CS 361 Sample Final Exam

NAME: SOLUTIONS

NETID:

CIRCLE YOUR DISCUSSION:

Thu 2-3 **Thu 4-5** **Fri 10-11** **Fri 11-12**

- Be sure that your exam booklet has 8 pages including this cover page
- Make sure to write your name exactly as it appears on your i-card
- Write your netid and circle your discussion section on this page
- **Show your work**
- This is a closed book exam
- You are allowed one handwritten 8.5 x 11-inch sheet of notes (both sides)
- You may **not** use a calculator or any other electronic device
- Turn off your phone and store it in your backpack
- Store away any other electronic devices including earphones and smartwatches
- Absolutely no interaction between students is allowed
- Use backs of pages for scratch work if needed
- Show your i-card when handing in your exam

Problem	1	2	3	4	5	6	7	Total
Possible	30	30	30	30	30	30	20	200
Score								

Problem 1 (30 pts)

1. (15 points) You buy a carton of a dozen eggs and find that exactly one has a double yolk. What is the variance of the number of yolks per egg in this carton? Draw a box around your answer.

mean $= 13/12$

$$\text{var} = \frac{1}{12}\left(11\left(1-\frac{13}{12}\right)^2 + 1\left(2-\frac{13}{12}\right)^2\right)$$

$$= \frac{1}{12}\left(\frac{11}{144}+\frac{121}{144}\right) = \frac{1}{\cancel{12}}\cdot\frac{\cancel{132}^{\,11}}{144} = \boxed{\frac{11}{144}}$$

2. (15 points) Suppose dataset $\{\mathbf{x}\}$ has the covariance matrix shown below. Calculate the correlation coefficient $\text{corr}(\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}\})$. Draw a box around your answer.

$$\text{Covmat}(\{\mathbf{x}\}) = \begin{bmatrix} 9 & -3 \\ -3 & 4 \end{bmatrix}$$

$$\text{corr}(\{x^{(1)}, x^{(2)}\}) = \frac{\text{cov}(\{x^{(1)}\}, \{x^{(2)}\})}{\text{std}(\{x^{(1)}\})\,\text{std}(\{x^{(2)}\})}$$

$$= \frac{-3}{\sqrt{9}\sqrt{4}}$$

$$= \boxed{-0.5}$$

Problem 2 (30 pts)

1. (15 points) The word game Scrabble contains 100 tiles, of which there are exactly 12 E, 4 U, 4 S and 2 C tiles. If you draw 7 tiles randomly from the original 100 without replacement, what is the probability that you can arrange the tiles to form the word SUCCESS (without using blank tiles)? You may use factorials and choose notation in your answer. Draw a box around your answer.

$$\boxed{\frac{\binom{12}{1}\binom{4}{1}\binom{4}{3}\binom{2}{2}}{\binom{100}{7}}}$$

no. of ways to make SUCCESS

no. of ways to draw 7 tiles

2. (15 points) You have been invited to a fishing game with a twist. If you catch no fish in an hour, you will pay \$100. If you catch 1 fish in that hour, you will win \$40. If you catch 2 or more fish, you will win \$80. Some research tells you that catching fish in this pond is a Poisson process with intensity of 3 fish per hour. Should you play the game? Justify your answer with calculations. You may approximate e^{-3} as 0.05.

$$P(0) = \frac{e^{-3}3^0}{0!} = 0.05 \qquad P(1) = \frac{e^{-3}3^1}{1!} = 0.15$$

$$P(2 \text{ or more}) = 1 - 0.05 - 0.15 = 0.8$$

$$\text{Expected winnings} = -100(0.05) + 40(0.15) + 80(0.8)$$
$$= -5 + 6 + 64 = 65 > 0$$

The expected winnings are \$65, so I should play the game

Problem 3 (30 pts)

1. (15 points) Suppose I offer to give you a prize if I toss a fair coin n times and it comes up heads k of those times. If I require that $n \geq 2$, what values of n and k should you choose to maximize your chance of winning? What is your maximum probability of winning? Draw a box around your answer.

$$P(\text{winning}) = \binom{n}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{n-k} = \binom{n}{k}\left(\frac{1}{2}\right)^n$$

This function is maximized at $\boxed{n=2,\ k=1}$

and the probability is $\boxed{\frac{1}{2}}$

2. (15 points) You asked 5 of your classmates what their scores were on Midterm 2 and then used Student's t-distribution to calculate that the 95% confidence interval for the population mean score is [96, 144]. Why is it wrong to conclude that the population mean score falls in the interval [96, 144] with probability 95%?

