DS102 Fall 2019 - Final Exam

First and Last Name: ______________________________

Student ID: ______________________________

- Please write your first and last name as well as your student ID at the top of the first sheet. Also write your student ID on the bottom of each page.
- You have ? minutes: there are five questions on this exam, with each question being worth an equal amount of points.
- Make sure you have ? pages. If you do not, let us know immediately.
- Question 1 (true/false) is required.
- For the remaining four questions (Questions 2-?), we will grade all of them and take the top ? among these ?. You may attempt all questions or skip one depending on time.
- Even if you are unsure about your answer it is better to write down as many details as possible so we can give you partial credit.
- You may, without proof, use theorems and facts that were given in the discussions, lectures or notes.
- We will only grade work on the front of each page unless you indicate otherwise. The exam is printed 1-sided so that you can use the back sides for scratch paper. If you do run out of space on the front, continue on the back side of the page and make a note at the bottom of this cover sheet to let us know.
- Make sure to write clearly. We can't give you credit if we can't read your solutions.

1. (10 points) For each of the following, answer true or false. **Circle T for true and F for false**. You don't need to justify your answer.

   (a) (1 point) (T / F) In causal inference, a valid instrumental variable must be uncorrelated with the outcome.

   Solution: False. It should only affect the outcome through the variable of interest, though.

   (b) (1 point) (T / F) $g(X) = \mathbb{E}[Y|X]$ minimizes the mean squared error: $\mathbb{E}[(g(X) - Y)^2]$ over all bounded functions g of X.

   Solution: True, this can be found for example, in notes 6.

   (c) (1 point) (T / F) With good priors, Thompson sampling can be much more sample-efficient than UCB.

   Solution: True, we saw this in lecture.

   (d) (1 point) (T / F) Confidence intervals for the mean of a random variable X derived from Chebyshev's inequality are smaller if the variance of X is large.

   Solution: False, they are small if the variance is small.

   (e) (1 point) (T / F) In ϵ-differential privacy, smaller ϵ means more privacy.

   Solution: True.

   (f) (1 point) (T / F) Even if the UCB algorithm is run infinitely long, it is possible to never pull some arm a after a finite number of rounds R.

   Solution: False, UCB will pull all arms infinitely often.

   (g) (1 point) (T / F) In Q-learning a low discount value will mean that the learner will prioritize rewards that can occur sooner in time.

   Solution: True, since a low discount factor will cause the value of rewards to rapidly decay the further in the future they are.

   (h) (1 point) (T / F) To achieve sublinear regret with the upper confidence bounds algorithm you need to know the smallest gap between the expected reward of the optimal arm and the expected reward of sub-optimal arms.

Solution: False, the whole point of the algorithm is that you do not need to know the gaps.

(i) (1 point) (T / F) Chebyshev's inequality is equivalent to Markov's inequality applied to $(X - \mathbb{E}[X])^2$.

Solution: True.

(j) (1 point) (T / F) The function f described below is a linear function of θ.

$$f(x, \theta) = \theta e^{-x} + x\theta - \frac{1}{1+\theta^2}$$

Solution: False, $\frac{1}{1+\theta^2}$ is not a linear function of θ.

2. (10 points) P. Diddy has recorded a new hit, “Data 102”, and wants to sell it online. In a market study, he collects from n people the price $x_i \in [0, 1]$ they would be willing to pay for a download of the song. He collects the data into a data set S. Assuming that respondents answered truthfully, a reasonable estimate for the revenue P. Diddy would get from selling the downloads of ”Data 102” at price p is:

$$q(p; S) = p \cdot \#\{i : x_i \geq p\}.$$

P. Diddy would like to learn a price $p^* \in \mathcal{P} = \{\$0.01, \$0.02, \ldots, \$0.99\}$ that maximizes the revenue, $p^* = \arg\max_p q(p; S)$. However, P. Diddy is also a responsible data scientist and wants to protect the privacy of his fans, so he wants to learn p^* in a differentially private way.

To do so, he uses the *exponential mechanism*. In this problem, we go through the derivation of this mechanism, and prove that it satisfies differential privacy.

(a) As for the Laplace mechanism, we will need a notion of *sensitivity*. In this setting, we define sensitivity as:

$$\Delta = \max_{p \in \mathcal{P}} \max_{S, S' \text{ neighboring}} |q(p; S) - q(p; S')|.$$

Recall that S and S' being neighboring data sets means that they differ in a single individual. What is the numerical value of Δ?

> **Solution:** The maximum change in revenue happens if $\arg\max_p q(p; S) = \$0.99$, and we replace a user who is willing to pay \$0.99 by a user who is not willing to pay as much. Thus, $\Delta = \$0.99$.

