

CS109 Midterm Exam (World Cup Edition)

This is a closed calculator/computer/phone/smart-watch/smart-toothbrush exam. You are, however, allowed to use 3 pages (front and back) of handwritten notes in the exam. You have 75 minutes to take the exam. The exam is 65 points. You may want to use the point allocation for each problem as a rough indicator for pacing yourself on the exam.

In the event of an incorrect answer, any explanation you provide of how you obtained your answer can potentially allow us to give you partial credit for a problem. For example, describe the distributions and parameter values you used, where appropriate. It is fine for your answers to include summations, products, factorials, exponentials, and combinations. You can leave your answer in terms of Φ (the CDF of the standard normal) or Φ^{-1} . For example $\Phi(3/4)$ is an acceptable final answer. You can only leave your answer in **terms of code** when the problem explicitly says so.

Stanford Honor Code: The Honor Code is an undertaking of the Stanford academic community, individually and collectively. Its purpose is to uphold a culture of academic honesty. Students will support this culture of academic honesty by neither giving nor accepting unpermitted academic aid on this examination.

This course is participating in the proctoring pilot overseen by the Academic Integrity Working Group (AIWG), therefore proctors will be present in the exam room. The purpose of this pilot is to determine the efficacy of proctoring and develop effective practices for proctoring in-person exams at Stanford.

I acknowledge and accept the letter and spirit of the honor code. I pledge to write more neatly than I have in my entire life:

Signature: _____

First and Last Name (print): _____

Stanford Email (@stanford.edu): _____

Exam Break Sign-out

I pledge that during my exam break:

- I will not bring any paper, electronic devices (phone, smart watch, smart glasses, etc), or aid (permitted or unpermitted) *out of or into* the exam room, nor access any aid during the break.
- I will not communicate with anyone other than the course instructional staff about the content of the exam.

Signature Confirming Honor Code	Exit Time	Return Time	Proctor Initial	Length (mins)

If you are feeling unwell and are not able to complete the exam, please speak with the proctor.

1 Starting Lineup [10 Points]

A coach has a squad of 20 players: 3 goalkeepers, 7 defenders, 6 midfielders, and 4 forwards. Before each match she selects a starting lineup, which is an *unordered* set of 11 players.

You may not use code to answer any part of this problem.

- (a) (6 points) A valid lineup must consist of exactly 1 goalkeeper, 4 defenders, 4 midfielders, and 2 forwards. How many different starting lineups are possible?

Each position group is chosen independently, so by the product (step) rule we multiply the number of ways to choose each group:

$$\binom{3}{1} \binom{7}{4} \binom{6}{4} \binom{4}{2} = 3 \cdot 35 \cdot 15 \cdot 6 = 9450.$$

- (b) (4 points) Using the same position breakdown as in part (a), how many of these lineups include the team's star player, Sofia, who is one of the 6 midfielders?

Sofia must be one of the 4 chosen midfielders. Fix her in the lineup and choose the remaining 3 midfielders from the other 5; the other position groups are unchanged:

$$\binom{3}{1} \binom{7}{4} \binom{5}{3} \binom{4}{2} = 3 \cdot 35 \cdot 10 \cdot 6 = 6300.$$

Check: since 4 of the 6 midfielders are chosen, a given midfielder is included in a fraction $\frac{4}{6}$ of the lineups, and $9450 \cdot \frac{4}{6} = 6300$.

2 World Cup Penalty Kicks [14 Points]

It is the World Cup final, and once again France is playing Argentina. After a tightly contested match, the game goes to penalty kicks. Each team gets five shot attempts on the opponent's goal, and whichever team scores more goals wins. France scores on any one of its attempts with probability p_F , and Argentina scores on any one of its attempts with probability p_A . All ten attempts are independent.

You may not use code to answer any part of this problem.

- (a) (6 points) Let F be the number of goals France scores. State the distribution of F (including its parameters), give $E[F]$, and write an expression for the probability that France scores *exactly* 3 goals.

Each of France's 5 attempts is an independent success with probability p_F , so

$$F \sim \text{Bin}(n = 5, p = p_F), \quad E[F] = 5p_F.$$

Using the Binomial PMF,

$$P(F = 3) = \binom{5}{3} p_F^3 (1 - p_F)^2.$$

- (b) (8 points) What is the probability that France *wins* the match (i.e., scores strictly more penalty kicks than Argentina)? Let A be the number of goals Argentina scores. Write your answer as a sum in terms of the Binomial PMFs of F and A ; you do not need to simplify.

Let $A \sim \text{Bin}(5, p_A)$, independent of $F \sim \text{Bin}(5, p_F)$. France wins when $F > A$. Conditioning on Argentina's total a and summing over the ways France can beat it:

$$P(F > A) = \sum_{a=0}^4 \left(P(A = a) \sum_{f=a+1}^5 P(F = f) \right),$$

where

$$P(A = a) = \binom{5}{a} p_A^a (1 - p_A)^{5-a}, \quad P(F = f) = \binom{5}{f} p_F^f (1 - p_F)^{5-f}.$$

(The outer sum stops at $a = 4$ because France cannot exceed 5.)

3 Team USA Goals [13 Points]

The number of goals scored by Team USA in a single game follows a Poisson distribution with $\lambda = 2$. The number of goals scored in different games is independent.

You may not use code to answer any part of this problem.

- (a) (3 points) What is the probability that Team USA scores *exactly* 2 goals in a single game?

With $G \sim \text{Poi}(\lambda = 2)$,

$$P(G = 2) = \frac{e^{-2} 2^2}{2!} = 2e^{-2}.$$

- (b) (6 points) What is the probability that, over the next two games, Team USA scores *at least* 2 goals in total?

The sum of independent Poisson random variables is Poisson with the summed rate. Over two games, the total $T \sim \text{Poi}(\lambda = 2 + 2 = 4)$. Using the complement,

$$P(T \geq 2) = 1 - P(T = 0) - P(T = 1) = 1 - e^{-4} - 4e^{-4} = 1 - 5e^{-4}.$$

- (c) (4 points) What is the probability that Team USA scores *exactly* two goals in *at least one* of the next two games? (Scoring exactly two goals in *both* games also satisfies this condition.)

Let A be the event that game 1 has exactly two goals and B the event that game 2 has exactly two goals. From part (a), each game independently satisfies this with probability $p = 2e^{-2}$, and the two games are independent. By inclusion–exclusion,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = p + p - p^2 = 2(2e^{-2}) - (2e^{-2})^2 = 4e^{-2} - 4e^{-4}.$$

(Numerically, $P(A \cup B) \approx 0.468$.)

Alternative solution (using the complement):

Let A be the event that Team USA scores exactly two goals in game 1, and let B be the event that Team USA scores exactly two goals in game 2. From part (a),

$$P(A) = P(B) = 2e^{-2}.$$

We instead compute the complement:

$$P(A \cup B) = 1 - P(A^c \cap B^c).$$

Since the games are independent,

$$P(A^c \cap B^c) = P(A^c)P(B^c) = (1 - 2e^{-2})^2.$$

Therefore,

$$P(A \cup B) = 1 - (1 - 2e^{-2})^2 = 4e^{-2} - 4e^{-4}.$$

4 Ticket Prices [15 Points]

You want to buy a ticket to the Team USA game. The number of tickets still available, X , is a continuous random variable distributed as $X \sim \text{Exp}(\lambda = 1/1000)$ (so X has mean 1000). Because of dynamic pricing, the price of a ticket is

$$\text{price} = \frac{100000}{X}.$$

You may not use code to answer any part of this problem.

- (a) (4 points) You are only willing to spend \$200. What is the probability that you will be able to buy a ticket (i.e., that the price is at most \$200)?

You can afford a ticket when

$$\frac{100000}{X} \leq 200 \iff X \geq 500.$$

Using the Exponential CDF

$$P(X \leq x) = 1 - e^{-\lambda x},$$

we have

$$\begin{aligned} P(X \geq 500) &= 1 - P(X < 500) \\ &= 1 - \left(1 - e^{-\frac{1}{1000} \cdot 500}\right) \\ &= e^{-1/2}. \end{aligned}$$

Because X is continuous, $P(X < 500) = P(X \leq 500)$.

- (b) (6 points) Suppose instead you want at least a 90% chance of being able to buy a ticket. What is the smallest budget b (maximum price you are willing to pay) that achieves this? Give a closed-form expression.

With budget b , you can afford a ticket when $\frac{100000}{X} \leq b \iff X \geq \frac{100000}{b}$. So

$$P(\text{afford}) = P\left(X \geq \frac{100000}{b}\right) = e^{-\frac{1}{1000} \cdot \frac{100000}{b}} = e^{-100/b}.$$

Setting this equal to 0.9:

$$e^{-100/b} = 0.9 \implies -\frac{100}{b} = \ln(0.9) \implies b = \frac{-100}{\ln(0.9)} = \frac{100}{\ln(10/9)} \approx \$949.$$

- (c) (5 points) Suppose you were in fact able to buy a ticket with your original budget of \$200 from part (a) (that is, the price came out to at most \$200). Given this, what is the probability that you actually spent *less than* \$100?

Spending less than \$100 means $\frac{100000}{X} < 100 \iff X > 1000$, while being able to buy at \$200 means $X \geq 500$. Since $\{X > 1000\} \subseteq \{X \geq 500\}$, the conditional probability is

$$P(X > 1000 \mid X \geq 500) = \frac{P(X > 1000)}{P(X \geq 500)} = \frac{e^{-1}}{e^{-1/2}} = e^{-1/2}.$$

Notice this equals the answer to part (a). By the *memorylessness* of the Exponential, once we know $X \geq 500$ the number of remaining tickets beyond 500 has the same $\text{Exp}(1/1000)$ distribution, so the chance of at least another 500 is again $e^{-1/2}$.

5 Shot Speed [13 Points]

A star striker's shot speed S (in km/h) is Normally distributed with mean $\mu = 100$ and variance $\sigma^2 = 64$.

You may not use code to answer any part of this problem.

- (a) (3 points) A shot is called “world-class” if its speed exceeds 112 km/h. What is the probability that a single shot is world-class?

Here $\sigma = \sqrt{64} = 8$. Standardizing,

$$P(S > 112) = P\left(Z > \frac{112 - 100}{8}\right) = P(Z > 1.5) = 1 - \Phi(1.5).$$

- (b) (5 points) In a game the striker takes two independent shots, with speeds S_1 and S_2 (each distributed like S). Let $T = S_1 + S_2$ be their total speed. State the distribution of T (including its parameters) and write an expression for $P(T > 224)$.

The sum of independent Normal random variables is Normal, with means and variances adding:

$$T = S_1 + S_2 \sim \mathcal{N}(\mu = 200, \sigma^2 = 64 + 64 = 128),$$

so the standard deviation is $\sqrt{128} = 8\sqrt{2}$. Then

$$P(T > 224) = P\left(Z > \frac{224 - 200}{8\sqrt{2}}\right) = P\left(Z > \frac{24}{8\sqrt{2}}\right) = 1 - \Phi\left(\frac{3}{\sqrt{2}}\right) \approx 1 - \Phi(2.12).$$

- (c) (5 points) Let B be the event that *both* of the striker's shots are world-class (each exceeds 112 km/h). Is $P(B)$ less than, equal to, or greater than $P(T > 224)$ from part (b)? Justify your reasoning. (You may use English, math, or both.)

$$P(B) < P(T > 224).$$

If both shots are world-class then $S_1 > 112$ and $S_2 > 112$, which forces $T = S_1 + S_2 > 224$. So B is a *subset* of the event $\{T > 224\}$, and a subset can have no larger probability. The containment is strict: the total can exceed 224 even when one shot is below 112 (for example $S_1 = 100$, $S_2 = 130$), so $\{T > 224\}$ contains outcomes that are not in B . Hence $P(B) < P(T > 224)$.

For reference, $P(B) = (1 - \Phi(1.5))^2 \approx 0.0045$ while $P(T > 224) \approx 1 - \Phi(2.12) \approx 0.017$.

You're done! We hope the World Cup Edition went smoothly. A few fun facts: the penalty-kick problem is the same kind of head-to-head Binomial comparison that shows up in sports analytics all the time. The Poisson distribution really is a standard model for the number of goals scored in a soccer match. The "Ticket Prices" problem is a reminder that a transformation of a random variable is often best handled by working back to an event about the original variable — and that the Exponential distribution is *memoryless*. We hope you enjoyed the theme — good luck to your favorite side!