=\text{Bedroom},\\ 0.90 & L=\text{Living Room},\\ 0.60 & L=\text{Kitchen},\\ 0.20 & L=\text{Office}.\end{cases}$$

You activate *"Ring My Phone"* and hear the phone, so $R=1$.
Using Bayes' rule,

$$P(L \mid R=1)=\frac{P(R=1 \mid L)P(L)}{\sum_{\ell} P(R=1 \mid L=\ell)P(L=\ell)},$$

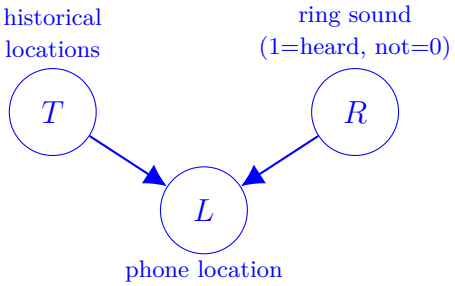
calculate $P(L \mid R=1)$ for each of the four possible locations.

d. Based on the posterior probabilities:

- Where should you search for the phone first?
- Explain how hearing the phone changed your belief about its location compared with the prior probabilities.

Answer:

a. Let T denote historical locations (the past location frequencies), R the observed ring sound, and L the phone's unknown location. The arrows below show the two inputs used to infer L:



The $T \to L$ arrow represents prior information; $R \to L$ represents updating the inferred location after hearing (or not hearing) the ring. Credit the three labeled circular nodes/variables and both arrows pointing into L. Note that the arrangement of the network can be slightly different.

b. This diagram displays *information flow toward the inference*, $P(L \mid T, R)$: past locations supply the prior and the ring observation updates it. The probability model used in the calculation still has $P(L, R \mid T) = P(L \mid T)P(R \mid L)$, because the actual location affects audibility. Historical tendencies determine the *prior* $P(L \mid T)$ (written $P(L)$ when T is fixed). The *likelihood* of an observed ring outcome r for each candidate location is $P(R=r \mid L)$. Conditioning on the observed outcome produces the *posterior*:

$$P(L=\ell \mid R=r, T)=\frac{P(R=r \mid L=\ell)P(L=\ell \mid T)}{\sum_{j} P(R=r \mid L=j)P(L=j \mid T)}.$$

For a heard ring, use $r=1$; for an unheard ring, use $r=0$ and $P(R=0 \mid L)=1-P(R=1 \mid L)$.

c. Multiply each location's prior by its chance of producing the observation $R=1$. In Bedroom, Living Room, Kitchen, Office order, the unnormalized weights are $0.40(0.30)=0.12$, $0.30(0.90)=0.27$, $0.20(0.60)=0.12$, and $0.10(0.20)=0.02$. Their sum, 0.53, is the overall probability $P(R=1)$. Divide

every weight by 0.53 so the posterior probabilities sum to one:

L	Bedroom	Living Room	Kitchen	Office
$P(L \mid R = 1)$	0.2264	0.5094	0.2264	0.0377

Credit the products, denominator, and four posterior values (rounding accepted).

d. Search the Living Room first; its posterior 0.5094 is highest. Before ringing, the Bedroom was most likely (0.40). Hearing a ring favors locations where a ring is easier to hear: the Living Room rises from 0.30 to 0.5094, and the Kitchen from 0.20 to 0.2264. The Bedroom falls from 0.40 to 0.2264, and the Office from 0.10 to 0.0377.

Q3. (2 points, 10 min.)

A football defense models the probability θ that an opposing offense calls a deep pass in a particular down-and-distance situation. The prior is

$$\theta \sim \text{Beta}(3, 12).$$

Game analysis finds 9 deep-pass attempts among 21 comparable plays. The defensive coordinator defines the high-risk state as $H = 1$ when $\theta > 0.40$. The defense may use a conservative two-high safety coverage (a_+) or stay with its standard coverage (a_-).

a. Find the posterior distribution of θ and compute $q = P(\theta > 0.40 \mid D)$ (select the correct value from the R outputs below).

b. Suppose the cost of using two-high coverage when the high-risk state is absent is $C_{\text{FP}} = 3$ expected yards, while the cost of staying in standard coverage when the high-risk state is present is $C_{\text{FN}} = 10$ expected yards. Find the threshold, choose the Bayes action, and compare expected losses.