(b) P. Diddy wants level of differential privacy equal to ϵ. He outputs a value $\hat{p}(S)$, which takes value $p \in \mathcal{P}$ with probability:

$$\mathbb{P}(\hat{p}(S) = p) = \frac{e^{\frac{\epsilon}{2\Delta} q(p;S)}}{\sum_{p' \in \mathcal{P}} e^{\frac{\epsilon}{2\Delta} q(p';S)}} \propto e^{\frac{\epsilon}{2\Delta} q(p;S)}.$$

Which value has the highest probability of being released as $\hat{p}(S)$?

> **Solution:** The most likely output is $\arg\max_p q(p; S) = p^*$.

(c) Show that, for neighboring data sets S and S',

$$\frac{e^{\frac{\epsilon}{2\Delta} q(p;S)}}{e^{\frac{\epsilon}{2\Delta} q(p;S')}} \leq e^{\epsilon/2}.$$

Solution:

$$\frac{e^{\frac{\epsilon}{2\Delta}q(p;S)}}{e^{\frac{\epsilon}{2\Delta}q(p;S')}} \leq e^{\epsilon/2} = e^{\frac{\epsilon}{2\Delta}(q(p;S)-q(p;S'))} \leq e^{\frac{\epsilon}{2\Delta}\Delta} = e^{\epsilon/2}.$$

(d) Show that, for neighboring data sets S and S',

$$\frac{\sum_{p'\in\mathcal{P}} e^{\frac{\epsilon}{2\Delta}q(p';S')}}{\sum_{p'\in\mathcal{P}} e^{\frac{\epsilon}{2\Delta}q(p;S)}} \leq e^{\epsilon/2}.$$

(Hint: Prove that $\frac{\sum_{i=1}^k a_i}{\sum_{i=1}^k b_i} \leq \max_{1\leq i\leq k} \frac{a_i}{b_i}$.)

Solution: First we prove the hint. Let $r_{\max} = \max_i \frac{a_i}{b_i}$. That means for for all $i \in \{1, \ldots, k\}$, $a_i \leq r_{\max} b_i$, hence $\sum_i a_i \leq r_{\max} \sum_i b_i$. Rearranging completes the proof.

Now we can apply the hint:

$$\frac{\sum_{p'\in\mathcal{P}} e^{\frac{\epsilon}{2\Delta}q(p';S')}}{\sum_{p'\in\mathcal{P}} e^{\frac{\epsilon}{2\Delta}q(p;S)}} \leq \max_{p'} e^{\frac{\epsilon}{2\Delta}q(p';S') - \frac{\epsilon}{2\Delta}q(p';S)} \leq e^{\frac{\epsilon}{2\Delta}\Delta} = e^{\epsilon/2}.$$

(e) Conclude that the exponential mechanism is ϵ-differentially private. That is, show that:

$$\mathbb{P}(\hat{p}(S) = p) \leq e^{\epsilon}\mathbb{P}(\hat{p}(S') = p),$$

for all neighboring data sets S and S', and all $p \in \mathcal{P}$.

Solution: We simply combine the previous two parts to get $e^{\epsilon/2}e^{\epsilon/2} = e^{\epsilon}$.

3. (10 points) Suppose that we are testing some number of hypotheses, and we are making decisions (discovery (1) vs no discovery (0)) according to some unknown decision rule.

   (a) Prove that $\mathbf{1}\{\text{at least one false discovery}\} \geq \text{FDP}$, where FDP denotes the false discovery proportion.

   Solution: If $\mathbf{1}\{\text{at least one false discovery}\} = 0$, then no false discovery has been made, in which case the FDP is clearly 0. If $\mathbf{1}\{\text{at least one false discovery}\} = 1$, then there is at least one false discovery, so $\text{FDP} = \frac{\#\text{ false disc.}}{\#\text{discoveries}}$, but since the number of discoveries is at least as big as the number of false discoveries, $\text{FDP} \leq 1$.

   (b) Prove that the family-wise error rate (FWER), i.e. the probability of making at least one false discovery, is at least as big as the false discovery rate (FDR):

$$\text{FWER} \geq \text{FDR}.$$

   Solution: We take expectations on both sides of the inequality from part (a) to get:

$$\mathbb{E}[\mathbf{1}\{\text{at least one false discovery}\}] \geq \mathbb{E}[\text{FDP}] \Leftrightarrow \text{FWER} \geq \text{FDR}.$$

   (c) Suppose we want to test possibly infinitely many hypotheses in an online fashion. At time $t \geq 1$, a p-value P_t arrives, and we proclaim a discovery if $P_t \leq \alpha_t$, where $\alpha_t = \left(\frac{1}{2}\right)^t \alpha$. Does this rule control the FWER under α? Give a proof or counterexample.

   Solution: We use the usual union-bound argument:

$$\text{FWER} \leq \sum_{i \in \text{nulls}} \mathbb{P}(P_i \leq \alpha_i) = \sum_i \left(\frac{1}{2}\right)^i \alpha = \alpha.$$

   Therefore, the rule does indeed control the FWER.

   (d) Does the rule from part (c) control the FDR under α?

