CS 70 Discrete Mathematics and Probability Theory
Summer 2025 Tate

HW 5

Due: Saturday, 8/2, 4:00 PM
Grace period until Saturday, 8/2, 6:00 PM
Remember to show your work for all problems!

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.) If you used an LLM, place transcripts of your chats here.

1 Sumanth and the High E Note

Note 15
Note 19

Every morning, Sumanth practices his flute. On any given day, he has a fixed probability p of successfully hitting the high E note. Each day's attempt is independent of the others.

(a) Let X be the number of days until he hits the high E note for the first time. What is $\mathbb{E}[X]$? (You may cite a known result.)

(b) Now suppose Sumanth vows to quit flute forever if he doesn't get it within the first k days. What is the expected number of days he will practice in this case?

(c) Suppose Sumanth plays two different flutes, one in the morning and one in the evening. Each gives him an independent chance p of hitting the high E. What is the expected number of days until he hits the note on either flute?

(d) Now suppose Sumanth plays n different flutes each day, each giving him an independent probability p of hitting the high E note. What is the expected number of days until he hits it on at least one flute?

2 Ritwik's NVIDIA Gauntlet

Note 15
Note 19

Ritwik is applying to an NVIDIA internship. There are three interviews: Coding (C), Systems (S), and HR (H). Define the following random variables with possible values $\{0,1\}$:

- $C = 1$ if he passes coding, with $\mathbb{P}[C = 1] = 0.8$.
- $S = 1$ if he passes systems. Given $C = 1$, $\mathbb{P}[S = 1] = 0.6$; given $C = 0$, $\mathbb{P}[S = 1] = 0.1$.
- $H = 1$ if he passes HR. Given $S = 1$, $\mathbb{P}[H = 1] = 0.9$; given $S = 0$, $\mathbb{P}[H = 1] = 0.3$.

Define random variables $X = C + S + H$ and $Y = 2C + 3S + H$.

(a) Compute the joint distribution table for (C, S) and (S, H).

(b) Compute $\mathbb{E}[X]$.

(c) Which is more likely: Ritwik passes all three rounds, or only Systems and HR (i.e., $C = 0, S = 1, H = 1$)? Justify.

(d) Compute $\mathbb{E}[Y]$.

3 Anikait's Completion Saga

Note 16
Note 19

Anikait, a huge anime fan recovering from a basketball injury, has been spending his recovery time collecting limited-edition anime figurines. Each booster pack contains **one** random figurine, and there are n different figurines in total.

Let's define the following scenario:

Anikait opens booster packs one at a time:

- The probability that a pack contains a figurine he **hasn't** collected yet is proportional to the number of figurines he hasn't seen yet.
- Let T_i represent the number of packs he needs to open to get the i^{th} new figurine.
- Let $S_n = \sum_{i=1}^{n} T_i$ be the total number of packs needed to collect all n figurines.
- You may use the following approximation for harmonic numbers: $\sum_{j=a}^{b} \frac{1}{j} \approx \ln b - \ln a$. In particular, the n-th harmonic number is defined as $H_n = \sum_{j=1}^{n} \frac{1}{j} \approx \ln n + \gamma$, where γ is the Euler–Mascheroni constant (you may ignore γ for approximation purposes).

(a) Show that $\mathbb{E}[T_i] = \frac{n}{n-i+1}$ (you should identify the type of distribution T_i has).

(b) Compute $\mathbb{E}[S_n]$.

(c) Compute $\text{Var}(S_n)$.

(d) Suppose Anikait is only aiming to collect k distinct figurines instead of all n. Find $\mathbb{E}[S_k]$ and explain how it scales with k and n.

(e) Suppose Anikait collects figurines his friend Pranav, and both open one pack per hour independently until either of them completes the full set of n figurines. Let $X \sim S_n$ be the number of packs Anikait needs, and let $Y \sim S_n$ be the number of packs Pranav needs. Compute or bound $\mathbb{E}[\min(X, Y)]$ using symmetry.

