DS 102 Fall 2020 Final

- You have 170 minutes to complete this exam. There are 7 questions, totaling 67 points.
- The exam is open notes and open internet.
- You can choose to write your solutions inside the exam sheet or write them on a separate sheet. You can use use either pen and paper, use a tablet or typeset your solutions.
- Even if you are unsure about your answer, it is better to write down partial solutions so we can give you partial credit.
- Make sure to write clearly. We can't give you credit if we can't read your solutions.
- If you have clarification questions first check the **Clarification Page** on Piazza, where we will add in real time clarification based on student questions. If you still have a question, make a private Piazza post.
- After 170 minutes have passed, please submit your solutions to Gradescope. You will have 10 minutes to do so, after which the Gradescope submission will close; these 10 minutes are for uploading solutions and not for continuing to work on the exam. *If you have gradescope issues you can email the solutions to us at data102.exam@gmail.com*

Honor Code

I will respect my classmates and the integrity of this exam by following this honor code.

I affirm:

- All of the work submitted here is my original work.
- I did not collaborate with anyone else on this exam.

Signature: ____

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) Thompson sampling always accumulates less regret than UCB.

Solution: False, if the priors on the arms are bad TS can perform very poorly.

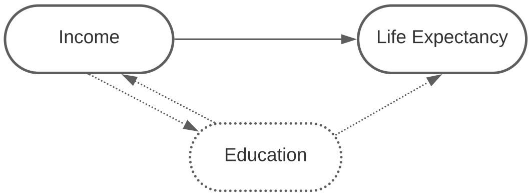
(b) (1 point) (T / F) In the band-optimal stable matching between drummers and bands, each band gets their most preferred drummer.

Solution: False, band-optimal means that the bands get the best attainable drummer, where attainable drummers are a subset of all the drummers.

(c) (1 point) (T / F) Increasing the number of hidden layers in a neural network increases the bias.

Solution: False, typically more expressive model tend to have lower bias, even if this isn't always true.

(d) (1 point) (T / F) Assume that you are trying to estimate the causal effect of Income on Life Expectancy in the US. Education is often a confounder, as people with more education tend to have both higher income and longer life expectancy. Given the causal diagram below, is it true or false that if Education is omitted, then you would likely over-estimate the causal effects of Income on Life Expectancy?



Solution: True

(e) (1 point) (T / F) In the GridWorld environment the value of a state $V(s)$ is the sum of Q values corresponding to all possible actions from state s; i.e., $V(s) = \sum_{a \in \text{all possible actions}} Q(s, a)$.

Solution: False, given policy π we have that $V(s) = \sum_a \pi(a|s)Q(s, a)$

(f) (1 point) (T / F) The Bellman equation states that the value of the current state can be decomposed as the reward at the current state/action and the (discounted) value of the future state.

Solution: True.

(g) (1 point) (T / F) Consider a random variable that has a finite mean, but an infinite variance. The Chebyshev inequality provides a correct, but vacuous, bound on tail probabilities in this case (i.e., the bound on the probability exceeds one).

Solution: True.

(h) (1 point) (T / F) The ϵ-Greedy algorithm for a multi-armed bandit achieves sublinear regret.

Solution: False.

(i) (1 point) (T / F) Under Gaussian likelihood ($y^{(i)} \sim N(\beta^T X^{(i)}, 1)$), the MLE estimator $\hat{\beta}_{MLE}$ is identical to the OLS estimator $\hat{\beta}_{OLS}$.

Solution: True

(j) (1 point) (T / F) If a test fails to reject a hypothesis $H_0 : \mu = 0$ at a 5% significance level, then it automatically fails to reject at 99% confidence level.