   Solution: From part (b), we know that $\text{FDR} \leq \text{FWER}$, so if the FWER is under α, then so is the FDR.

4. (10 points) (a) Esther and Tijana separately (and independently from each other) take n i.i.d draws each from a Gaussian distribution with unknown mean μ and known variance σ^2. They individually compute frequentist confidence intervals for the mean from their samples, with confidence level $1-\alpha$ each. Show that the probability that their confidence intervals do not overlap is less than 2α.

Solution: We know that each interval contains μ with probability $\geq 1-\alpha$. We also know that if both intervals contain μ, they must be overlapping.

$$\begin{aligned}
\mathbb{P}[\text{intervals overlap}] &\geq \mathbb{P}[\text{both intervals contain } \mu] \\
&= \mathbb{P}[\text{Esther's interval contains } \mu] \cdot \mathbb{P}[\text{Tijana's interval contains } \mu] \\
&\geq (1-\alpha)^2 \\
&\geq 1 - 2\alpha
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}[\text{intervals do not overlap}] &= 1 - \mathbb{P}[\text{intervals overlap}] \\
&\leq 2\alpha
\end{aligned}$$

Note that a union bound on the events that either interval does not contain the mean will not suffice, as the two intervals could both not contain the mean but still overlap.

(b) Karl and Eric want to estimate the mean of the same Gaussian distribution, but they will use credible intervals. However, they do not agree on their priors. In particular, Karl specifies a Normal prior distribution on μ centered at μ_1, with standard deviation σ_p, $\mu \sim \mathcal{N}(\mu_1, \sigma_p)$. Eric, on the other hand, specifies prior distribution $\mu \sim \mathcal{N}(\mu_1, \sigma_p)$ where $\mu_2 > \mu_1$, with the same standard deviation σ_p. Karl and Eric will use the *same* n i.i.d draws $x_1, .., , x_n$ from the true distribution $\mathcal{N}(\mu, \sigma^2)$ distribution on the parameter μ.

i. Show that the the posterior distributions that each Karl and Eric will calculate after seeing the n samples $x_1, ..., x_n$ with sample average $m\bar{u} = \frac{1}{n}\sum_{i=1}^n x_i$are given by:

$$\mathcal{N}\left(\left(\frac{1}{\sigma_p^2} + \frac{n}{\sigma^2}\right)^{-1}\left(\frac{\mu_1}{\sigma_p^2} + \frac{n\bar{\mu}}{\sigma^2}\right), \left(\frac{1}{\sigma_p^2} + \frac{n}{\sigma^2}\right)^{-1}\right)$$

and

$$\mathcal{N}\left(\left(\frac{1}{\sigma_p^2} + \frac{n}{\sigma^2}\right)^{-1}\left(\frac{\mu_2}{\sigma_p^2} + \frac{n\bar{\mu}}{\sigma^2}\right), \left(\frac{1}{\sigma_p^2} + \frac{n}{\sigma^2}\right)^{-1}\right) .$$

Solution: TODO: write out. Solutions are in a discussion worksheet.

ii. Karl and Eric will construct *credible intervals* using their posterior distributions, taking as their intervals 2 standard deviations from the mean in either direction (roughly a 95% confidence interval).
As a function of μ_1, μ_2, σ_p and σ, calculate the smallest number of samples n for which we are guaranteed that Karl and Eric's calculated credible intervals overlap.

Solution: The intervals overlap if

$$\left(\frac{1}{\sigma_p^2}+\frac{n}{\sigma^2}\right)^{-1}\left(\frac{\mu_1}{\sigma_p^2}+\frac{n\bar{\mu}}{\sigma^2}\right)+2\left(\frac{1}{\sigma_p^2}+\frac{n}{\sigma^2}\right)^{-1/2}$$
$$\geq \left(\frac{1}{\sigma_p^2}+\frac{n}{\sigma^2}\right)^{-1}\left(\frac{\mu_2}{\sigma_p^2}+\frac{n\bar{\mu}}{\sigma^2}\right)-2\left(\frac{1}{\sigma_p^2}+\frac{n}{\sigma^2}\right)^{-1/2}$$

Simplifying a bit,

$$\left(\frac{\mu_1}{\sigma_p^2}+\frac{n\bar{\mu}}{\sigma^2}\right)\geq\left(\frac{\mu_2}{\sigma_p^2}+\frac{n\bar{\mu}}{\sigma^2}\right)-4\left(\frac{1}{\sigma_p^2}+\frac{n}{\sigma^2}\right)^{1/2}$$
$$\frac{\mu_1}{\sigma_p^2}\geq\frac{\mu_2}{\sigma_p^2}-4\left(\frac{1}{\sigma_p^2}+\frac{n}{\sigma^2}\right)^{1/2}$$

Solving for n:

$$\left(\frac{1}{\sigma_p^2}+\frac{n}{\sigma^2}\right)\geq\left(\frac{\mu_2-\mu_1}{4\sigma_p^2}\right)^2$$
$$n\geq\sigma^2\left(\left(\frac{\mu_2-\mu_1}{4\sigma_p^2}\right)^2-\frac{1}{\sigma_p^2}\right)$$

5. (10 points) You have a hypothesis that drinking boba tea after a workout causes better muscle recovery, and thus better performance later on, in runners. However, you also know that young people drink more boba tea, on average, and have faster mile times, on average. Thus, the causal diagram you consider looks like the following: Where a denotes age, b whether an individual drinks boba after a run, and t, mile time. The variable z is our designed incentive in the part .

