Data 102, Spring 2022 Midterm 2

- You have 110 minutes to complete this exam. There are 7 questions, totaling 40 points.
- You may use one 8.5×11 sheet of handwritten notes (front and back), and the provided reference sheet. No other notes or resources are allowed.
- You should write your solutions inside this exam sheet.
- You should write your name and Student ID on every sheet (in the provided blanks).
- Make sure to write clearly. We can't give you credit if we can't read 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.
- You may, without proof, use theorems and facts that were given in the discussions or lectures, **but please cite them**.
- There will be no questions allowed during the exam: if you believe something is unclear, clearly state your assumptions and complete the question.
- Unless otherwise stated, no work or explanations will be graded for multiple-choice questions.
- Unless otherwise stated, you must show your work for free-response questions in order to receive credit.

Last name	
First name	
Student ID (SID) number	
Calcentral email (`@berkeley.edu`)	
Name of person to your left	
Name of person to your right	
Row (main exam room only)	

Honor Code: I will respect my classmates and the integrity of this exam by following this honor code. I affirm that all of the work submitted here is my original work, and I did not collaborate with anyone else on this exam.

Signature: ______________________________

1. (7 points) For each of the following, determine whether the statement is true or false. For this question, no work will be graded and no partial credit will be assigned.

   (a) (1 point) In a k-armed bandit setting, the first k rounds of the UCB algorithm pull different arms.

   ◯ A. TRUE ◯ B. FALSE

   (b) (1 point) Consider a k-armed bandit setting where one arm has an expected reward of 1 and all other arms have expected reward of 0. In this setting, the pseudo-regret of any algorithm is equal to the number of rounds where it selects an arm with expected reward of 0.

   ◯ A. TRUE ◯ B. FALSE

   (c) (1 point) The stable unit treatment value assumption (SUTVA) requires that the potential outcomes are conditionally independent of the treatment.

   ◯ A. TRUE ◯ B. FALSE

   (d) (1 point) When evaluating the causal relationship between a treatment Z and outcome Y, suppose that variable T is not directly caused by Z or Y (in other words, it is not a child of either in the causal DAG). T can still be a collider.

   ◯ A. TRUE ◯ B. FALSE

   (e) (1 point) Random Forests are more interpretable than linear regression because their decisions can be expressed as majority votes of other classifiers.

   ◯ A. TRUE ◯ B. FALSE

   (f) (1 point) An advantage of using random forests over decision trees is that they have lower variance.

   ◯ A. TRUE ◯ B. FALSE

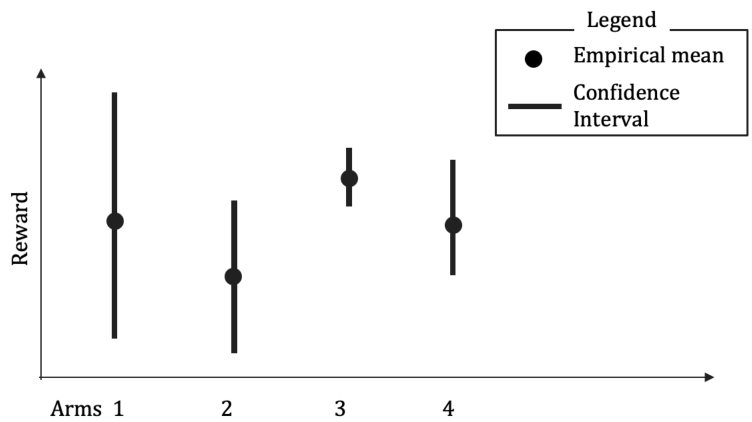
   (g) (1 point) In order to have good prediction accuracy with a two-layer neural net with prediction $y = W_1\sigma(W_2x + b_2) + b_1$, we must assume that the data are generated using that same process.

   ◯ A. TRUE ◯ B. FALSE

Name: ______________________ SID: ____________

2. (2 points) Multi-Armed Bandits

For a 4-armed bandit, you run the UCB algorithm. The diagram below shows a snapshot of the empirical rewards and their confidence intervals at the beginning of round $t = 10$. Answer the following questions using this diagram (no justification is needed). The confidence intervals were computed using the formula from lecture: $[\hat{\mu}_i(t) - \sqrt{C_t/T_i(t)}, \hat{\mu}_i(t) + \sqrt{C_t/T_i(t)}]$, where $T_i(t)$ is the number of time arm i is pulled before time t.



(a) (1 point) Which arm has been pulled most often by UCB up to time $t = 10$?