Solution: True

2. (10 points) **Distribution Shift**

A team of researchers is working on face recognition software. They have obtained photographs of 10,000 individuals in Company A (one photo per individual), across four demographic groups: *darker skin females*, *darker skin males*, *lighter skin females* and *lighter skin males* (each individual belongs to one group). Table 1 contains the breakdown across the four categories for the training dataset:

	Darker Skin	Lighter Skin
Female	10%	10%
Male	30%	50%

Table 1: Gender-Shade Distribution in Company A

The classifier trained on this data has the following classification accuracy across the four demographic groups:

	Darker Skin	Lighter Skin
Female	0.55	0.7
Male	0.85	0.95

Table 2: Gender-Shade Accuracy

(a) (2 points) What is the average accuracy of the classifier?

Solution: The average accuracy is: $0.55 \times 0.1 + 0.7 \times 0.1 + 0.85 \times 0.3 + 0.95 \times 0.5 = 0.855$

(b) (2 points) Encouraged by the average accuracy in part (a), Company B decides to adopt the classifier. The demographics in Company B are as follows:

	Darker Skin	Lighter Skin
Female	10%	40%
Male	40%	10%

Table 3: Gender-Shade Distribution in Company B

Compute the expected average accuracy for individuals in Company B.

Solution: The expected accuracy is: $0.55 \times 0.1 + 0.7 \times 0.4 + 0.85 \times 0.4 + 0.95 \times 0.1 = 0.677$

(c) (2 points) Why is the accuracy lower in Company B?

Solution: Company B has a smaller proportion of lighter skin males compared to the training data, and the classifier performs better for such individuals.

(d) (2 points) A journalist argues that the classifier should not be used because, referring to Table 2, she notes that it yields worse accuracy for darker skin individuals, and indeed this holds for both women and for the men w(the two rows of the table). A lawyer at Company B investigates this by computing the average accuracy (averaging across gender) for darker skin individuals and lighter skin individuals in the data from Company B.

Do this computation yourself—compute the average accuracy for lighter skin individuals and the average accuracy for darker skin individuals (averaging across gender). Which accuracy is larger?

Solution: The average accuracies are: 0.75 for lighter skin individuals and 0.77 for darker skin individuals.

(e) (2 points) The lawyer claims that the accuracy is roughly the same for darker skin and lighter skin individuals; indeed, it's slightly larger for darker skin individuals. Is he correct? If so, is this claim consistent with the journalist's observation that the classifier is worse for darker skin individuals? The lawyer concludes that the classifier can be deployed after all, and the journalist disagrees. Who do you support?

Solution: This is an example of Simpson's paradox. They should support the journalist. (Or give a good reason why they don't).

3. (10 points) **Differentially private Bayesian estimator.**

The GSIs of Data 102 would like to estimate the proportion of students who attend office hours at least once. We denote the proportion as μ. Assume that there are n students in all. Based on past experience, we let the prior for μ be $\mu \sim Beta(\alpha, \beta)$, for some $\alpha > 2, \beta > 2$. We assume that for each individual student, $i \in [n]$, the attendance is distributed as a Bernoulli distribution: $X_i \,|\, \mu \sim Bernoulli(\mu)$. Recall that the density function for the beta distribution is parameterized by shape parameters $\alpha > 0$ and $\beta > 0$, and is given by

$$f(z; \alpha, \beta) \propto z^{\alpha-1}(1-z)^{\beta-1}, \quad 0 < z < 1.$$

Recall also that the mean of the beta distribution is equal to $\alpha/(\alpha+\beta)$. Finally, recall that the Bernoulli distribution, $Bernoulli(\mu)$, is the distribution of a random coin toss that takes the value 1 with probability μ and the value 0 with probability $1-\mu$.

(a) (3 points) Derive the posterior distribution $P(\mu|X_1, \cdots, X_n)$.

Solution: The posterior distribution is $Beta(\alpha + \sum X_i, \beta + n - \sum X_i)$.

(b) (2 points) Derive the mean $\hat{\mu}_{pos}$ of the posterior distribution (the *posterior mean*), in terms of α, β, and the X_i.