Suppose you had an observational data set with elite professional runners and their post-run boba drinking habits. In this data set, you find that older runners (50+) on average drink less boba, and have slower mile times than younger runners (50 or less), who drink more boba. For each age group, boba drinkers have slower miletimes

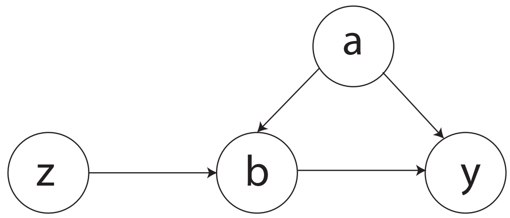

Figure 1: Causal Diagram for part b.

on average than the non-boba drinkers. However, ignoring age, you would find that runners who drink more boba have faster mile times.

Which of the following are true (circle all that apply):

1. There are more older runners than younger runners in the data set.
2. The population of this study is representative of the likely effect of boba on running for the UC Berkeley Student body.
3. The study gives evidence that there may be difference in mile times and boba drinking rates between young and old runners.

Solution: The first is false; the phenomenon observed is possible if there are more younger runners than older runners in the data set. The second is false; not all UC Berkeley students are professional runners. The third is true; there is evidence for difference in mile times and boba drinking across age groups.

(b) To your hypothesis, you enroll 100 participants in a study. The participants all do the same running workout, and after the workout you provide 50 of the participants with free boba, which they can either drink or not. Denote the fraction of the 50 treatment group participants who drink the offered boba as:

$$\bar{b}(1) = \frac{1}{50} \sum_{i \in \text{treatment}} \mathbb{I}[b_i = 1]$$

The second 50 participants (control group) you ask not to drink boba. However, they could still go get boba on their own, and so you also record the fraction of the 50 participants in the control group who complied with your intention:

$$\bar{b}(0) = \frac{1}{50} \sum_{i \in \text{control}} \mathbb{I}[b_i = 0]$$

You may assume that your participants answered whether they drank boba (b_i truthfully), and that there are no defyers in your study; anyone who acted against your request or offer who would have acted that way no matter what. A few days later, you time all 100 participants in a mile race and record their paces y_i, and calculate the average mile-times for the treatment group ($z = 1$, offered boba) and for the control group ($z = -1$, no boba):

$$\bar{y}(1) = \frac{1}{50} \sum_{i \in \text{treatment}} y_i$$
$$\bar{y}(0) = \frac{1}{50} \sum_{i \in \text{control}} y_i$$

Recall for this setting, that the two-stage least squares estimator for the causal effect of drinking boba on mile time (treating z as an instrumental variable) is:

$$\hat{\beta}_{IV} = (z^\top b)^{-1} z^\top y$$

Show that for this problem, this is equivalent to the average treatment effect estimator:

$$\hat{\tau}_c := \frac{\bar{y}(1) - \bar{y}(0)}{\bar{b}(1) - \bar{b}(0)}$$

That is, show that $\hat{\beta}_{IV} = \hat{\tau}_c$.

Solution: Since $z \in \{-1, 1\}$ then

$$\begin{aligned} z^\top b &= z_i b_i \\ &= \sum_{i:z_i=1} b_i - \sum_{i:z_i=-1} b_i \\ &= 50\bar{b}(1) - 50\bar{b}(0) \end{aligned}$$

Similarly, $z^\top y$ decomposes as

$$\begin{aligned} z^\top y &= z_i y_i \\ &= \sum_{i:z_i=1} y_i - \sum_{i:y_i=-1} b_i \\ &= 50\bar{y}(1) - 50\bar{y}(0) \end{aligned}$$

So that

$$(z^\top x)^{-1} z^\top y = \frac{50\bar{y}(1) - 50\bar{y}(0)}{50\bar{b}(1) - 50\bar{b}(0)} = \frac{\bar{y}(1) - \bar{y}(0)}{\bar{b}(1) - \bar{b}(0)}$$

(c) Suppose you run the regression from part (b), and you find that $\hat{\tau}_c = 0.5$. Interpret this coefficient in terms of the problem setting (at most one sentence).

Solution: We find that the estimated causal effect of drinking boba tea after a workout causes is a 30 second increase to miletime, as compared to not drinking boba after a workout.

(d) Suppose that the treatment effect of drinking boba on mile time change could be different for people of different age groups. In one sentence or less, describe a new experiment design that would account for this.