Solution: Arm 3, because the interval is narrowest.
1 point for correct choice, no justification needed.

(b) (1 point) Which arm will be pulled by UCB at round $t = 10$?

Solution: Arm 1, because it has the highest upper confidence bound.
1 point for correct choice, no justification needed.

3. (6 points) Consider an MDP with five adjacent states. The far left and far right states are terminal states, and the center state is the start state. The reward for entering each state is as follows:

10	1	0	-4	1000

There are two actions, left and right. Assume that the actions are deterministic: they move to the left one square and to the right one square respectively. Unless otherwise stated, assume that the discount factor is $\gamma = 0.5$.

(a) (2 points) Compute the optimal value function for the starting state, $V^*(\text{start})$. You do not need to simplify your answer arithmetically.

Hint: you don't need to compute the value function for all states.

Solution:

$$V^*(\text{start}) = -4 + 1000 * 0.5 = 496$$

(b) (2 points) **For this part only, we change the discount factor to $\gamma = 1$.** In this case, what is the optimal policy? Give your answer by filling in the empty squares below with a left arrow for the left action and a right arrow for the right action. Briefly justify your answer.

$\times$				$\times$

Solution:

$\times$	$\rightarrow$	$\leftarrow$	$\leftarrow$	$\times$

Because future rewards are not discounted, the agent can obtain infinite reward by going back and forth between the 1 state and the 0 state: this is higher than the reward of 996 that it would obtain by moving right.

Note: Because the optimal policy will never visit the state directly to the right of the starting state, we awarded full points to those that left that cell blank.

(c) (2 points) Suppose we use Q-learning. We plan on collecting trajectory data by running a greedy policy π in this environment many times:

$$\pi(s) = \underset{a}{\operatorname{argmax}} \max_{s'} R(s, a, s')$$

If we apply a Q-learning update for every data point collected using this policy, will our estimate converge to the optimal Q function? Why or why not?

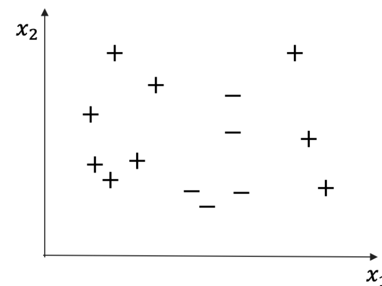
Solution: It will not, because the greedy policy always moves left.

1 point for "will not"

1 point for something along the lines of "the greedy policy never moves right.'

4. (5 points) Neural Networks and Backpropagation

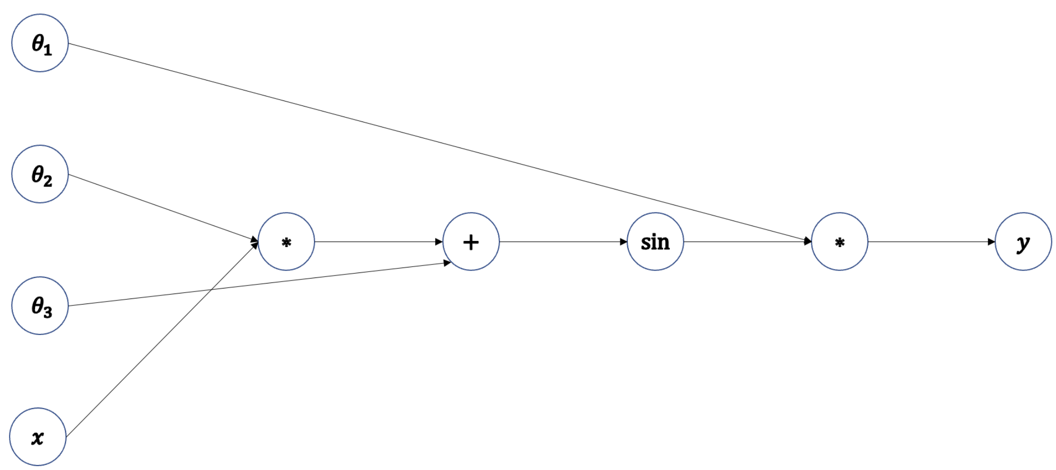
(a) (2 points) Consider a neural network with a sigmoid activation function at every hidden node and at the output (similar to what you saw in lecture). In order to achieve perfect classification accuracy on the dataset shown below, what is the smallest number of hidden layers needed? Briefly justify your answer. You don't need to compute the weights of the network; just explain how many hidden layers and why.