Solution:

$$\hat{\mu}_{pos} = \frac{\alpha + \sum X_i}{\alpha + \beta + n}.$$

(c) (3 points) Now suppose we want to compute the posterior mean in an ϵ-differentially private way. Recall that a differentially private algorithm $\hat{\mu}$ is an algorithm which, for all databases S' that are neighbors of our database $S = \{X_1, \cdots, X_n\}$ (meaning they differ in only one entry; one can change one and only one of the X_i to a different value), satisfies the following:

$$\frac{\mathbb{P}(\hat{\mu}(S) = a)}{\mathbb{P}(\hat{\mu}(S') = a)} \leq e^{\epsilon}.$$

We adopt the *Laplace mechanism*, which outputs $\hat{\mu}_{Lap}(X_1, \cdots, X_n) = \hat{\mu}_{pos}(X_1, \cdots, X_n) + \xi_\epsilon$, where $\xi_\epsilon \sim Lap(0, p)$. Find the smallest scale parameter p such that $\hat{\mu}_{Lap}$ is ϵ-differentially private.

Solution: We know that X_i is always bounded in between 0 and 1. Thus the sensitivity is $\Delta_f = \sup_{\text{neighboring } S,S'} |\hat{\mu}(S) - \hat{\mu}(S')| \leq 1/(\alpha + \beta + n)$. From Discussion 13, we know that one can pick $p = 1/(\epsilon(\alpha + \beta + n))$

(d) (2 points) How does the noise level p change with respect to α, β and n? Explain intuitively why this is the case.

Solution: The noise level p decreases as α, β, n grow larger. This is intuitively correct. We know that α, β can be viewed as 'phantom samples' obtained in the prior, and n is the number of samples we acquired. More samples mean that an individual sample will reveal less information. Thus it suffices to add some Laplace distribution with smaller variance.

4. (10 points) **Upper confidence bounds**

Netflix would like to compare a set of recommendation algorithms by running an experiment in which algorithms are tried out on small group of users. They also want to ensure that users are getting a good overall experience during this experiment. Thus, they use a bandit algorithm. In each round, they randomly select one recommendation algorithm (arm) $a \in [K]$ out of K algorithms (arms), launch it, and observe the performance X_a, which is some measure of the engagement of the users. Assume that for each arm $a \in [K]$, the reward X_a is a random variable with mean $\mu_a = \mathbb{E}[X_a]$. Netflix uses the upper confidence bound algorithm (UCB) to help them balance the exploration and exploitation.

(a) (3 points) Rather than making the usual assumption of bounded reward, Netflix prefers to only assume that the reward of any arm a has bounded variance:

$$\mathbb{E}[(X_a - \mu_a)^2] \leq \sigma^2, \tag{1}$$

for all a and for some $\sigma > 0$. Using a concentration inequality, provide an upper bound on $P(|X_a - \mu_a| \geq \gamma)$ that is as tight as possible. (Here $\gamma > 0$. Your final bound can depend on γ and σ.)

Solution: From Chebyshev's inequality, we know that

$$P(|X_a - \mu_a| \geq \gamma) \leq \frac{\sigma^2}{\gamma^2}. \tag{2}$$

(b) (3 points) Recall that the general method for constructing an upper confidence bound for the true mean μ_a of an arm a, given $T_a(t)$ i.i.d. samples $X_a^{(1)}, ..., X_a^{(T_a(t))}$, is to find a value of $C_a(T_a(t), \delta)$ such that:

$$P(\mu_a < \hat{\mu}_{a,T_a(t)} + C_a(T_a(t), \delta)) > 1 - \delta, \tag{3}$$

where $\hat{\mu}_{a,T_a(t)}$ is a sample mean: $\hat{\mu}_{a,T_a(t)} = \frac{1}{T_a(t)} \sum_{i=1}^{T_a(t)} X_a^{(i)}$. Now under the same variance bound assumption (Equation (1) in (a)), construct a tight upper confidence bound $C_a(T_a(t), \delta)$ for arm a, after observing $T_a(t)$ sample rewards from arm a. (Your final answer should depend on σ. δ and $T_a(t)$.)

