CSCI 3022
Midterm Exam
Spring 2018

Name: $\underbrace{\hspace{5cm}}_{\text{By writing my name I promise to abide by the Honor Code}}$

Read the following:

- **RIGHT NOW**! Write your name on the top of your exam.
- You are allowed one 8 1/2 × 11in sheet of **handwritten** notes (both sides). No magnifying glasses!
- You may use a calculator provided that it cannot access the internet or store large amounts of data.
- You may **NOT** use a smartphone as a calculator.
- Clearly mark answers to multiple choice questions on the provided answer line.
- Mark only one answer for multiple choice questions. If you think two answers are correct, mark the answer that **best** answers the question. No justification is required for multiple choice questions.
- If you do not know the answer to a question, skip it and come back to it later.
- For free response questions you must clearly justify all conclusions to receive full credit. A correct answer with no supporting work will receive no credit.
- If you need more space for free-response questions, there are blank pages at the end of the exam. If you choose to use the extra pages, make sure to **clearly** indicate which problem you are continuing.

Problem	Points	Score
1	3	
2	3	
3	3	
4	3	
5	3	
6	3	
7	3	
8	3	
9	3	
10	3	
11	3	
12	3	
13	3	

Problem	Points	Score
Mult. Choice Total	39	
14	20	
15	20	
16	20	
For luck!	1	1
Total	100	

1. (3 points) Suppose you are a CU student and you have built a website that allows other students to put their faces in some sort of online book, primarily so that you can serve them advertisements and weird news. Your app is currently rolled out to 5,000 CU students, 1,000 Mines students, and 4,000 CSU students. You are designing a survey to determine voter registration status by asking users if they are registered democrats, registered republicans, registered but with no party affiliation, or unregistered. Which of the following is the variable of interest?

   A. Voter registration status
   B. The percentage of democrats
   C. A stratified sample
   D. College students from CU, Mines, and CSU
   E. College students from CU, Mines, and CSU, but in proportion to their numbers

1. A

2. (3 points) What is Q1 (the first quartile) for the following dataset?

$$[39, 7, 6, 40, 42, 41, 49, 43, 20\pi, 14, 36]$$

   A. 23
   B. 15
   C. 25
   D. 6
   E. In this case, Q1 is undefined since π is irrational.

$14+36 = \frac{50}{2} = 25$

39, 7, 6, 40, 42, 41

6, 7, 14, 36, 39, 40, 41, 42, 43, 20π

2. C

3. (3 points) Let A and B be events in a sample space Ω. Suppose that the probability that A occurs is 0.4, the probability that B occurs is 0.7, and the probability that *neither A nor B occurs* is 0.2. What is the probability that both A **AND** B occur?

   A. 0.08
   B. 0.28
   C. 0.3
   D. 0.8
   E. None of the Above

$P(A) = 0.4, P(B) = 0.7 \quad P(A^c) \cap P(B^c) = 0.2$

$P(A \cup B) + P(A \cup B)^c = 1$

$P(A \cup B) = 1 - P(A \cup B)^c$

$= 1 - 0.2 = 0.8$

3. ____

$P(A^c) \cap P(B^c) = P(A \cup B)^c$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$0.8 = 0.4 + 0.7$

4. (3 points) You are stuck in a YouTube clickhole, late on a Thursday night, and cannot stop yourself from clicking Next...Next...Next...on all the videos. Homework is due for CSCI 3022 the next day, but you simply cannot get enough of these amazing clips of river otters playing in the snow. What a bunch of goofballs! A new notification appears on your mobile, and it says:

$$f(x) = kx - x^2 \text{ for } x = 1, 2, 3 \text{ and } f(x) = 0 \text{ for any other integer value of } x$$

What value or values of k make $f(x)$ a valid probability mass function?

$$\int_1^3 (kx - x^2)\,dx = \frac{kx^2}{2} - \frac{x^3}{3}\Big|_1^3$$

A. 3

B. 15/6

C. 6/15

D. $15/6, -15/6$

E. No such value of k exists.

4. E

5. (3 points) The Pareto distribution with parameter $\alpha > 0$ is often used to model the size of cities. It's probability density function is given by $f(x) = \alpha/x^{\alpha+1}$ for $x \geq 1$ and 0 otherwise. What is the cumulative distribution function for the Pareto distribution?

