Math 5710 Opportunity #4 ♬ Deadline: 1 p.m. on Thursday, August 6, 2020

1. Please print your name legibly. Brock Francom A02052161

2. Examine each of the following propositions, determine whether or not it is true, display your choice by circling either "T" or "F"; for each either prove that your decisions is correct or write at least one paragraph that explains why you decided that the proposition is true or why you decided that the proposition is false:

A. $X \in \{$ discrete random variables of $\Omega \} \Rightarrow X \in \{$ sets $\}$

T F Entry #038 in our glossary contains the definition of a discrete random variable. The last bit of that definition states $X: E \to \mathbb{R}$. There fore X is a function mapping an event of Ω to $\mathbb{R}$. Therefore X is a set.

B. (A card is randomly drawn from standard 52-card poker deck) ⇒ (the odds that a club is drawn is 0.25)

T F This is false because the probability that a club is drawn is .25, But the odds of drawing a club is not .25. Using definition 40 from our glossary will show us that the odds of drawing a club is $\frac{.25}{.75} = \boxed{\frac{1}{3}}$

C. $\{$ Bernoulli random variables $\} \cap \{$ binomial random variables $\} = \varnothing$

T F Bernoulli random variables are Binomial random variables. There fore, {Bernoulli random variables} ∩ {binomial random variables}

⇒ = {Bernoulli random variables}

non inclusive

D. A number x is randomly selected from $[\,^-1,\ 1\,) \Rightarrow$ the probability that $x = {}^-1 >$ the probability that $x = 1$

0

T F Because the set of numbers contained in $[-1, 1)$ is uncountably infinite, the probability that $x =$ any given number in that range is 0. Therefore, the probability that $x = -1$ is 0, as is the probability that $x = 1$ because $X = 1$ is not even in the range $[-1, 1)$.

E. A number x is randomly selected from $(\,0, 6\,) \Rightarrow$ the probability that $x \geq \pi$ is 0.5

T F The probability that $x < \pi$ is approxamatly $\frac{\pi}{6} \approx .5236$. Therefore, the probability that $x \geq \pi$ is $1 - .5236 = .4764$. Of course, this is assuming that your sample space is all numbers in the interval $(0, 6)$ and each number is equally likley to be chosen.

F. $X \in \{$ continuous random variables of $\Omega\,\} \Rightarrow |\,X\,| = \mathfrak{C}$

T F

If X is a continuous random variable of Ω, it is a function mapping events in Ω to a number in $\mathbb{R}$. Because X is a continuous random variable, there are uncountably infinite number of events. Therefore $|X| = \mathfrak{C}$.

G. $X \in \{ \text{random variables of } \Omega \} \Rightarrow X \in \{ \text{events of } \Omega \}$

T F

When we take a look at our definitions for random variables in our glossary, we will notice that X, a random variable, is actually a function that maps an event to a number in $\mathbb{R}$. That means X is not an element of {events of Ω}, but a function mapping events of Ω to a number in $\mathbb{R}$.

H. Results from an interval measurement can be tenably interpreted as if they were nominal.

T F

Interval measurment has categories just like nominal measurments. For instance, say there are 10 guys whose height is between 64-67 inches. That would be an interval measurment, But you could interperet it as nominal and call the category 64-67 inches.

J. Measurement usefulness is a necessary condition for measurement reliability.

T F

Measurment usefullness is a suffiecent condition for measurment reliability. Necessesary conditions for measurment reliability include: Internal Consistency, and both intra and inter observer consistency.

3. Please solve the following problem; display the computations – including the probability distribution:

$n=4$

$P(\text{Success})=.2$

A person is randomly selected from a population and tested for COVID-19 infection. A positive test result is labeled a "success" and coded as 1; a negative test result is labeled a "failure" and coded as 0. Again a person is randomly selected from that *same* population (Thus, the first person is still in the population; so the two events are independent). The trial is repeated twice more. The number of successes is recorded. As of June 2, 2020, one seemingly credible estimate is 20% of the people worldwide are infected; use that figure for this problem. Display the probability distribution for the random variable for this experiment.

This is going to be a binomial distribution, and so we will use therom 11 from our glossary to help us come up with our probability distribution.

