Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

ECON/ACCT/BUSA 222—Bethany College

**Exam 02**

* There are 110 possible points on this exam. The test is out of 100.
* You have two hours to complete this exam, but you should be able to complete it in less than that
* Please turn off all cell phones and other electronic equipment.
* You are allowed a calculator for the exam. This calculator cannot be capable of storing equations. This calculator cannot double as a cell phone.
* Be sure to read all instructions and questions carefully.
* Remember to show all your work.
* Recall basic logic. “Water is wet” is a true statement. “Water is wet and leopards have stripes” is a false statement.
* You are allowed one 3” by 5” note card with the exam. You are allowed any information you deem important on it.
* *Please print clearly and neatly.*

**Part I: Multiple Choice.** *Choose the best answer to the following.*

4 points each.

1. Which of the following is an advantage of sampling over just surveying the population?
	1. It’s more accurate
	2. It’s often more practical
	3. It’s cheaper
	4. **B & C**
	5. None of the above

*Surveying the population is, by definition, more accurate than a sample. But sampling is both cheaper and more practical (you can’t survey the whole population of cars for crash testing because then you would have no cars to drive).*

1. Francis is curious how Wal-Mart patrons feel about their shopping experience. He surveys every tenth person upon leaving the store. What kind of sample is this?
	1. **Systematic**
	2. Stratified
	3. Simple Random
	4. A & C
	5. None of the above

*This one’s a bit tricky. Recall that systematic sampling is sampling every element at a fixed increment, which is what Francis is clearly doing. But it’s not a simple random sample. While the starting point is random, some combinations of patrons may not be selected. For example, in a couple leaving together it is impossible they will both be selected. A simple random sample (rolling a 10-sided die for each one and selecting them for sampling if a “10” is rolled) would allow a 1% they are both selected. Under systematic, that chance is zero percent.*

1. Which of the following is an example of a discrete probability distribution?
	1. The number of viewers of *American Idol* in a given week.
	2. The number of items in a refrigerator.
	3. The number of credit cards in a wallet.
	4. **B & C**
	5. None of the above

*While (A) may seem like the right answer—you can’t have half a viewer—the average of (A) is so large, it is functionally continuous. However (B) and (C) are both discrete and the average is not small enough to be functionally continuous.*

1. Each factory which makes Lays Potato Chips makes thousands of chips every day. Suppose you sample 100 chips to ensure the quality control machines are running correctly (like all things, some mistakes are expected even when they is running well). What technique should you use to determine if the mistakes are numerous enough to determine if the machine is broken?
2. Binomial distribution
3. Poisson distribution
4. **z-test**
5. t-test
6. None of the above

*Both (A) and (B) are used to determine probabilities that certain conditions are met. To compare to a population/expectation, you need either a z-test or a t-test. Since you know the population standard deviation, you use a z-test.*

1. Which of the following is an example of Type 1 Error?
2. Not seeing a new television show which you would have enjoyed
3. Turning down a job offer from Facebook before the company took off on the basis that Facebook will never “make it big.”
4. Eating the same thing you always eat even though sometime better is available.
5. A & B
6. **None of the above**

*In all of these, you accepted the null hypothesis when you should have rejected it. A Type I Error is the opposite.*

1. Consider a uniform distribution with a maximum value of 10 and a minimum value of 4. What is the standard deviation of this distribution?
	1. √0.5
	2. **√3**
	3. √12
	4. √18
	5. None of the above

*This is just some math: the square root of (10 – 4)2/12* = √3.

1. Donald owns a real estate company and wants to reward his best performing salesmen (“best” based on highest sales). He can afford to give a $5,000 bonus to the top 10% of salesmen. Sales follow a normal distribution with an average of $20,000 per week and a standard deviation of $3,000. To determine the cutoff point, Donald pulls up Excel for a normal inverted distribution. The function arguments look like this:



What, if anything, is wrong with what Donald is doing here?

* 1. The function should be NORMSINV (for standard normal), not NORMINV.
	2. **The probability should be 0.9, not 0.1.**
	3. The mean should be $5,000, not $20,000.
	4. A & C
	5. None of the above

*Donald is measuring from the bottom 10%, not the top 10%. Note that the calculation for this ($16,155.35) is below the mean. He wants the cutoff for the bottom 90%, or 0.9.*

1. I once told a friend that hybrid cars aren’t as eco-friendly as one might think. Because they save so much gas, people with hybrids end up driving more. My friend responded: “That’s absurd! I have a hybrid car and I don’t drive any more than I otherwise would.” What mistake was my friend making?
2. **His sample is biased**
3. He is accepting the null hypothesis when he should have rejected it
4. He is assuming his behavior hasn’t change but he may be wrong
5. A & B
6. None of the above

*This individual has a sample size of one, and that sample is not accurate, since he chose it based solely on the fact that it’s him. The problem is not merely that the sample is one (a small sample), but that it wasn’t random.*

1. Under all circumstances, a normal distribution:
	1. Has a mean of zero
	2. Has most of its observations within one standard deviation of the mean
	3. Is bell-shaped
	4. **B & C**
	5. None of the above

*About 68% of observations are within one standard deviation of the mean and of course a normal distribution is bell shaped. Option (A) only applies to a standard normal distribution.*