A. $F(x) = x^\alpha$ for $x \geq 1$

B. $F(x) = 1 - 1/x^\alpha$ for $x \geq 1$

C. $F(x) = -\alpha(\alpha + 1)/x^{\alpha+2}$ for $x \geq 1$

D. $F(x) = 1 - 1/x^{\alpha+2}$ for $x \geq 1$

5. B

6. (3 points) Kyle, arriving at a bus stop, just misses the bus. Suppose that he decides to walk if the (next) bus takes longer than 8 minutes to arrive. Suppose also that the time in minutes between the arrivals of buses at the bus stop is a continuous random variable with a $U(4, 9)$ distribution. What is the probability that Kyle will end up walking.

A. 1/5

B. 1/2

C. 3/5

D. 4/5

E. 1

6. A

7. (3 points) Suppose you're playing Yahtzee with your friends. In this game, you roll five fair six-sided dice all at once. You want to model the number of rolls it takes until you get the very best roll—all sixes, called a Yahtzee!—via the random variable X. What distribution would you choose for X to follow?

   A. Binomial
   B. Geometric
   C. Poisson
   D. Uniform
   E. Exponential
   F. Negative Binomial
   G. None of the Above

7. B

8. (3 points) You thoroughly shuffle a standard deck of cards and then begin flipping over cards from the top of the deck one at a time. Let X be the random variable representing the number of cards you flip over up to and including the appearance of a first two Jacks. What distribution does X follow?

   A. Binomial
   B. Geometric
   C. Poisson
   D. Uniform
   E. Exponential
   F. Negative Binomial
   G. None of the Above

8. G

9. (3 points) The background radiation rate as measured in a particular room with a particular Geiger counter is known to be 800 counts per hour on average. You run the Geiger counter in the room for a five second interval and write down the count. You run the Geiger counter again for another five second interval and write down that count. You repeat, and write down a third count. And a fourth. And a fifth … . As you store your data in a convenient pandas dataframe, you begin to notice that your measurements appear to be drawn from which of the following distributions?

   A. Binomial
   B. Poisson
   C. Geometric
   D. Uniform
   E. Exponential
   F. None of the Above

9. B

10. (3 points) Consider the following function related to incoming texts to a cellphone during class. What distribution does the return value of the function belong to?

```
def text_me_im_so_stoked_about_3022(q):
    j = 0
    t = np.random.exponential(1/q)
    while t <= 1:
        j += 1
        t += np.random.exponential(1/q)
    return j
```

A. Binomial

B. Geometric

C. Poisson

D. Uniform

E. Exponential

F. None of the Above

10. C

11. (3 points) Consider the following function related to a series of coin flips with a biased coin that lands Heads with probability p. What distribution does the return value of the function belong to?

```
def flippity_flop(p):
    lambda = 20
    j = 0
    k = 0
    while j < lambda:
        k += np.random.choice([0,1], p=[1-p, p]):
    return k
```

A. Bernoulli

B. Binomial

C. Geometric

D. Poisson

E. Exponential

F. None of the Above

11. B

12. (3 points) Consider simulating the roll of a fair, six-sided die. Which of the following quantities does the following function estimate?

```
def rolling_rollers(num_samples):
    rolls = np.random.choice([1,2,3,4,5,6], size=num_samples)
    return np.sum(np.logical_and((rolls % 2)==1),(rolls > 2)) / np.sum((rolls % 2)==1)
```

A. $P(X > 2 \mid X \text{ is Odd})$

B. $P(X \text{ is Odd} \mid X > 2)$

C. $P(X > 2 \cap X \text{ is Odd})$

D. $P(X > 2 \cup X \text{ is Odd})$

12. A

13. (3 points) It's Saturday night and you are at a party on the rooftop of Norlin Library. Everybody is there. The temperature is uniformly distributed between ideal and perfect. The music is loud, if you're into loud music, but also it is more on the quiet side, if you're more into that. Suddenly the music of your preferred volume stops and a voice comes over the speakers:

"Suppose there exists a continuous random variable X with probability density function $f(x)$!"