$P(k) = \binom{n}{k}(m(1)^k)(1-m(1))^{n-k}$

$P(0)=\binom{4}{0}(.2)^0(.8)^4 = .4096$

$P(1)=\binom{4}{1}(.2)^1(.8)^3 = .4096$

$P(2)=\binom{4}{2}(.2)^2(.8)^2 = .1536$

$P(3)=\binom{4}{3}(.2)^3(.8)^1 = .0256$

$P(4)=\binom{4}{4}(.2)^4(.8)^0 = .0016$

4. Comprehend the following case:

An educational H&PE researcher conducted a study to assess the relationship between the psychomotor agility of second-grade children and their computational fluency. She administered a psychomotor agility test as well as a computational fluency test to a single random sample of 86 second-grade students. The resulting string of bivariate data X is of the following form:

$$X = ((v_1, d_1), (v_2, d_2), (v_3, d_3), \ldots, (v_{86}, d_{86}))$$

The resulting sample statistics are as follows:

$$n = 86 \wedge r = .10$$

She tested the following null hypothesis via a t-test for correlations:

$$H_o : \rho_X = 0$$

The calculation from http://vassarstats.net/textbook/ch4apx.html provided the following results:

N = 86	r = 0 1
Reset	Calculate
t	**df**
0 921	84
Probability	
directional	0 1798095
non-directional	0 359619

Because the t value was such that $p > 0.05$, the researcher failed to reject H_o.

Examine each of the following propositions to determine its truth value; indicate your choice by circling either "T" or "F" and then write a paragraph defending your choice:

i. The results of the t-test indicated that the correlation coefficient is not statistically significant.

T F

Because the p-value was such that $p > 0.05$, the Correlation coefficient is not statistically significant. Smaller p-values mean more statistical significance, and our p-value was large.

ii. The results of the t-test indicated that there is no relationship between second grade students' psychomotor agility and their computational fluency at least among those children represented by the study sample.

T F

Because she failed to reject H_0, we did not find evidence that there is a relationship between the psychomotor agility and computational fluency. We should conclude that the test was inconclusive.

iii. The results of the t-test indicated that $|r|$ is so close to 0, that H_o should be accepted.

T F

The results of the t-test indicated that we should not reject H_0. However, this does not prove that H_0 is correct and that it should be accepted.

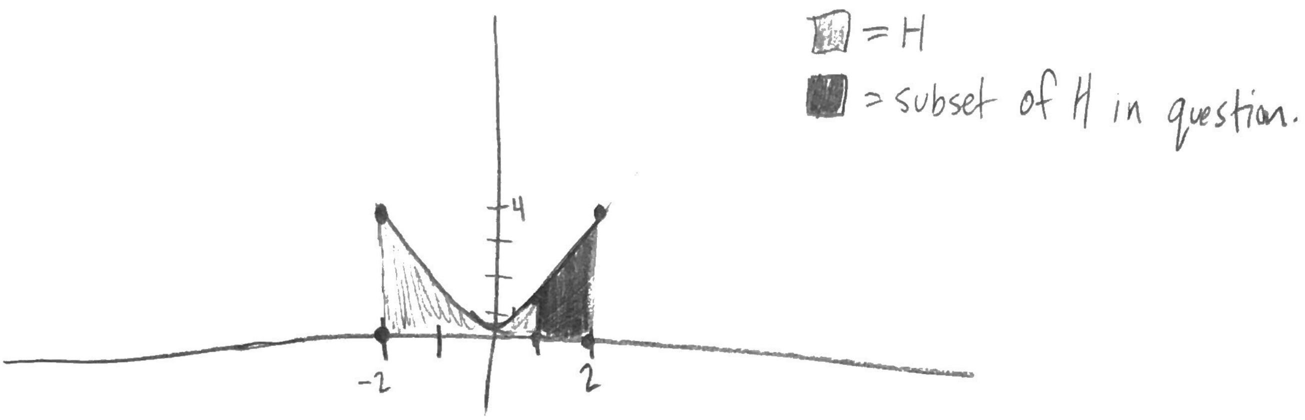
iv. Based on the results of the study, a Type II error is possible but it is impossible to have a Type I error.

T F

Type II error occurs when we fail to reject a false null hypothisis. Because we failed to reject our H_0, Type II error could have ocured.

5. Consider the graph of $f \ni f: [{}^{-}2, 2] \to \mathbb{R} \ni f(x) = x^2$. H is a region that is bounded under the curve of f and is determined by the following 4 points with coordinates $({}^{-}2, 0)$, $({}^{-}2, 4)$, $(2, 0)$, and $(2, 4)$. An experiment is conducted in which a point is randomly selected from H. What is the probability that the selected point lies in the subset of H that is determined by 4 points with coordinates $(1, 0)$, $(1, 1)$, $(2, 0)$, and $(2, 4)$?

Display an illustration that is useful in helping people comprehend the question. And display the computations that will help people understand how you computed the problem in question.