In the frequentist point-of-view, the population mean score is a number, not a random variable.

[The correct statement is: if you repeat this sampling randomly, the resulting confidence interval will contain the population mean score with probability 95%.]

not required for full points

Problem 4 (30 pts)

1. (15 points) You hypothesize that the average product rating on Amazon is 4 stars. From a sample of 100 items, you find a sample mean of 4.4 stars and a sample (unbiased) standard deviation of 0.1 stars. Assess the evidence against the claim.

$$\text{stderr} = \frac{0.1}{\sqrt{100}} = 0.01$$

$$\text{test statistic} = \frac{4.4-4}{0.01} = 40$$

$$\text{p-value} = 1 - \frac{1}{\sqrt{2\pi}} \int_{-40}^{40} \exp\left(-\frac{x^2}{2}\right) dx$$

$$\simeq 0$$

Since the p-value is close to 0, we reject the hypothesis

2. (15 points) Last week you flipped a coin 10 times. The coin came up on the same side all 10 times, but you've forgotten if it was all heads or all tails. Write down the likelihood function $L(\theta)$ for the coin's probability of coming up heads. Draw a box around your answer.

observation

$$L(\theta) = P(D|\theta)$$

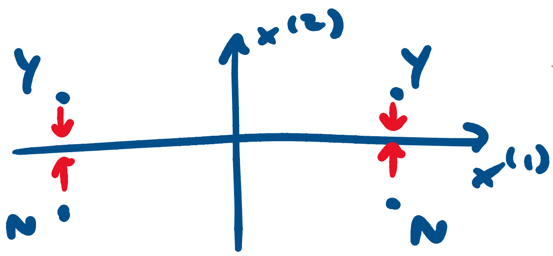
$$= P(\text{all heads}|\theta) + P(\text{all tails}|\theta)$$

since these events are disjoint

$$= \boxed{\theta^{10} + (1-\theta)^{10}}$$

Problem 5 (30 pts)

1. (15 points) Suppose you want to train a classifier using the training data below. Explain why you should **not** start by projecting the dataset $\{\mathbf{x}\}$ on to its first principal component.



$\mathbf{x}^{(1)}$	$\mathbf{x}^{(2)}$	label
5	1	Y
−5	1	Y
5	−1	N
−5	−1	N

The 1st principal component is $x^{(1)}$. Projecting on to $x^{(1)}$ would superimpose the Y & N points, making it impossible to classify them.

2. (15 points) Given a dataset $\{0, 2, 4, 6, 24, 26\}$, initialize the k-means clustering algorithm with 2 cluster centers $c_1 = 3$ and $c_2 = 4$. What are the values of c_1 and c_2 after one iteration of k-means? What are the values of c_1 and c_2 after the second iteration of k-means? Draw a box around your answer.

First iteration

c_1 gets $\{0,2\}$

c_2 gets $\{4,6,24,26\}$

$c_1 := 1$

$c_2 := 15$

Second iteration

c_1 gets $\{0,2,4,6\}$

c_2 gets $\{24,26\}$

$c_1 := 3$

$c_2 := 25$

Problem 6 (30 pts)

1. (15 points) Suppose you want to train a linear regression model $y = \beta_1 x + \beta_2$ using the training data below. Write down X and $\mathbf{y}$ so that the least-squares estimate for the coefficients is

$$\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = (X^T X)^{-1} X^T \mathbf{y}$$