From R, inputs in blue, outputs in red:

```
> 1 - pbeta(0.40, shape1 = 12, shape2 = 24, lower.tail = FALSE) → [1] 0.80483
> 1 - pbeta(0.40, shape1 = 3, shape2 = 15, lower.tail = TRUE) → [1] 0.01232
> 1 - pbeta(0.40, shape1 = 12, shape2 = 24, lower.tail = TRUE) → [1] 0.19518
> 1 - pbeta(0.30, shape1 = 12, shape2 = 24, lower.tail = TRUE) → [1] 0.65156
> 1 - pbeta(0.70, shape1 = 12, shape2 = 24, lower.tail = TRUE) → [1] 2.8536e-06
```

Answer:

a. The Beta prior contributes shape parameters $(3, 12)$. The 21 plays contribute 9 deep passes and $21-9 = 12$ other plays, hence

$$\theta \mid D \sim \text{Beta}(3 + 9, 12 + 12) = \text{Beta}(12, 24).$$

The high-risk event is a statement about the unknown rate θ, so integrate the posterior above 0.40:

$$q = P(\theta > 0.40 \mid D) = 1 - F_{\text{Beta}(12,24)}(0.40) \approx 0.19518.$$

The R output with `lower.tail = TRUE` → 0.19518. In R, `pbeta(..., lower.tail = TRUE)` is the cumulative probability $P(\theta \leq 0.40 \mid D)$; subtracting it from one gives the upper tail. The 0.80483 line subtracts an upper-tail result from one and therefore returns the lower-tail probability. I expect some people will make this mistake, need to take points out!

b. Two-high coverage a_+ incurs its 3-yard false-positive cost only when $H = 0$, which has probability $1 - q$. Standard coverage a_- incurs its 10-yard false-negative cost only when $H = 1$, which has probability q.

Thus

$$\mathcal{R}(a_+) = 3(1-q) \qquad \text{and} \qquad \mathcal{R}(a_-) = 10q.$$

Comparing them, choose a_+ if

$$q > \frac{C_{\text{FP}}}{C_{\text{FP}} + C_{\text{FN}}} = \frac{3}{13} \approx 0.23077,$$

with a tie at equality. Since $0.19518 < 0.23077$, choose **standard coverage** a_-. Numerically,

$$\mathcal{R}(a_+) \approx 3(0.80482) = 2.41446 \text{ yards and } \mathcal{R}(a_-) \approx 10(0.19518) = 1.9518 \text{ yards.}$$

Standard coverage has the smaller posterior expected loss. Credit threshold, action, and losses.

Q4. (2 pts + 1 extra pt, 10 min.)

A student models the probability θ that their bus will be more than 10 minutes late on a weekday morning. Before tracking this route,

$$\theta \sim \text{Beta}(4, 12).$$

On 16 comparable mornings, the bus was more than 10 minutes late 4 times. Tomorrow, the student can take a rideshare (a_+) or take the bus as usual (a_-).

a. Find the posterior predictive probability that the bus will be more than 10 minutes late tomorrow.
b. Suppose taking a rideshare when the bus would have been on time costs $C_{\text{FP}} = \$3.00$, while taking the bus when it is late costs $C_{\text{FN}} = \$12.00$ (for example, you will arrive late for your shift and won't get paid for the time, or you will have to make up the hours). Assume all other losses are zero. Find the action threshold and the Bayes action. Verify the decision using expected losses.
c. **(Extra point)** Find the critical false-negative cost above which taking a rideshare becomes optimal, holding $C_{\text{FP}} = \$3.00$ fixed.

Answer:

a. Four late mornings are successes and the other 12 are failures. Updating the Beta prior gives

$$\theta \mid D \sim \text{Beta}(4+4, 12+12) = \text{Beta}(8, 24).$$

For one new comparable morning, average the conditional late probability θ over its posterior. Thus the posterior predictive probability is

$$p = P(\text{late tomorrow} \mid D) = E(\theta \mid D) = \frac{8}{8+24} = 0.25.$$

b. A rideshare incurs the \$3 unnecessary-ride cost only if the bus would be on time (probability $1-p$). The bus incurs the \$12 lateness cost only if it is late (probability p). The expected losses are therefore $\mathcal{R}(a_+) = 3(1-p)$ and $\mathcal{R}(a_-) = 12p$ dollars. Comparing $3(1-p) < 12p$ gives the rideshare threshold $p > \frac{3}{3+12} = 0.20$; at equality either action is optimal. Since $p = 0.25$, choose rideshare. Its expected loss is $3(0.75) = \$2.25$, compared with $12(0.25) = \$3.00$ for the bus.
c. Hold $p = 0.25$ and $C_{\text{FP}} = \$3$ fixed and solve for the false-negative cost that makes the actions equally costly: $3(1-0.25) = C_{\text{FN}}(0.25)$. Hence $C^*_{\text{FN}} = \$9.00$. Rideshare is strictly optimal above \$9.00; at \$9.00 both actions tie, and below \$9.00 the bus is optimal.