PSTAT 5A / **MIDTERM EXAM 2** / Spring 2023   Instructor: **Ethan Marzban**

Name:_____   **UCSB NetID:**_____
　　　　First, then Last   *NOT your Perm Number!*

Circle the section you <u>attend</u>:

Yuan 10 - 10:50am   Jason 11 - 11:50am   Nickolas 12 - 12:50pm   Nickolas 1 - 1:50pm

Your Seat Number:_____

Person Sitting to your Left:_____

Person Sitting to your Right:_____

Instructions:

- You will have **65 minutes** to complete this exam.

- You are allowed the use of a single **8.5 × 11-inch** sheet, front and back, of notes. You are also permitted the use of **calculators**; the use of any and all other electronic devices (laptops, cell phones, etc.) is prohibited.

- For Multiple Choice Questions: fill in the bubble corresponding to your answer directly on the exam. Partial credit will **not** be awarded.

- For Free Response Questions: be sure to include **all** of your work! Correct answers with no supporting work will **not** receive full points.

- ⎡**PLEASE DO NOT DETACH ANY PAGES FROM THIS EXAM.**⎦

- Good Luck!!!

Honor Code: In signing my name below, I certify that all work appearing on this exam is entirely my own and not copied from any external source. I further certify that I have not received any unauthorized aid while taking this exam.

　　　✕_____

Problem 1. A 97% confidence interval for some population parameter θ has been constructed using a particular dataset. Would we expect the width of a 72% confidence interval for θ, constructed using the same dataset, to be wider than or narrower than the width of the 97% confidence interval?

[1pts.]

     ○ Wider

     ○ Narrower

     ○ The same width

Problem 2. 500 fair six-sided dice are rolled independently of each other, and the outcomes are recorded. What is the expected number of dice (out of these 500) that land on the number 1?

[1pts.]

     ○ 0.1667

     ○ 1.0000

     ○ 250.0000

     ○ 83.3333

     ○ None of the above.

Problem 3. Suppose 34% of all residents in a particular country live in a two-story home. In a representative sample of 300 residents, 30% are found to live in a two-story home. **Select the answer choice** that best fills in the blanks in the following sentence:

[1pts.]

    | The value of 34% is a ___*(Blank 1)*___ , and the value of 30% is a ___*(Blank 2)*___ .

     ○ **Blank 1:** population parameter;   **Blank 2:** population parameter

     ○ **Blank 1:** population parameter;   **Blank 2:** sample statistic

     ○ **Blank 1:** sample statistic;   **Blank 2:** population parameter

     ○ **Blank 1:** sample statistic;   **Blank 2:** sample statistic

Problem 4. Does the expected value of a random variable have to be an element of the random variable's state space?
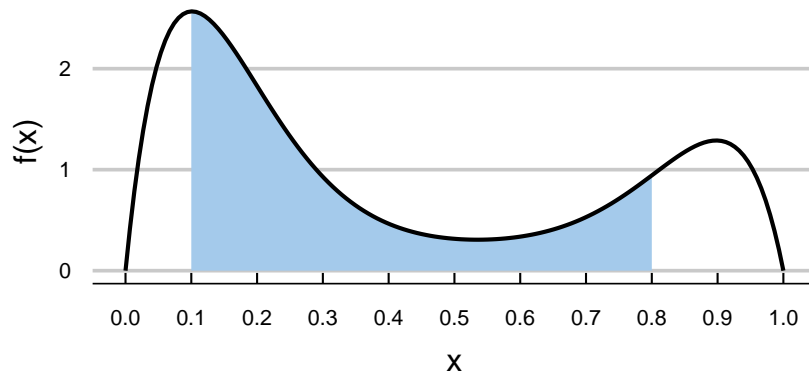
[1pts.]

     ○ Yes

     ○ No

Problem 5. What module contains the function `t.ppf()`?

[1pts.]

     ○ `numpy`

     ○ `datascience`

     ○ `scipy.stats`

     ○ `matplotlib.pyplot`

     ○ None of the above

Problem 6. Consider a random variable X with the following density curve: [1pts.]

What is the probability that is shaded in the figure above? (If the figure is hard to read: the region shaded is the region underneath the density curve between $x = 0.1$ and $x = 0.8$.)