Solution: One hidden layer is necessary and sufficient. The positive and negative classes are not linearly separable, because the positive class is not a convex region. Therefore, 1 hidden layer is necessary. One layer is also sufficient: Each node in the hidden layer captures one of halfpsaces separating the negative points from one of the adjacent sets of positive points and the last layer combines the two hidden nodes.

(b) (1 point) Consider the following function to predict y from x: $y = \theta_1 \sin(\theta_2 x + \theta_3)$. Draw a computational graph for this function.

Solution:



(c) (2 points) When computing the gradients of the function from the previous part (with respect to θ_1, θ_2, and θ_3), which of the following is equal to the number of partial derivatives that must be computed when applying backpropagation? Choose the single best answer by filling in the circle next to it.

- ◯ A. The number of nodes in the graph
- ◯ B. The number of edges in the graph
- ◯ C. The number of nodes on the path from x to y
- ◯ D. The number of edges on the path from x to y
- ◯ E. $n_1 + n_2 + n_3$, where n_i is the number of nodes on the path from θ_i to y.
- ◯ F. $e_1 + e_2 + e_3$, where e_i is the number of edges on the path from θ_i to y.

Solution:

The correct answer is B, the number of edges: in backpropagation, we compute gradients at each edge starting at the output and working back, multiplying as we go.

We accepted A as well, because for this particular graph, if students count $\partial y/y$ then the number of nodes will end up equal to the edges and will be an acceptable solution.

Name: ______________________________ SID: ______________

5. (9 points) Matei is investigating whether watching Netflix causes lower grades. He finds data from a study that followed 1000 students for a semester, with the average number of hours they watched Netflix per week (W), their GPA (G), and whether they had high-speed internet (H). A separate study (not shown here) found that having high-speed internet caused students to earn higher grades.

W	G	H
1	3.4	1
15	3.7	1
15	2.9	0

(a) (2 points) **For this part only**, assume that the data come from a randomized experiment, where each student was randomly assigned to watch either 1 or 15 hours per week. Matei wants to know the average treatment effect (ATE) of watching Netflix for 15 hours/week versus 1 hour/week on GPA. Assuming that the dataframe above is called `netflix`, either write 1-4 lines of code to compute an unbiased estimate for the ATE, or explain why this is impossible from the information given.

Solution:
```
treatment_mean = netflix[netflix['W'] == 15]['G'].mean()
control_mean = netflix[netflix['W'] == 1]['G'].mean()
treatment_mean - control_mean
```

For the remainder of the question, we assume the data come from an observational study. Matei creates a new column W_b that is 0 if $W \leq 10$ and 1 if $W > 10$. He calls these categories "light watching" and "heavy watching" respectively, and wants to investigate whether heavy watching causes lower GPA than light watching.

(b) (2 points) Matei computes the average GPA for heavy watching students (3.4) and for light watching students (3.1), and incorrectly concludes that **heavy watching causes students' GPA to increase by 0.3 compared to light watching**.

Which of the following statements are reasons that Matei's conclusion (bolded above) is incorrect?

☐ A. High-speed internet is a confounder for the treatment (light/heavy watching) and the outcome (GPA).

- ☐ B. High-speed internet is a <u>collider</u> for the treatment (light/heavy watching) and the outcome (GPA).
- ☐ C. Hours of Netflix watched and GPA could have a nonlinear relationship.
- ☐ D. Choosing a threshold of 10 hours violates the stable unit treatment value assumption (SUTVA).

Solution: High-speed internet can have a causal effect on both treatment (how much Netflix someone watches) and outcomes (their GPA), so it is a confounder. This means that in an observational study, the difference in means is not an unbiased estimate for the ATE.

Name: ______________________ SID: __________

(c) (3 points) Matei finds out that Netflix randomly offered a free promotion (P) to half the students in the data: they were given \$100 if they watched more than 10 hours a week during the semester.

Matei decides to use the promotion as an instrumental variable (IV). For each assumption required to use IVs, explain why the promotion does or doesn't satisfy it. You should draw a causal DAG with the variables W_b, H, G, and P to support your answer.

Solution:

```
            H
           / \
          v   v
   P ->  Wb -> G
```

1. The promotion only affects the outcome through the treatment (no direct effect): this is true because the extra \$100 will not have any impact on grades. We accepted properly justified answers arguing the opposite (e.g., spending the money on a tutor).
2. The promotion affects the treatment: if you get \$100 as an incentive, you are much more likely to watch more.
3. The promotion is independent of the confounder (high speed internet): this is true. We accepted answers saying this wasn't true if students spent their money on buying high-speed internet.