Solution: Since X_a has variance bounded by σ^2, we have

$$\mathbb{E}[(\hat{\mu}_{a,T_a(t)} - \mu_a)^2] \leq \frac{\sigma^2}{T_a(t)}.$$

Thus by Chebyshev's inequality, we have

$$P(|\hat{\mu}_a - \mu_a| \geq \gamma) \leq \frac{\sigma^2}{T_a(t)\gamma^2}.$$

By setting $\delta = \frac{\sigma^2}{T_a(t)\gamma^2}$, we can see $C_a(T_a(t), \delta) = \gamma = \sigma/\sqrt{\delta T_a(t)}$.

(c) (1 point) Let $\delta = 1/2$. At time step $t+1$, we have access to the count of samples for each arm, $T_a(t)$, and we have the mean reward $\hat{\mu}_{a,T_a(t)}$. Describe the UCB algorithm by filling the blank below.

At time step t, the UCB algorithm chooses arm A_{t+1} as

$$A_{t+1} = \underset{a\in[K]}{\operatorname{argmax}} \underline{\hspace{4cm}}.$$

Solution:

$$A_{t+1} = \underset{a\in[K]}{\operatorname{argmax}}\, \hat{\mu}_{a,T_a(t)} + \frac{\sqrt{2}\sigma}{\sqrt{T_a(t)}}.$$

(d) (3 points) Recall that UCB algorithm aims at getting sublinear pseudo-regret, where the pseudo-regret is defined as

$$R_t = t\mu_* - \mathbb{E}\left[\sum_{s=1}^{t} X_{A_s}\right]. \tag{4}$$

Here $\mu_* = \max_a \mathbb{E}[X_a]$ is the maximum expected mean among all rewards, and A_s is the choice of arm by the UCB algorithm at time s. By "sublinear," we mean $R_t/t \to 0$ as $t \to \infty$.

Consider the case of $K = 2$. Suppose that we have a wrong modeling assumption, and X_1 is distributed as

$$X_1 = \begin{cases} 0, & \text{with probability } 0.99, \\ 1000\sigma, & \text{with probability } 0.01. \end{cases} \tag{5}$$

Here $Var[X_1]$ is much larger than σ^2, which violates the assumption in (1). However we do not have the knowledge of $Var[X_1]$ and use the same algorithm derived in (b) and (c) with $\delta = 1/2$.

Find a distribution of the reward X_2 such that the UCB algorithm in (c) may fail to guarantee sublinear pseudo-regret. Justify your answer.

Solution: We let X_1 be distributed as

$$X_1 = \begin{cases} 0, & \text{with probability } 0.99, \\ 1000\sigma, & \text{with probability } 0.01, \end{cases} \tag{6}$$

and X_2 be a random variable which always takes value 5σ. We can see that $\mathbb{E}[X_1] > \mathbb{E}[X_2]$. So we shall choose arm 1 in the long run to get sublinear pseudo-regret. However, at the first two rounds when we pull both arm 1 and arm 2, with probability 0.99 we see 0 reward for arm 1, 5σ reward for arm 2. The corresponding confidence bound is $\sqrt{2}\sigma$. In the third round, we compare the mean + confidence bound for both arms, since $0 + \sqrt{2}\sigma < 5\sigma$, we will continue to pull arm 2 again and again without getting to pull arm 1. Thus overall we achieve linear regret with probability 0.99, which gives also linear pseudo-regret.

5. (10 points) **Stable matchings.**

(a) (3 points) Suppose there are two bands, b_1 and b_2, and two drummers, d_1 and d_2. Consider the following two matchings, M and M', of these bands and drummers:

$$\begin{aligned} M &= \{(b_1, d_1), (b_2, d_2)\} \\ M' &= \{(b_1, d_2), (b_2, d_1)\} \end{aligned}$$

Write down an example of preference lists for the bands and the drummers such that both of these matchings are stable.