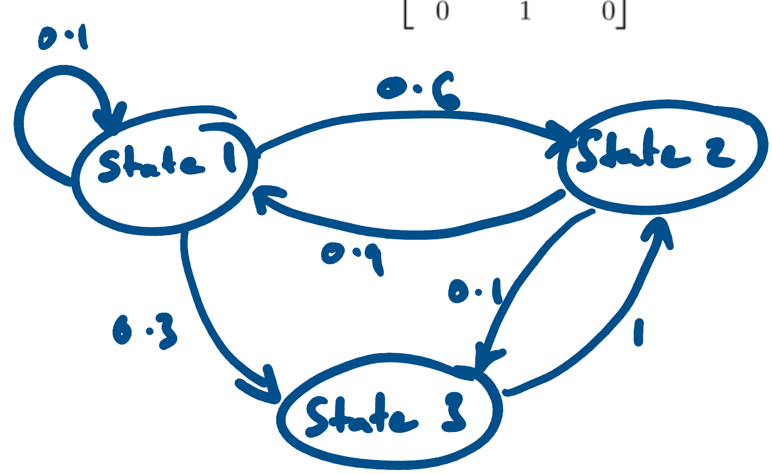
Draw a box around your answer.

x	y
1	17
2	15
3	10
4	2

$$\boxed{X = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix} \qquad y = \begin{bmatrix} 17 \\ 15 \\ 10 \\ 2 \end{bmatrix}}$$

2. (15 points) Draw the state transition diagram (i.e. the directed graph) for the Markov chain with transition probability matrix P given below. Make sure to label the edges of the graph with the appropriate probabilities.

$$P = \begin{bmatrix} 0.1 & 0.6 & 0.3 \\ 0.9 & 0 & 0.1 \\ 0 & 1 & 0 \end{bmatrix}$$



Note: when a dataset contains only 2 points, the mean is the midpoint and the std dev. is half the distance between the points

Problem 7 (20 pts)

1. (20 points) For the training data below, model each conditional probability of the form $P(\mathbf{x}^{(i)}|y)$ as a normal distribution. Then use the naïve Bayes assumption to write an expression for

$$\frac{P(y=0|\mathbf{x})}{P(y=1|\mathbf{x})}$$

Draw a box around your answer.

$\mathbf{x}^{(1)}$	$\mathbf{x}^{(2)}$	y
4	7	0
−4	5	0
2	10	1
10	4	1

$P(y=0) = 1/2$

$P(y=1) = 1/2$

$P(x^{(1)}|y=0) \sim \mathcal{N}(0,4)$

$P(x^{(1)}|y=1) \sim \mathcal{N}(6,4)$

$P(x^{(2)}|y=0) \sim \mathcal{N}(6,1)$

$P(x^{(2)}|y=1) \sim \mathcal{N}(7,3)$

$$\frac{P(y=0|x)}{P(y=1|x)} = \frac{P(x^{(1)}|y=0)\,P(x^{(2)}|y=0)\,\cancel{P(y=0)}}{P(x^{(1)}|y=1)\,P(x^{(2)}|y=1)\,\cancel{P(y=1)}}$$

$$= \frac{\frac{1}{\cancel{\sqrt{2\pi}}\,4}\exp\left(-\frac{(x^{(1)}-0)^2}{2\cdot 4^2}\right)\frac{1}{\cancel{\sqrt{2\pi}}\,1}\exp\left(-\frac{(x^{(2)}-6)^2}{2\cdot 1^2}\right)}{\frac{1}{\cancel{\sqrt{2\pi}}\,4}\exp\left(-\frac{(x^{(1)}-6)^2}{2\cdot 4^2}\right)\frac{1}{\cancel{\sqrt{2\pi}}\,3}\exp\left(-\frac{(x^{(2)}-7)^2}{2\cdot 3^2}\right)}$$

$$= \boxed{3\,\frac{\exp\left(-\frac{(x^{(1)})^2}{32}\right)\exp\left(-\frac{(x^{(2)}-6)^2}{2}\right)}{\exp\left(-\frac{(x^{(1)}-6)^2}{32}\right)\exp\left(-\frac{(x^{(2)}-7)^2}{18}\right)}}$$