E = the point picked lies in the subset of H.

$$P(E) = \frac{\int_1^2 x^2\,dx}{\int_{-2}^2 x^2\,dx}$$

What we want to find is the $\frac{\text{Area of Subset}}{\text{Area of H}}$. That is easily done using integration.

$$P(E) = \frac{\left.\frac{x^3}{3}\right|_1^2}{\left.\frac{x^3}{3}\right|_{-2}^2} = \frac{\frac{8}{3} - \frac{1}{3}}{\frac{8}{3} - \frac{-8}{3}} = \frac{\frac{7}{3}}{\frac{16}{3}} = \boxed{\frac{7}{16} \approx .4375}$$

6. Note that our *Glossary* includes a definition for *binomial random variable*:

043A. Definition for a binomial random variable:

Given $n \in \mathbb{N} \wedge \Omega = \{ 1, 2, 3, ..., n \} \wedge$
$X \in \{$ Bernoulli random variables of $\{ 0, 1\}\} \wedge$
(A string of n experiments are conducted with $X \ni (X(i) = 0 \veebar (X(i) = 1$ depending on the results of the i^{th} experiment $\wedge |\{ (i, X(i)) : X(i) = 1 \}| = k$), ($Y \in \{$ binomial random variables of $\Omega \} \Leftrightarrow$

$Y \in \{$ discrete random variables of $\Omega \} \ni Y(i) = \sum_{i}^{n} X(i) = k$

However our *Glossary* does not include a rigorous formal set-theoretic definition for *hypergeometric random variable.* Instead we only have these notes:

044D. Note on *hypergeometric* experiments and related discrete probability distributions:

A hypergeometric experiment is quite similar to a binomial experiment but with one crucial exception: In a binomial experiment, the selected events are independent from one another; whereas in a hypergeometric experiment the selected events are dependent on one another because events are selected one at a time without replacement. Consider the following example of a hypergeometric experiment:

> Five cards are randomly selected from a standard 52-card poker deck and this is done *without replacement.* The goal of the experiment is to determine the probability that exactly two of the selected cards are red.

Now consider a similar experiment that is binomial rather than hypergeometric:

> Five cards are randomly selected one at a time from a standard 52-card poker deck; after the first card is selected, its color is recorded and returned to the deck, and the deck is reshuffled. The same algorithm is repeated until five cards have been drawn. Thus, this is done *with replacement.* The goal of the experiment is to determine the probability that exactly two of the selected cards are red.

For a hypergeometric random variable X, the formula for the discrete probability function p can be expressed as follows where N is the number of elements in the population (e.g., 52), k = the number of successful events in the population (e.g., the number of possible events in which exactly two cards are red), n = the number of element in each event (e.g., 5) , and x is the number of successes in the random sample (e.g., 2):

$$p(X = x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

And now we have a confession: We believe that Math 5710 should build our talents for reading and writing mathematics w/r the theory of probability. And now it's time to help us assess our progress toward that lofty goal. So here is the task:

> Attempt to develop a set-theoretic rigorous definition for *hypergeometric random variables*. Maybe we should consider beginning by taking our definition for *binomial random variables* and modify it so that it mutates into something much different (i.e., a definition of *hypergeometric random variable*).

Write a two-paragraph report on your efforts with this daunting task that may or may not culminate in a satisfactory definition.

Given $n \in \mathbb{N} \wedge \Omega = \{1,2,3,\dots,n\} \wedge X \in \{$Bernoulli random variables of $\{0,1\}\}$ $\wedge$ (A string of n experiments are conducted with $X \ni (X(i)=0) \veebar (X(i)=1)$ depending on the results of the i^{th} experiment $\wedge\ |\{(i,X(i)) : X(i)=1\}| = k)$,

$(Y \in \{$hypergeometric random variables of $\Omega\}) \Leftrightarrow$

$Y \in \{$discrete random variables of $\Omega\} \ni Y(i) = \sum_{1}^{n} X(i) = k \wedge$ (After each experiment i, element i is removed from the sample space/population).

This task was difficult, and I am not completly satisfied with my definition. I found that although I understood the definitions we have used in this class, I am not fluent enough to be able to write my own. I feel like I understand very clearly the difference between a binomial random variable and a hypergeometric random variable, but my understanding is not translating to set theory.

7. Smile; you've taken advantage of Opportunity #4.

I Know how we arrive at the hypergeometric pmf as well. I am able to compare each peice of that function to a binomial pmf and show why each peice is necessary. I ended up writing as best as I could in words what I wanted to express in set notation. Cheers, Brock.