Solution: Preference lists, in order of most preferred to least preferred, could be

$$\begin{aligned} b_1 &: d_1, d_2 \\ b_2 &: d_2, d_1 \\ d_1 &: b_2, b_1 \\ d_2 &: b_1, b_2 \end{aligned}$$

or

$$\begin{aligned} b_1 &: d_2, d_1 \\ b_2 &: d_1, d_2 \\ d_1 &: b_1, b_2 \\ d_2 &: b_2, b_1 \end{aligned}$$

(b) (3 points) In the same problem setting as Part (a), write down an example of preference lists for the bands and the drummers such that M is stable but M' is unstable.

Solution: Preference lists, in order of most preferred to least preferred, could be any of the following:

$b_1 : d_1, d_2$	$b_1 : d_1, d_2$	$b_1 : d_1, d_2$	$b_1 : d_1, d_2$
$b_2 : d_2, d_1$	$b_2 : d_2, d_1$	$b_2 : d_1, d_2$	$b_2 : d_1, d_2$
$d_1 : b_1, b_2$	$d_1 : b_1, b_2$	$d_1 : b_1, b_2$	$d_1 : b_1, b_2$
$d_2 : b_2, b_1$	$d_2 : b_1, b_2$	$d_2 : b_2, b_1$	$d_2 : b_1, b_2$

$$\begin{array}{lcl} b_1 : d_1, d_2 & \qquad & b_1 : d_2, d_1 \\ b_2 : d_2, d_1 & & b_2 : d_2, d_1 \\ d_1 : b_2, b_1 & & d_1 : b_1, b_2 \\ d_2 : b_2, b_1 & & d_2 : b_2, b_1 \end{array}$$

(c) (4 points) For a general matching problem, suppose there is a band b and a drummer d that have ranked each other first on their preference lists. Prove that in all stable matchings, b and d are matched together.

Hint: Use a proof by contradiction. That is, start by assuming there does exist a stable matching M where b and d are not matched to each other, and find a contradiction that arises from this assumption.

Solution: Assume there exists a stable matching M where b and d are not matched to each other. Then b prefers d to whomever b is matched to in M, and d prefers b to whomever d is matched to in M, so (b, d) are a blocking pair for M, which is a contradiction.

6. (10 points) **Grid World**

Consider the following grid representation of a game:

R_1	R_2	
	$\times$	`start`

where **start** represents our initial state and $\times$ is a state the agent cannot access. R_1 and R_2 represent two terminal states with rewards R_1 and R_2, respectively, where $R_1 > 0$ and $R_2 > 0$. At each state, we have four possible actions (up, down, left and right). Suppose the transitions are deterministic, meaning that an action in a specific direction always moves us in that direction. If the state that results from the proposed action is not accessible, then the agent remains at the same state. For instance, if the agent chooses action **left** when they are at state **start**, they will remain at state **start**. Recall that for any given $\gamma \in (0,1)$, **the optimal value function** $V^*(s)$ is defined as:

$$V^*(s) = \max_{a \in A} \sum_{s' \in S} \mathbb{P}(s' \mid s, a) \left[R(s, a, s') + \gamma V^*(s') \right],$$

and the optimal Q-function $Q^*(s, a)$ is defined as:

$$Q^*(s, a) = \sum_{s' \in S} \mathbb{P}(s' \mid s, a) \left[R(s, a, s') + \gamma V^*(s') \right].$$

(a) (2 points) For this part only, assume $R_1 = 10$, $R_2 = 2$. Compute the optimal value function $V^*(s)$ for all states s, when (i) $\gamma = 0.9$; (ii) $\gamma = 0.1$. You don't need to justify your answers. You only need to complete the following two grid representations with the values you computed. We already filled out some values for you. You do not need to simplify your answers arithmetically.

N/A	N/A	
10	N/A	

$\gamma = 0.9$

N/A	N/A	
10	N/A	

$\gamma = 0.1$

Solution:

N/A	N/A	10×0.9^5
10	N/A	10×0.9^4
10×0.9	10×0.9^2	10×0.9^3

$\gamma = 0.9$

N/A	N/A	2
10	N/A	0.2
1	0.1	0.02

$\gamma = 0.1$

(b) (2 points) Let's consider what would happen if $\gamma > 1$. For this part only, assume $\gamma > 1$. Describe what your optimal moves would be if $R_1 > R_2 > 0$. [Hint: Do you want to end the game early?]