○ $\mathbb{P}(X \leq 0.1) - \mathbb{P}(X \leq 0.8)$
○ $\mathbb{P}(X \leq 0.8) - \mathbb{P}(X \leq 0.1)$
○ $\mathbb{P}(X \geq 0.1) - \mathbb{P}(X \leq 0.8)$
○ $\mathbb{P}(X \leq 0.1) - \mathbb{P}(X \geq 0.8)$
○ None of the above.

Problem 7. The time it takes Johnny to bike from his apartment to UCSB varies uniformly between 18.3 minutes and 23.5 minutes. What is the probability that Johnny will take exactly 20 minutes to bike from his apartment to UCSB today? [1pts.]

○ 0.0000
○ 0.1923
○ 0.3269
○ 0.6730
○ None of the above.

Problem 8. What is the correct syntax to import the `numpy` module with the nickname `np`? [1pts.]

○ `import numpy as np`
○ `import numpy with name np`
○ `import.numpy.np`
○ `import numpy with nickname np`

Problem 9. Karla would like to compute the standard deviation of the list of numbers [1, 4, 3, 6, 10]. To do so, she has imported the `numpy` module (not with any nickname), and is going to run `std([1, 4, 3, 6, 10])`. Will this give her the desired result? As a reminder, the formula we use in this class to compute standard deviation is

$$\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

○ Yes, it will give her the desired result.

○ No, because she needs to pass in the argument `ddof = 1` to her call to `std()`.

○ No, because she has copied down the wrong list of numbers.

○ None of the above.

Problem 10. Consider the following code cell:

```
y = 0

for k in np.arange(0, 10):
    y = k - 1
```

What will be the value of `y` after running the above code cell once?

○ 5

○ 6

○ 7

○ 8

○ 9

○ 10

○ None of the above

Free Response Questions

Problem 11. The duration of a randomly selected flight out of SBA (Santa Barbara Airport) is uniformly distributed between 50 minutes and 200 minutes.

   (a) Define the random variable of interest, and call it X.   [1pts.]

   (b) Sketch the density curve of X. Make sure you clearly label any points of importance on your graph.   [2pts.]

   (c) What is the probability that a randomly selected flight will have duration between 70 and 140 minutes?   [2pts.]

(d) A sample of 10 flights is taken (with replacement). What is the probability that this sample contains exactly 4 flights with duration between 70 and 140 minutes? [4pts.]

Problem 12. An electrician collects a representative sample of 45 *GauchoBuzz*-brand batteries and records their average lifetime (in years). The average lifetime of her sample of 45 batteries is 1.92 years, and the standard deviation of lifetimes in her sample is 0.5 years.

(a) Define the random variable of interest, $\overline{X}$. [1pts.]

(b) What distribution would we use to construct a 95% confidence interval   [3pts.]
   for the true average lifetime of a *GauchoBuzz*-brand battery? Be sure to
   include any/all relevant parameter(s), and check any/all conditions!

(c) Construct a 95% confidence interval for the true average lifetime of a *Gau-*   [3pts.]
   choBuzz-brand battery, and interpret your interval.

Problem 13. Cillian, the owner of *Cillian's Ice Cream*, believes that 77% of his customers pay with credit card. To test his suspicion, he takes a representative sample of 170 customers and finds that 74% of these customers pay with credit card. With this data, he wishes to conduct a two-sided hypothesis test at an $\alpha = 0.05$ level of significance.

(a) Write down the null and alternative hypotheses for this test. Be sure to use mathematical notation, and define any variables you use. [2pts.]

(b) Compute the value of the test statistic. [3pts.]

(c) What is the distribution of the test statistic, assuming the null hypothesis   [2pts.]
is correct? Be sure to check any relevant conditions, and be sure to specify
any/all necessary parameter(s) in your final answer!

(d) Find the critical value of the test.   [2pts.]

(e) Now, conduct the hypothesis test and state your conclusions in the context   [2pts.]
of the problem.

Problem 14. A random variable X has the following probability mass function (p.m.f):

k	-1	0.5	2
$\mathbb{P}(X = k)$	0.1	a	0.5

(a) What is the state space S_X of X? [1pts.]

(b) Is X a discrete random variable or a continuous random variable? Justify your answer. [1pts.]

(c) Find the value of a. [2pts.]

(d) Compute $\mathbb{P}(X \leq 1)$. [2pts.]

(e) Compute $\mathbb{E}[X]$, the expected value of X. (Make sure you plug in the value for a you found in part (d) above.) [3pts.]

(f) Compute $\text{Var}(X)$, the variance of X. (Make sure you plug in the value for a you found in part (d) above.) [4pts.]