Solution: Blocking on age.

6. (10 points) The table below contains eight samples where each sample is of the form (x_i, y_i) where $x_i \in \{0,1\}^3$ are its features and $y_i \in \{0,1\}$ is its label.

Feature 1	Feature 2	Feature 3	Class
0	0	0	0
1	0	0	0
0	1	0	0
1	1	0	0
0	0	1	0
1	0	1	0
0	1	1	1
1	1	1	1

Furthermore, given a dataset with two class labels where p_0 is the proportion of elements with label 0 and p_1 is the proportion of elements with label 1, recall that the Gini purity is defined as

$$\phi(p_0, p_1) = p_0(1 - p_0) + p_1(1 - p_1).$$

(a) Describe the procedure to create a decision tree using the CART algorithm and the Gini purity.

(b) What is the Gini purity of the dataset above?

(c) Which feature should we split on when constructing a tree for the dataset? Why?

(d) Draw a complete decision tree constructured using CART for this dataset where each node consists of a decision rule based on one feature.

(e) Describe a way in which we might avoid overfitting in decision trees.

7. (10 points) Assume that we have the following gridworld

-900		S	1
-900		1	
-900			10

where S represents our starting point, and the 1, 10, and -100 cells represent terminal states with corresponding rewards. For parts a-e, assume deterministic state transitions, meaning that an action in a specific direction always moves us in that direction (unless it's toward the boundary of the world in which case we remain stationary).

(a) Write down the optimal value function at each empty cell below when the discount factor γ is 0.9. You may leave your answer in terms of powers of numbers.

-900			1
-900		1	
-900			10

Solution:

N/A	$.9^3 10$	$.9^4 10$	N/A
N/A	$.9^2 10$	N/A	10
N/A	9	10	N/A

(b) Compute the optimal Q-function at our starting point for the action of going up, down, left, and right when the discount factor $\gamma = 0.9$.

(c) Write down the optimal value function at each empty cell below when the discount factor γ is 0.1. You may leave your answer in terms of powers of numbers.

-900			1
-900		1	
-900			10

Solution:

N/A	0.1	1	N/A
N/A	0.1	N/A	10
N/A	1	10	N/A

(d) Compute the optimal Q-function at our starting point for the action of going up, down, left, and right when the discount factor $\gamma = 0.1$.

(e) What are the optimal moves to make at the starting point given discount factors of 0.1 and 0.9? Are they the same? Give intuition for why or why not.

(f) Let the discount factor $\gamma \in (0, 1]$. Now suppose the state transitions are stochastic at every cell except the starting cell. At the starting cell you will go in the direction you want with probability 1. At every other cell you will go in your specified direction with probability 0.7 and you have a probability of 0.1 of going in any other direction. For example if you decide to go up you will have a 0.7 probability of going up, a 0.1 probability of going left, a 0.1 probability of going right, and a 0.1 probability of going down. Without computing Q-functions or value functions what do you think is the best action to perform at the starting point? Does your answer depend on the value of the discount factor? Explain your reasoning.

Solution: You should go right no matter the discount factor. If we try to go for the reward of 10 we could easily fall in the pit, no matter what the discount factor is the probability is significant enough that our Q-function will always be negative by going left.

8. (10 points)

```
import numpy as np

def min_coins(coins, total):
    if total < 0:
        return np.inf

    num_coins = 0
    for coin in coins:
        if min_coins(coins, total - coin) + 1 < num_coins:
            num_coins = min_coins(coins, total - coin) + 1

    return num_coins
```

(a) There are **two** bugs in this code that will cause it to give the wrong answer. What are they? How would you fix the code to make it produce the correct output?

> **Solution:** The first bug is that we do not account for the base case where total is 0, in that case we should return 0. The second bug is that we should be initializing num_coins to inf.

(b) Let n be the total monetary amount for which we are asking change. Is the number of recursive calls the code has to make going to be closer to 2^n or n? Why? Given your answer do you think this code would be reasonable to deploy in production so that it can be used to compute change for customers at point of sale terminals?

> **Solution:** It will be closer to 2^n since each function call will make at least two sub-function calls, that will in turn make at least two sub-function calls each until we hit a base-case. Hence the number of recursive calls will grow exponentially.

(c) There are two ways in which the code can be sped up. One will lead to a major speed-up the other will lead to a minor speed-up. Describe what those two ways are and what changes you would need to make to implement them.

> **Solution:** The first minor way to speedup code is to cache the result of the recursive call to min_coins in a temp variable and then use that temp variable in both the comparison and the assignment. The major way to speedup the code is to use memoization by caching the minimum number of coins to make change in an array of length n where the $i - th$ element will hold the number of coins to make change for $i.

(d) Given the speedup changes is the number of recursive calls the code has to make going to be closer to 2^n or n? Why?

Solution: n since we are clearly not making an exponential number of calls anymore.