People scream, but you do not; you have been training for this moment.

"The variance of X is 10 and the expected value of X is 2!"

You smile. You're pretty sure you know where this is going.

"Compute $\int_{-\infty}^{\infty} x^2 f(x)\, dx$ or the party is over! Answer me!"

What do you answer to get the party restarted? (Assume that you are trying to answer correctly because you want the rooftop party to continue.)

A. 6

B. 8

C. 14

D. 104

E. The limit does not exist!

$\mathrm{Var}(X) = E[X^2] - E[X]^2$

$10 = E[X^2] - (2!)^2$

$\Rightarrow E[X^2] = 10 + 4 = 14$

13. C

The rest of this page may be used for calculations, but will not be graded. There is additional space at the very end of this exam.

1 2 3 4 5 6 7 8 9 10 11

14. (20 points) Consider the following data: 8, 8, 10, 11, 11, 11, 11, 12, 13, 16, 17

(a) Compute the quartiles Q_1, Q_2, and Q_3 as well as the IQR for this data set.

$Q_2 = 11$

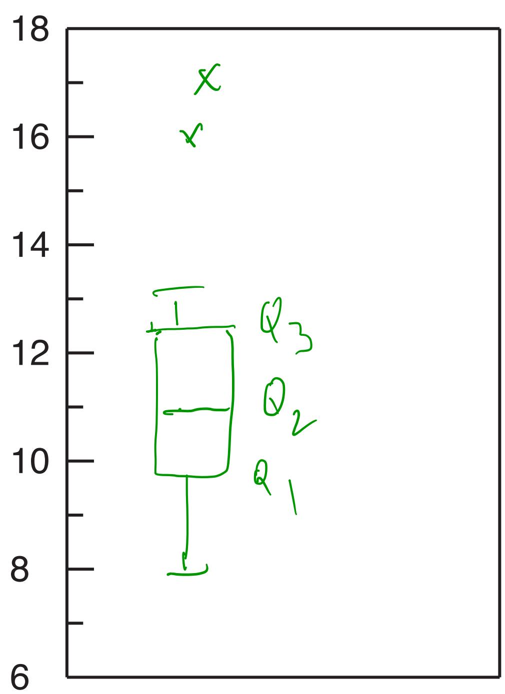
$Q_1 = \frac{10+11}{2} = 10.5$

$Q_3 = \frac{12+13}{2} = 12.5$

$IQR = Q_3 - Q_1 = 12.5 - 10.5 = 2$, $IQR \times 1.5 = 3$

whisker extend $[7.5, 15.5]$

(b) Draw a box-and-whisker plot for the data set using the provided axes. Use the conventions discussed in lecture.



(c) Classify the distribution of the data as symmetric, postively skewed, or negatively skewed. Clearly justify your response.

$\bar{X} = 11$, $\tilde{X} = 11.64$

so it's positively skewed

15. (20 points) You have a box with 10 coins in it. Six of the coins are biased, such that heads comes up twice as often as tails. The remaining coins are unbiased, so heads and tails come up equally as often.

(a) You choose a coin at random. What is the probability that it is a **biased** coin?

$P(H) = 2P(T)$
$P(H) + P(T) = 1$ $\Rightarrow 2P(T) + P(T) = 1$
$P(T) = \frac{1}{3}$, $P(H) = \frac{2}{3}$

$P(B) = \frac{6}{10} = \frac{3}{5}$

(b) You choose a coin at random and flip it. What is the probability that you get a tails?

$$P(T) = P(T|B)P(B) + P(T|U)P(U)$$
$$= \frac{1}{3}\cdot\frac{3}{5} + \frac{1}{2}\cdot\frac{2}{5} = \frac{2}{5}$$

$P(U) = \frac{4}{10} = \frac{2}{5}$

(c) You choose a coin at random. You flip that coin three times. It comes up heads two times and tails the other time. What is the probability that the coin that you chose was one of the **unbiased** coins? You may express your answer as a fraction or as a decimal rounded to 3 decimal places.