(d) (2 points) Matei decides to use inverse propensity weighting (IPW). He computes the IPW estimate for the ATE, removing points with propensity scores below 0.1

and 0.9 (as in Lab 8). Assume all necessary assumptions for using IPW are satisfied, and that he implements it correctly.

Now, he wants to quantify the uncertainty in his estimate. Is bootstrap appropriate? Select all answers that apply.

- ☐ A. Yes, because the IPW estimate is not sensitive to removal/inclusion of a small number of points.
- ☐ B. No, because the IPW estimate is not sensitive to removal/inclusion of a small number of points.
- ☐ C. Yes, because the random sampling in bootstrap eliminates the effect of any confounding variables.
- ☐ D. No, because the random sampling in bootstrap eliminates the effect of any confounding variables.
- ☐ E. No, because the number of parameters in the IPW estimate is equal to the number of data points.
- ☐ F. None of the above

Solution: The IPW estimator, especially after removing points with predicted propensity scores outside of 0.1 to 0.9, is not sensitive to the removal/inclusion of specific points.

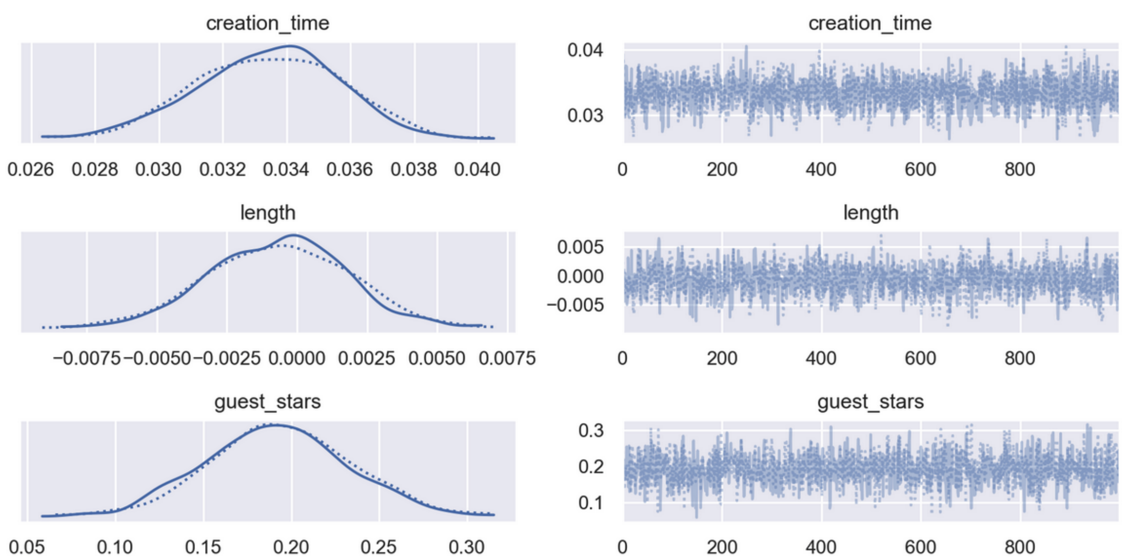
The random sampling in bootstrap has nothing to do with the randomness required for a randomized experiment.

Name: ______________________________ SID: ______________

6. (6 points) Kaiea wants to predict the number of views that his YouTube videos will get. For each of his past videos, he collects the number of views, the time he spent creating it (in hours), its length (in minutes), and the number of guest stars. He builds the following dataframe, and applies negative binomial regression.

	views	creation_time	length	guest_stars
0	288	27.3	54.3	0
1	638	40.8	59.7	0
2	3022	44.8	17.2	4

(a) (3 points) **Bayesian**: He implements the model correctly in PyMC3 and visualizes the coefficient posterior distributions (not all are shown):



Which of the following are valid conclusions from these plots? Select all answers that apply.

☐ A. Using a 99% credible interval for the coefficient of length, according to the model, longer videos are associated with more views.

☐ B. Using a 99% confidence interval for the coefficient of length, according to the model, longer videos are associated with more views.

☐ C. According to the model, each additional guest star is associated with an increase of between 0.05 and 0.35 additional views.

☐ D. According to the model, adding one guest star is associated with the same effect on the number of views as spending an additional T hours making the video, and T is between 5 and 7.

Solution: A/B: The credible interval for length (since this is a Bayesian analysis) includes positive and negative values, so we can't determine association between length and views.