9. (10 points) During the project you learned a mixture of Gaussian distributions to describe the distribution of the time of day when two populations of riders rent bikes. However, by observing the data you noticed the distributions was skewed. In this problem you will analyze learning a mixture of Poisson distributions. You first round all the customer's arrivals to the nearest minute of the day, and then assume that the Customers and Subscribers are being generated from a mixture of different Poisson distributions.

$$Poisson(k;\lambda) = \frac{\lambda^k}{k!}e^{-\lambda}$$

Thus, for each data point $i = 1, ..., n$:

$$x_i \sim Bernouilli(\theta)$$
$$y_i \sim Poisson(\lambda_{x_i})$$

You have observed (x_i, y_i) (i.e the user type and the rental time) but you do not know θ, or the parameters λ of the Poisson distributions.

(a) Express the Likelihood function $p(x, y; \theta, \lambda_0, \lambda_1)$ in terms of the data $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$ and the parameters of the distributions given.

Solution:

$$p(x, y; \theta, \lambda_0, \lambda_1) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{(1-x_i)} Poisson(y_i; \lambda_{x_i})$$

(b) Write an expression for the log-likelihood of the data as a function of the data and the parameters of the distribution.

Solution:

$$\begin{aligned}\log p(x, y; \theta, \lambda_0, \lambda_1) &= \sum_{i=1}^{n} x_i \log\theta + (1-x_i)\log(1-\theta) + \log Poisson(y_i|\lambda_{x_i}) \\ &= \sum_{i=1}^{n} x_i \log\theta + (1-x_i)\log(1-\theta) + y_i \log\lambda_{x_i} - \log y_i! - \lambda_{x_i}\end{aligned}$$

(c) Write an expression for the maximum likelihood estimates of θ (θ_{MLE}). As a function of the observed data $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$. (Hint: you may want to write θ_{MLE} as a function of $C = \sum_{i=1}^{n} x_i$.

Solution: This is a freebie, its the same as for Gaussians. Differentiate the log-likelihood w.r.t. θ:

$$\frac{\partial l}{\partial \theta}\ell(\theta, \lambda_0, \lambda_1 | x, y) = \sum_{i=1}^{n} \left\{ \frac{x_i}{\theta} - \frac{1 - x_i}{1 - \theta} \right\} \qquad = \frac{C}{\theta} + \frac{n - C}{1 - \theta}$$

Setting $\frac{\partial l}{\partial \theta}\ell(\theta, \mu_0, \mu_1 | x, y; \sigma_0, \sigma_1) = 0$ and solving for θ reveals that

$$\hat{\theta}_{MLE} = \frac{C}{n}$$

(d) Derive the estimates $\hat{\mu_0}_{MLE}$ and $\hat{\mu_1}_{MLE}$ as a function of the observed data $x_1, ..., x_n$ and $y_1, ..., y_n$, and the parameters λ_0 and λ_1.

Solution: Taking the derivative of the log-likelihood with respect to μ_0, we get:

$$\frac{\partial l}{\partial \lambda_0}\ell(\theta, \lambda_0, \lambda_1 | x, y) = \sum_{i=1}^{n} \mathbb{I}\{x_i = 0\} \left(\frac{y_i}{\lambda_0} - 1 \right)$$

Let $C_0 = \sum_{i=1}^{n} \mathbb{I}\{x_i = 0\}$. Setting equal to zero and solving gives:

$$\begin{aligned}
0 &= \sum_{i=1}^{n} \mathbb{I}\{x_i = 0\} \left(\frac{y_i}{\lambda_0} - 1 \right) \\
C_0 &= \sum_{i=1}^{n} \mathbb{I}\{x_i = 0\} \frac{y_i}{\lambda_0} \\
\lambda_0 &= \sum_{i=1}^{n} \mathbb{I}\{x_i = 0\} \frac{y_i}{C_0}
\end{aligned}$$

Similarly:

$$\lambda_1 = \sum_{i=1}^{n} \mathbb{I}\{x_i = 1\} \frac{y_i}{C_1}$$

10. (10 points) We will now investigate the regret of UCB on a 2-armed bandit. There are

two arms each with 1-subgaussian reward distributions. Arm 1 has the higher mean (i.e. $\mu_1 > \mu_2$). The upper confidence bounds for $i = 1, 2$ are:

$$UCB_i(t) = \hat{\mu}_i(T_i(t)) + \sqrt{\frac{2}{T_i(t)} \log \frac{1}{\delta}}$$

where $T_i(t) \geq 1$ is the number of times arm i has been pulled up to time t and $\hat{\mu}_{i,t}$ is the empirical mean of arm i up to time t:

$$\hat{\mu}_i(T_i(t)) = \frac{1}{T_i(t)} \sum_{k=1}^{t} r_k \mathbb{I}\{A_k = i\}.$$

(a) Let n be a positive integer, $\delta \in (0, 1)$, and $T > n$ be fixed and define the event:

$$G = \{\mu_1 < \min_{t \leq T} UCB_1(T_1(t))\} \cap \left\{ \hat{\mu}_2(n) + \sqrt{\frac{2}{n} \log \frac{1}{\delta}} < \mu_1 \right\}$$

Show, using the definition of the UCB algorithm that arm 2 will never be chosen more than n times by time T by the UCB algorithm if the event G is true. (Hint: Argue by contradiction that if G is true, then $T_2(t) \leq n$).