Hint 1: Recall the identity for any two random variables X, Y:

$$\min(X, Y) = \frac{X + Y - |X - Y|}{2}.$$

This implies:

$$\mathbb{E}[\min(X, Y)] = \mu - \frac{1}{2}\mathbb{E}[|X - Y|],$$

where $\mu = \mathbb{E}[S_n] = nH_n$.

Hint 2: Use the Cauchy-Schwarz inequality to bound $\mathbb{E}[|X-Y|]$. For any random variables A, B,

$$\mathbb{E}[|AB|] \leq \sqrt{\mathbb{E}[A^2] \cdot \mathbb{E}[B^2]}.$$

4 Double-Check Your Intuition Again

Note 16

(a) You roll a fair six-sided die and record the result X. You roll the die again and record the result Y.

  (i) What is $\text{cov}(X+Y, X-Y)$?

  (ii) Prove that $X+Y$ and $X-Y$ are not independent.

For each of the problems below, if you think the answer is "yes" then provide a proof. If you think the answer is "no", then provide a counterexample.

(b) If X is a random variable and $\text{Var}(X) = 0$, then must X be a constant?

(c) If X is a random variable and c is a constant, then is $\text{Var}(cX) = c\,\text{Var}(X)$?

(d) If A and B are random variables with nonzero standard deviations and $\text{Corr}(A,B) = 0$, then are A and B independent?

(e) If X and Y are not necessarily independent random variables, but $\text{Corr}(X,Y) = 0$, and X and Y have nonzero standard deviations, then is $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$?

(f) If X and Y are random variables then is $\mathbb{E}[\max(X,Y)\min(X,Y)] = \mathbb{E}[XY]$?

(g) If X and Y are independent random variables with nonzero standard deviations, then is

$$\text{Corr}(\max(X,Y), \min(X,Y)) = \text{Corr}(X,Y)?$$

5 Achyut's Notre Dame Drone Challenge

Note 20

Achyut, a massive architect fan with a deep love for Notre Dame, is testing a new drone that scores points by successfully navigating through flying arches on a miniature replica of the cathedral.

Each **successful pass** earns a score drawn independently from a random variable X_i, where:

- $\mathbb{P}(X_i = 1) = 0.5$
- $\mathbb{P}(X_i = 2) = 0.3$
- $\mathbb{P}(X_i = 3) = 0.2$

He continues attempting passes until he accumulates at least 15 points. Let:

- N be the number of successful passes until the total score $S = \sum_{i=1}^{N} X_i \geq 15$
- The scores X_i are i.i.d., and independent of the stopping time N

(a) What is the expected score per successful pass, $\mathbb{E}[X_i]$?

(b) Using the approximation $\mathbb{E}[S] \approx \mathbb{E}[N] \cdot \mathbb{E}[X_i]$, estimate how many passes Achyut needs on average to reach 15 points.

(c) Let's define the MMSE (Minimum Mean Squared Error) estimator for S given $N = n$. Compute $\hat{S} = \mathbb{E}[S \mid N = n]$.

(d) Let T be the number of flights where Achyut scores exactly 3 points. Suppose n successful passes were made to reach total score s. Compute $\mathbb{E}[T \mid S = s]$ leaving your answer in the form n*(some conditional probability).

6 Balls in Bins Estimation

Note 20

We throw $n > 0$ balls into $m \geq 2$ bins. Let X and Y represent the number of balls that land in bin 1 and 2 respectively.

(a) Calculate $\mathbb{E}[Y \mid X]$. [*Hint*: Your intuition may be more useful than formal calculations.]

(b) What is $L[Y \mid X]$ (where $L[Y \mid X]$ is the best linear estimator of Y given X)? [*Hint*: Your justification should be no more than two or three sentences, no calculations necessary! Think carefully about the meaning of the conditional expectation.]

(c) Unfortunately, your friend is not convinced by your answer to the previous part. Compute $\mathbb{E}[X]$ and $\mathbb{E}[Y]$.

(d) Compute $\text{Var}(X)$.

(e) Compute $\text{cov}(X, Y)$.

(f) Compute $L[Y \mid X]$ using the formula. Ensure that your answer is the same as your answer to part (b).