C: Each additional guest star is associated with a multiplicative increase of $e^{0.18}$ views, not an additive one.

D: The coefficient for guest stars is about 6 times the coefficient for creation time.

(b) (3 points) **Frequentist**: He implements the model correctly in statsmodels and obtains the following results:

```
                 Generalized Linear Model Regression Results
==============================================================================
Dep. Variable:                  views   No. Observations:                  112
Model:                            GLM   Df Residuals:                      108
Model Family:        NegativeBinomial   Df Model:                            3
Link Function:                    log   Scale:                          1.0000
Method:                          IRLS   Log-Likelihood:                -819.63
Date:                Sat, 02 Apr 2022   Deviance:                       26.508
Time:                        13:16:51   Pearson chi2:                     30.5
No. Iterations:                     6
Covariance Type:            nonrobust
=================================================================================
                    coef    std err          z      P>|z|      [0.025      0.975]
---------------------------------------------------------------------------------
const             5.0187      0.304     16.488      0.000       4.422       5.615
creation_time     0.0334      0.005      6.983      0.000       0.024       0.043
length           -0.0007      0.005     -0.141      0.888      -0.011       0.009
guest_stars       0.1905      0.087      2.197      0.028       0.021       0.361
=================================================================================
```

Which of the following are valid conclusions from these results? **Note that the choices are the same as in the previous question, except for the numbers in choice C.** Select all answers that apply.

☐ A. Using a 99% credible interval for the coefficient of length, according to the model, longer videos are associated with more views.

☐ B. Using a 99% confidence interval for the coefficient of length, according to the model, longer videos are associated with more views.

☐ C. According to the model, each additional guest star is associated with an increase of between 0.021 and 0.361 additional views.

☐ D. According to the model, adding one guest star is associated with the same effect on the number of views as spending an additional T hours making the video, and T is between 5 and 7.

Solution: A/B: The confidence interval for length (since this is a frequentist analysis) includes positive and negative values, so we can't determine association between length and views.

C: Each additional guest star is associated with a multiplicative increase of $e^{0.19} \approx 1.2$ views, not an additive one.

D: The coefficient for guest stars is about 6 times the coefficient for creation time.

7. (5 points) Ritvik takes what he learned about concentration inequalities in Data 102 and uses it to become a professional poker player. He starts with \$1000, and plays 50 rounds of poker. Let X_i be a random variable representing the number of dollars earned (or lost) after the i^{th} round of a poker game. For example, if he loses \$80 on the second round, then $X_2 = -80$. Let $S_{50} = \sum_{i=1}^{50} X_i$ represent his total profit after 50 rounds of poker. Ritvik assumes that the X_i are i.i.d., and that $\mathbb{E}[X_i] = 2$. You should take these assumption as correct throughout this question.

   (a) (2 points) Ritvik wants to know the probability of winning big, and decides to use Markov's inequality with S_{50}. He computes that $\mathbb{E}[S_{50}] = 100$, and writes down the following calculation: $\Pr(S_{50} \geq 1000) \leq 100/1000$, but Ruhi points out that this isn't correct. Explain to Ritvik why and where he made a mistake.

   Solution: "X_i can be a negative random variable, so S_{50} can be a negative random variable, and Markov's inequality only applies to non-negative random variables."

   (b) (3 points) Ritvik decides to play at a table with limits on the bets, so that he can assume X_i is bounded between -10 and 10. Given the information so far, find the largest p for which $\Pr(S_{50} > 75) \geq p$ using what you learned in Data 102.

   Solution: Just as in the homework, we need to adjust the desired probability before we can apply the appropriate concentration inequality. We know that

$$P(S_{50} \geq 75) = 1 - P(S_{50} \leq 75)$$

   by taking the complement. Hoeffding's inequality (the best we can use, since we know each X_i is bounded) states that

$$P\left(\frac{1}{n}\sum(X_i - \mu) \leq t\right) \leq \exp\left\{\frac{-2nt^2}{(b-a)^2}\right\}$$

In this case, we can rearrange terms in the probability $P(S_{50} \leq 75)$ to match the form of Hoeffding's inequality:

$$P\left(\frac{1}{50}\sum(X_i - 2) \leq -1/2\right) \leq \exp\left\{-\frac{2 \times 50 \times (1/2)^2}{20 \times 20}\right\} = \exp(-1/16)$$

Combining this with our first statement,

$$P(S_{50} \geq 75) \geq 1 - \exp(-1/16)$$