Solution: Optimal actions are determined by the optimal value functions. Since $\gamma > 1$, the optimal value function gets larger for every step. Hence, your optimal moves would be moving across different states for as long as you can before you reach R_1. In other words, if there is no constraint on the total number of steps, you will never terminate.

(c) (2 points) Give an expression for $V^*(\text{start})$ in terms of R_1, R_2 and γ, where $R_1 > 0$, $R_2 > 0$ and $\gamma \in (0, 1)$. [Hint: You might use the function $\max(a, b) = \begin{cases} a, & a \geq b \\ b, & a < b \end{cases}$ in your expression.]

Solution: Since the transitions are deterministic and the only states with rewards are the terminal states, let the state above **start** be **start-up** and the state below **start** be **start-below** we have:

$$\begin{aligned} V^*(start) &= \max_{a \in A} \gamma V^*(s') \\ &= \gamma \max(V^*(\text{start-up}), V^*(\text{start-below})) \\ &= \gamma \max(\max(\gamma^5 R_1, R_2), \max(\gamma^3 R_1, \gamma^2 R_2))) \\ &= \gamma \max(\gamma^3 R_1, R_2) \end{aligned} \tag{7}$$

(d) (2 points) For this part only, assume $\gamma = 0.9$ and $R_2 = 10$. What values can R_1 take on such that $Q^*(\text{start}, \text{up}) > Q^*(\text{start}, \text{down})$?

Solution:

$$Q^*(\text{start}, \text{up}) = \gamma \max(\gamma^5 R_1, R_2) = \gamma \max(0.9^5 R_1, 10)$$

$$Q^*(\text{start}, \text{down}) = \gamma \max(\gamma^3 R_1, \gamma^2 R_2) = \gamma \max(0.9^3 R_1, 10 \cdot 0.9^2)$$

Therefore, $Q^*(\text{start}, \text{up}) > Q^*(\text{start}, \text{down})$ implies that

$$\max(0.9^5 R_1, 10) > \max(0.9^3 R_1, 10 \cdot 0.9^2)$$

Since $R_1 > 0$, we have $0.9^5 R_1 < 0.9^3 R_1$. Hence, in order for the inequality above to hold, we must have that $\max(0.9^5 R_1, 10) = 10$ and $10 > \max(0.9^3 R_1, 10 \cdot 0.9^2)$. Therefore,

$$10 \geq 0.9^5 R_1$$

$$10 > 0.9^3 R_1$$

This implies that $R_1 < 10/(0.9)^3 = 13.7$

7. (7 points) **Miscellaneous**

This problem consists of several short-answer questions.

(a) (2 points) Draw a computation graph for the following expression:

$$f(x_1, x_2, y) = y \cdot (x_1 \cdot x_2 + x_2),$$

where x_1, x_2 and y are real numbers.

Solution:

(b) (3 points) Suppose we want to model the number of visitors X_t to a website over time. Since the overall trend seems to be increasing roughly exponentially over time, we decide to fit a Poisson model with an exponential link function; i.e., $X_t \sim \text{Poisson}(e^{\lambda t})$.

(*i*) (1 point) What are some reasons the data might have greater variance than the Poisson model can predict? List at least 2 reasons.

Solution:

(*ii*) (2 points) How does such model inadequacy affect the uncertainty estimates and what is one way to try to address this?

Solution:

(c) (2 points) We observe a large sample $n = 10^9$ of real-valued data points, $X_1, \ldots, X_n$. We compute the sample mean, $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, and we also want to report a standard error. Suppose that, instead of running a bootstrap where we repeatedly resample n observations with replacement out of the original n, to save time we repeatedly *subsample* $m = 10^3$ observations (with replacement), and report the standard deviation of the averages across all subsamples. Is this subsampling procedure correct, or it will it overestimate or underestimate our uncertainty? Briefly justify your answer with 1-2 sentences.

Solution: