Chris Piech
CS109

Fall 2021
Oct 26th

# CS109 Midterm Exam

This is a closed calculator/computer exam. You are, however, allowed to use notes in the exam. You have 2 hours (120 minutes) to take the exam. The exam is 120 points, meant to roughly correspond to one point per minute of the exam. You may want to use the point allocation for each problem as an 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 $\Phi$ (the CDF of the standard normal). For example $\Phi(\frac{3}{4})$ is an acceptable final answer..

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: ______________________________________________

Family Name (print): ______________________________________________

Given Name (print): ______________________________________________

Email (preferably your gradescope email): ____________________

# 1 Color.com [20 points]

For covid testing, color.com produces 6 digit tracking codes. Code can start with 0s. Here are a few example codes: 123456, 000000, 091923.

a. (2 points) How many unique codes are possible with 6 digits?

b. (5 points) Color.com generates 3 random codes using a bad algorithm: each code is generated uniformly over all possible codes without checking for duplicates. What is the probability that none of them 3 random codes are the same?

c. (5 points) Every time a user enters a digit, there is a 1% chance that they enter it incorrectly. A code is incorrect if any digit is incorrect. What is the probability of an incorrectly entered 6-digit-code?

d. (8 points) 20,000 Stanford students enter their 6-digit-code independently. Use an approximation that could be used to efficiently compute the probability that more than 10 students will incorrectly enter a code.

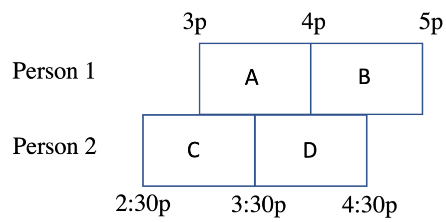
## 2 Doodle Poll [20 points]

You are trying to find a time when you can meet in a group. Assume a day is composed of 8 one hour blocks. Each block has $p = \frac{5}{8}$ probability of being busy for each person. We assume that whether each block is busy for a person is independent of whether other blocks are busy for them, and also independent of whether that block is busy for other people.

a. (5 points) What is the probability that a single person is busy for 6 or more blocks in a single day?

b. (7 points) Two people are trying to schedule a meeting. What is the probability that there is a block of time when they are both free on a given day?

c. (8 points) A person in India and a person in the UK are trying to schedule a one-hour meeting. In India, time zones are offset by half an hour relative to time zones in the UK. Here are their overlapping blocks:




The two people can only meet if blocks A, C, and D are all free, or if blocks A, B, and D are all free. What is the probability that the two people will be able to meet? [1]

[1] In case you are curious: all times in the figure are in India Standard Time. Person 2 is in the UK and can't meet before 2:30p because it would be too early in their local time.

## 3 I Heard That! [30 points]

We are trying to build a probabilistic at-home test to detect whether an 8 month old can hear a particularly soft sound. To do so, we are going to play the sound and then observe how much the baby's gaze moves in the following 3 seconds. If the baby hears the sound, they are more likely to have a large change in gaze. We have built a special camera which can record the change in gaze. Before observing any change in gaze, our prior belief is that:

$$\text{P(can hear the sound)} = \frac{3}{4}$$

$X$ is our measure of the angle that the baby's gaze moves over 3 seconds. We have estimated the following probability mass functions for $X$. These PMFs were obtained by observing tens of thousands of babies with known hearing ability. All values are discretized into buckets covering 5 degrees:

| Value of $X$ | PMF of $X$ given Baby can hear the sound | PMF of $X$ given Baby can **not** hear the sound |
|---|---|---|
| 0 to 5 | 0.08 | 0.40 |
| 5 to 10 | 0.15 | 0.30 |
| 10 to 15 | 0.35 | 0.12 |
| 15 to 20 | 0.20 | 0.08 |
| 20 to 25 | 0.12 | 0.05 |
| Above 25 | 0.10 | 0.05 |

a. (8 points) You play a sound and observe no movement ($X = 0$). What is the updated probability that she can hear the sound?

b. (6 points) You repeat the experiment to get a **second** independent reading and observe a 16 degree change — so you now have two independent observations, 0 degrees and 16 degrees. What is your newly updated probability? Let $p$ be your answer to part (a).