1. Lars Costanzia runs an oil change shop. It takes, on average, 14 minutes for one of his employees to change someone’s oil. Lars just hired Kelly. After sampling 9 oil changes, Lars calculated Kelly to have an average change time of 12 minutes. Suppose Lars wants to know if his new employee is faster than his other ones. If X is Kelly’s true average (projecting her average into future performance), what is the null hypothesis?
	1. X = 12
	2. **X = 14**
	3. X > 12
	4. X < 14
	5. None of the above

*The null hypothesis is that nothing interesting is going on; in this case, that means Kelly is no better nor worse than her co-workers. Option D is our alternative hypothesis—“if his new employee is faster”—which, notably, is one-tailed since we are only asking if she’s faster and not also possibly slower.*

1. Suppose the standard deviation of the population is 1 minute and the standard deviation of the sample is 2 minutes. Using the information from the previous question, what is the calculated value?
2. t = -3
3. t = -6
4. z = -3
5. **z = -6**
6. None of the above

*The main point here is knowing if you should do a z- or a t-test. Recall that you use the t-test if you don’t know the population’s standard deviation. But we do know it—1 minute—so we use a z test (thus A nor B can be right). So we subtract (12 – 14 = -2) and then divide the population’s standard deviation by the square root of 9 (1/ √9 = ⅓). Then we combine them as per the equation: -2 / ⅓ = -6.*

$$z= \frac{\overbar{x}-μ}{^{σ}/\_{\sqrt{n}}}= \frac{12-14}{^{1}/\_{\sqrt{9}}}= \frac{-2}{^{1}/\_{3}}=-6$$

1. Kelly selects 11 books at random from a shelf of 30 books. She doesn’t put them back. Twenty of these books are fiction. If you are curious what the likelihood is that Kelly will select nothing but books of fiction, what technique would you use?
2. Binomial distribution
3. Poisson distribution
4. z-test
5. t-test
6. **None of the above**

*What you actually want to hypergeometric, which is binomial but without replacement. Since your sample is more than 5% of your population, the probability of success isn’t constant between each book taken.*

**Part II: True/False.** *Answer true or false, and justify your answer.*

10 points each.

1. Suppose Carol randomly surveys people with a job about how much each person makes. If she uses that data to estimate the average income for *all* Americans, then her sample is accurate but not precise.

*False. While we don’t know how precise it will be (that’s largely a function of sample size), we know it will not be accurate. Carol’s sample will be biased since it will not include Americans that are not employed.*

1. By the Central Limit Theorem (CLT), the mean of many sample means should equal the population mean.

*True. We know that the CLT states many sample means will form a normal distribution. The average of this normal distribution is the population average.*

1. When you write your paper, an appropriate thesis statement could be “I explore how poverty rates affect CO2 emissions.”

*False. While I won’t go into exact paper details on the actual exam, you should know (without even having to take this course) that the above statement is not a thesis statement. A thesis statement is some kind of argument. “Exploring…” is not a thesis statement. If, however, you wrote “Greater poverty rates decrease pollution,” that is a thesis statement (an argument you would defend in the paper).*

**Part III: Short Answer.** *Answer the following.*

16 points each.

1. Suppose cargo captain Kasidy Yates wants to invest in a new starship. To determine what she can afford, she needs to figure out what she can expect to transport (in cubic meters). Suppose she estimates the cubic meters of goods she can expect to transport using the latest orders from her busiest run: from Earth to Bajor. Based on previous experience, she knows that the standard deviation in cubic meters is 3,000. (A war just broke out and it will decrease trade—so she can’t use her previous data to determine the new mean—but it’s reasonable to assume the standard deviation is the same.) With 95% confidence, she wants to limit her error to no more than 600 cubic meters on either side of her mean. How many orders will she have to sample? (HINT: at 95% confidence, z=1.96.)

*Recall we reworked the confidence interval equation so as to isolate n:*

$$n=\left(\frac{zσ\_{p}}{E}\right)^{2}$$

*So we plug and chug:*

*n = (1.96\*3,000/600)2 = (1.96\*5)2 = 96.04, or 97 samples.*

1. Consider Rachel, who works for an auto insurance company. She wants to know if men are worse drivers than women. In a survey of 100 men, 12 have caused a major accident in the past five years. In a survey of 110 women, 6 have caused a major accident in the past five years. At 95% confidence (z=1.96) do men have a higher accident rate than women?

*Recall that to see if two proportions are different, we use the following equation:*

$$z=\frac{p\_{1}-p\_{2}}{\sqrt{\frac{p\_{c}(1-p\_{c})}{n\_{1}}+\frac{p\_{c}(1-p\_{c})}{n\_{2}}}}$$

*Where,*

$$p\_{c}=\frac{x\_{1}+x\_{2}}{n\_{1}+n\_{2}}$$

*p1 = 12/100 = 0.120*

*p2 = 6/110 = 0.055*

*pc = 18/210 = 0.086*

*(0.12 – 0.055) / √[(0.086)(1 – 0.086)(1/100 + 1/110)]*

*0.065 / √[0.079(0.01 + 0.009)]*

*0.065 / √0.0015*

*0.065 / 0.039*

*1.67 = z*

*This is less than the critical value of 1.96 so we fail to reject the null hypothesis that there is no difference between groups.*