Solution: Suppose that G is true but $T_2(T) > n$, then there is a time k when $T_2(k) = n$ and arm 2 is chosen.
Therefore, at that time:

$$UCB_1(k) < UCB_2(k)$$

However, if G is true:

$$UCB_2(k) = \hat{\mu}_2(n) + \sqrt{\frac{2}{n} \log \frac{1}{\delta}} < \mu_1 < UCB_1(k)$$

which is a contradiction.

(b) Show that the expected regret of the UCB algorithm on the event G is bounded below $n\Delta$, where $\Delta = \mu_1 - \mu_2 > 0$. i.e show:

$$\mathbb{E}[R(T)|G] \leq n\Delta$$

Solution: Using the regret decomposition:

$$\mathbb{E}[R(T)|G] = \mathbb{E}[T_2(t)|G]\Delta \leq n\Delta$$

(c) Let us now analyze the complement of G:

$$G^c = \left\{\mu_1 > \min_{t \le T} UCB_1(T_1(t))\right\} \cup \left\{\hat{\mu}_2(n) + \sqrt{\frac{2}{n}\log\frac{1}{\delta}} > \mu_1\right\}$$

Suppose that we choose $\delta = \frac{1}{T^2}$, and $n = \frac{16\log T}{\Delta^2}$, use the Chernoff-Hoeffding bound on 1- sub-gaussian random variables to upper bound the probability of the second event in G^c:

$$G_2^c = \left\{\hat{\mu}_2(n) + \sqrt{\frac{2}{n}\log\frac{1}{\delta}} < \mu_1\right\}$$

Recall that the Chernoff-Hoeffding bound for 1-subgaussian random variables is given by:

$$Pr(\hat{\mu}_2(n) - \mu_2 > t) \le e^{-nt^2/2}$$

Solution: With these choices of the constants:

$$\begin{aligned} G_2^c &= \left\{\hat{\mu}_2(n) + \sqrt{\frac{2}{n}\log\frac{1}{\delta}} > \mu_1\right\} \\ &= \{\hat{\mu}_2(n) - \mu_2 > \Delta/2\} \end{aligned}$$

Using the Chernoff bound, we know that the probability of this event is upper bounded by:

$$Pr(G_2^c) \le e^{-4\log(T)/2} = \frac{1}{T^2}$$

(d) For the first event in G^c:

$$G_1^c = \left\{\mu_1 > \min_{t \le T} UCB_1(T_1(t))\right\}$$

Argue (in words or mathematically) why:

$$\begin{aligned} G_1^c = \left\{\mu_1 > \min_{t \le T} UCB_1(T_1(t))\right\} &\subseteq \bigcup_{k=1}^{T_1(T)} \left\{\hat{\mu}_1(k) - \mu_1 < -\sqrt{\frac{2}{k}\log\frac{1}{\delta}}\right\} \\ &\subseteq \bigcup_{k=1}^{T} \left\{\hat{\mu}_1(k) - \mu_1 < -\sqrt{\frac{2}{k}\log\frac{1}{\delta}}\right\} \end{aligned}$$

Solution: G_1^c is the even that the minimum value of the upper confidence bound over all time up to time T is less than the true mean. If the upper confidence bound given any of $k \leq T_1(T)$ samples from arm 1 is less than the true mean, then G_1^c must be true. This second even is trivially a subset of having picked the arm at every time from $t = 1$ to T.

(e) Use the Union Bound (and the Hoeffding bound on the lower tail) to show that:

$$Pr(G_1^c) \leq T_1(T)\delta$$

Recall that the Chernoff-Hoeffding bound for the lower tail 1-subgaussian random variables is given by:

$$Pr(\hat{\mu}(n) - \mu < -t) \leq e^{-nt^2/2}$$

Solution:

$$\begin{aligned} Pr\left(\bigcup_{k=1}^{T_1(T)} \left\{\hat{\mu}_1(k) - \mu_1 < -\sqrt{\frac{2}{k}\log\frac{1}{\delta}}\right\}\right) &\leq \sum_{k=1}^{T_1(T)} Pr\left(\hat{\mu}_1(k) - \mu_1 \leq -\sqrt{\frac{2}{k}\log\frac{1}{\delta}}\right) \\ &= T_1(T)\delta \end{aligned}$$

(f) Show that the expected regret of the UCB algorithm is upper bounded by:

$$Pr(G^c)\mathbb{E}[R(T)|G^c] \leq 2\Delta$$

when $n = \frac{16\log(T)}{\Delta^2}$, and $\delta = \frac{1}{T^2}$. (Hint: Upper bound $\mathbb{E}[T_2(T)|G^c] < T$).