**The Normal Assumption**: We choose to approximate eye movements with normal distributions. For babies who **can** hear sounds, we approximate their gaze movement after the sound is played as: $N(\mu = 15, \sigma^2 = 50)$. For babies who can **not** hear sounds, we approximate gaze movement as $N(\mu = 8, \sigma^2 = 50)$.

c. (8 points) For a new baby we observe a 0 degree movement after the sound is played. What is your belief that a baby can hear, under The Normal Assumption? Since this is a new baby, you should not carry forward results from part (a) or (b).

d. (8 points) We want to check if The Normal Assumption was accurate. Write an expression for the probability that $10 < X < 15$ given that a baby **can hear the sound**, under the Normal Assumption. You can leave your answer in terms of $\Phi$.

## 4 Sleep and Dreams [30 points]

The number of times a baby wakes up in the night (an 8 hour period) is Poisson with rate lambda = 2.

a. (6 points) What is the probability that the baby wakes up exactly once during the night?

b. (6 points) The amount of rest you get is 0 if the baby wakes up more than 3 times. Otherwise, the amount of rest you get is $15 - 2X$, where $X$ is the number of times the baby wakes up. What is the expectation of rest?

c. (6 points) The baby wakes up half way through the night. What is the probability that she will not wake up again?

d. (12 points) Let $X_1$ be the first time she wakes up, measured in hours after she went to sleep. Given that she wakes up exactly once in the night, $N = 1$, prove mathematically that $X_1 \sim \text{Uni}(0, 8)$.
*Hint:* either show that $P(X_1 < x|N = 1) = \frac{x}{8}$ or that $f(X_1 = x|N = 1) = \frac{1}{8}$. Both are equally hard.

https://www.overleaf.com/project/617676717aa3502063d23e2c

## 5 Longest Sequence of Heads [20 points]

We are going to define a new random variable type $X \sim \text{Run}(n)$ where $P(X = x)$ is the probability of getting a **longest run** of exactly $x$ **heads** in $n$ coin flips. Here are a few examples with $n = 8$ coin flips:

| | |
|---|---|
| $HHHHHHHH$ | $X = 8$ since they are all heads. |
| $HTHTHTHT$ | $X = 1$ since there are no consecutive runs with two heads. |
| $HTTHHHHT$ | $X = 4$ since the longest run has 4 heads (HHHH). |
| $HHTTTTHH$ | $X = 2$ since the longest run has 2 heads (HH). |

a. (2 points) Let $X \sim \text{Run}(n)$. What is the support of $X$ (the values that $X$ can take on) in terms of $n$ ?

b. (6 points) Let $X \sim \text{Run}(n = 3)$. Provide the full Probability Mass Function for $X$:

c. (6 points) You are given a function $\texttt{count2}(n)$ which returns the count of all unique orderings of $n$ coin flips that have a longest run of 2 or fewer heads.

Here are a few base case examples:
$\texttt{count2}(0) = 2^0 = 1$, by definition.
$\texttt{count2}(1) = 2^1 = 2 \quad \{H, T\}$
$\texttt{count2}(2) = 2^2 = 4 \quad \{HH, HT, TH, TT\}$

Let $X \sim \text{Run}(n = 8)$. Define the probability $P(X \leq 2)$ using the $\texttt{count2}$ function.

d. (6 points) What is $\texttt{count2}(n)$ if $n > 2$? You may define your formula using a recursive call to count where the parameter is a value less than $n$, for example $\texttt{count2}(n-2)$.
*Hint: consider sequences that start with T, HT, and HHT.*

*That's all folks! The goal of the Doodle Polls, and Color.com problems was to inspire new ways to see probability in your day to day life. Non-invasive ways to measure babies ability to hear is an open probability problem (and one that we might make some progress on). Getting rest as a new parent is real. Problem (5), Longest sequence of heads, is a step towards solving the challenge: given a poisson process of baby wake ups, what the the probability of getting a chunk of sleep at least k units long.*