Let E be the event of getting a t. and 2 heads

$$P(U|E) = \frac{P(E|U)P(U)}{P(E)} = \frac{P(E|U)P(U)}{P(E|U)P(U) + P(E|B)P(B)}$$
$$= \frac{P(E|U)\frac{2}{5}}{P(E|U)\frac{2}{5} + P(E|B)\frac{3}{5}} = \frac{2P(E|U)}{2P(E|U) + 3P(E|B)}$$

If I flip a biased coin, what is the prob. I get 2 heads 1 tail

$$P(E|B) = \binom{3}{1}\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right) = 3\left(\frac{4}{27}\right)$$

For unbiased coin:

$$P(E|U) = \binom{3}{1}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right) = 3\left(\frac{1}{8}\right)$$

$$P(U|E) = \frac{2\cdot 3\left(\frac{1}{8}\right)}{2\cdot\left(\frac{1}{8}\right) + 3\left(3\left(\frac{4}{27}\right)\right)} = \frac{9}{25}$$

16. (20 points) The probability mass function of a discrete random variable X is given by

$$P(X=-2)=\frac{3}{8}, \quad P(X=-1)=\frac{1}{4}, \quad P(X=0)=\frac{1}{8}, \quad P(X=1)=\frac{1}{4}$$

(a) Compute $E[X]$

$$E[X] = \sum_i a_i P(X=a_i) = -2\cdot\left(\frac{3}{8}\right) + -1\cdot\left(\frac{1}{4}\right) + 0\cdot\frac{1}{8} + 1\cdot\left(\frac{1}{4}\right)$$

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

(b) Let Y be the random variable $Y = X^2 - 1$. Write down the probability distribution of Y.

$(-2)^2-1 = 3$, $(-1)-1) = 0$, $(0)-1) = -1$, $(1)^2-1 = 0$

$P(Y=-1) = \frac{1}{8}$, $P(Y=0) = \frac{1}{4} + \frac{1}{4} = \frac{2}{4} = \frac{1}{2}$

$P(Y=3) = \frac{3}{8}$

(c) Compute $E[Y]$ and $\mathrm{Var}(Y)$

$E[Y] = \sum_i a_i P(Y=a_i) = -1\cdot\frac{1}{8} + 0\cdot\frac{1}{2} + 3\cdot\frac{3}{8} = 1$

$E[Y^2] = (-1)^2\cdot\frac{1}{8} + 0^2\cdot\frac{1}{2} + 3^2\cdot\frac{3}{8} = \frac{1}{8} + \frac{27}{8} = \frac{7}{2}$

$\mathrm{Var}(Y) = E[Y^2] - E[Y]^2 = \frac{7}{2} - 1^2 = \frac{5}{2}$

(d) Suppose you draw 3 times from Y to create a dataset. What is the probability that your dataset has a single non-negative mode?

A single non-negative mode mean that we have drawn one of the follow $\{0,0,0\}$, $\{0,0,0^c\}$, $\{3,3,3\}$, $\{3,3,3^c\}$

Additional Workspace

$$P(\{0,0,0\}) = \left(\frac{1}{2}\right)^2 = \frac{1}{3}$$

$$P(\{0,0,0\}) = \binom{3}{1}\left(\frac{1}{2}\right)^2\left(1-\frac{1}{2}\right) = \frac{3}{8}$$

$$P(\{3,3,3\}) = \left(\frac{3}{8}\right)^2 = \frac{27}{512}$$

$$P(\{3,3,3\}) = \binom{3}{1}\left(\frac{3}{8}\right)^2\left(1-\frac{3}{8}\right) = \frac{135}{512}$$

Additional Workspace

Additional Workspace