Solution: Using the regret decomposition and the union bound:

$$\begin{aligned} Pr(G^c)\mathbb{E}[R(T)|G^c] &= Pr(G^c)\Delta\mathbb{E}[T_2(T)|G^c] \\ &\leq \left(T\delta + \frac{1}{T^2}\right)\Delta\mathbb{E}[T_2(T)|G^c] \\ &\leq 2\Delta \end{aligned}$$

(g) Show that when $n = \frac{16\log(T)}{\Delta^2}$, and $\delta = \frac{1}{T^2}$:

$$\mathbb{E}[R(T)] \leq \lceil n \rceil\Delta + 2\Delta$$

where $\lceil n \rceil$ is the ceiling function applied to n (i.e n rounded up to the nearest integer).

Does this show that the regret is sub-linear?

Solution: Using the regret decomposition and the union bound:

$$\begin{aligned}\mathbb{E}[R(T)] &= Pr(G)\mathbb{E}[R(T)|G] + Pr(G^c)\mathbb{E}[R(T)|G^c] \\ &\leq \lceil n \rceil Pr(G)\Delta + 2\Delta \\ &\leq \lceil n \rceil \Delta + 2\Delta\end{aligned}$$

Yes this is sublinear because there is only a logarithmic T dependence.

11. (10 points) In this problem we will look at the Chebyshev, and Hoeffding bounds for sums of i.i.d exponential random variables. For $i = 1, ..., n$, let X_i be i.i.d random variables such that:

$$X_i \sim Exponential(\lambda)$$

where $Exponential(x; \lambda) = \lambda exp(-\lambda x)$ for $x \geq 0$. Recall that the mean and variance of an exponential random variable is:

$$\mathbb{E}[X] = \lambda \qquad Var(X) = \frac{1}{\lambda}$$

Define $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

(a) Use the Chebyshev bound to find an upper bound on:

$$Pr(\hat{\mu} - \lambda > (1+c)\lambda)$$

Solution:

$$Pr(\hat{\mu} - \lambda > (1+c)\lambda) \leq \frac{1}{n(1+c)^2\lambda^4}$$

(b) Write out an expression for the Moment Generating Function of an Exponential random variable:

$$M(t) = \mathbb{E}[e^{tX}]$$

Be sure to explicitly state for which values of t the moment generating is finite.

Solution:

$$M(t) = \mathbb{E}[e^{tX}]$$
$$= \int_0^\infty \lambda e^{tX - \lambda x} dx$$

Suppose that $t < \lambda$:

$$M(t) = \int_0^\infty \lambda e^{(t-\lambda)x} dx$$
$$= \frac{1}{(\lambda - t)}$$

(c) What is the moment generating function, $M_Z(t)$ of:

$$Z = \sum_{i=1}^{n} X_i$$

Solution: Since the moment generating function of a sum of random variables is a product of the moment generating functions, we have that:

$$M_Z(t) = \mathbb{E}[e^{tX}]$$
$$= \frac{\lambda^n}{(\lambda - t)^n}$$

(d) Use the Chernoff bound to derive an upper bound on:

$$Pr(\hat{\mu} - \lambda > (1 + c)\lambda)$$

Does the bound you derive decay faster than the Chebyshev bound for large values of n?

Recall that the Chernoff Bound is given by:

$$Pr(Z > c) \leq \inf_t \mathbb{E}[e^{tZ - tc}]$$

Solution: Rearranging, we first note that we need:

$$Pr(Z > cn\lambda)$$

Using the standard definition of the Chernoff bound:

$$\begin{aligned} Pr(Z > cn\lambda) &= Pr(e^{tZ} > e^{tcn\lambda}) \leq \inf_{t<\lambda} \mathbb{E}[e^{tZ-tcn\lambda}] \\ &= Pr(e^{tZ} > e^{tcn\lambda}) \leq \inf_{t<\lambda} \frac{\lambda^n}{(\lambda-t)^n} e^{-tcn\lambda} \end{aligned}$$

We now need to minimize over t. Taking the derivative of the expression and setting equal to zero gives:

$$\begin{aligned} 0 &= \frac{n\lambda^n}{(\lambda-t)^{n-1}} e^{-tcn\lambda} - cn\lambda^{n+1} \frac{1}{(\lambda-t)^n} e^{-tcn\lambda} \\ c\lambda &= \lambda - t \\ t &= (1-c)\lambda \end{aligned}$$

Plugging this in gives:

$$\begin{aligned} Pr(\hat{\mu} - \lambda > (1+c)\lambda) &\leq \frac{1}{c^n} e^{-(c-c^2)n\lambda^2} \\ &\leq e^{-(c-c^2)n\lambda^2 - nlog(c)} \end{aligned}$$

Yes this decays faster because the exponential decays faster than $1/n$ when $c < 1$.