Name: **Key**

ECON/ACCT/BUSA 222—Bethany College

**Exam 01**

* 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. What was the Stanford Prison Experiment?
	1. When the Stanford University administration experimented with supplementing their income by converting some dorms into prison cells.
	2. A famous experiment which demonstrated the importance of standard deviation.
	3. **An experiment which demonstrated the importance of strict ethics rules.**
	4. B & C
	5. None of the above

*Recall this is the experiment where half the subjects likely did severe emotional and psychological harm to the other half.*

1. Suppose you had some data concerning daily oil production (in gallons) for 500 different wells in the United States. If you wanted to get an idea for the distribution of production (including if there are multiple modes), which data display would be most appropriate?
	1. Dot plot
	2. **Histogram**
	3. Stem-and-leaf
	4. Box plot
	5. None of the above

*There’s going to be a wide variety of production, including fractions since we’re dealing with the daily production of a liquid, so a dot plot wouldn’t be a good idea. We have too many observations for a stem-and-leaf plot and a box plot won’t tell us if the distribution is bimodal. Histogram is the best option.*

1. What is a major disadvantage to using the median to describe data?
	1. It’s disproportionately influenced by large outliers
	2. If an increase in the sample shares a similar value, it may not change
	3. It has difficulty describing continuous variables
	4. B & C
	5. **None of the above**

*The major problem is it can mask some important changes, e.g. the lower half of the population lose half their value. Note the median would change if B happens.*

1. If the growth rate in one year is 6% and the growth rate in the following year is 2%, about what is the geometric mean?
2. 1.04%
3. 3.46%
4. **3.98%**
5. 4.00%
6. None of the above

*Convert 6% and 2% into decimal form (0.06 and 0.02, respectively), add one (1.06 and 1.02, respectively), multiply them together and take the square root (1.0398), subtract one (0.0398) and then convert it back to percent (3.98%).*

1. Which of the following is a normative statement?
	1. “I suspect the turn we need to make should be coming up soon.”
	2. “The more you tax people, the less incentive they have to work.”
	3. **“The best policy for America right now is to remove immigration barriers.”**
	4. A & C
	5. None of the above

*Just because the sentence has the word “should” in it, doesn’t mean it’s normative. Similarly, it doesn’t need the word “should” to be normative. It’s about context. Option B is a positive statement and A is hypothesis.*

1. Astrophysicist Neil deGrasse Tyson argues in favor of increasing NASA’s budget because it does valuable things and its budget only consumes half a penny for every taxpayer dollar. Assume he is correct about the size of the budget but his argument is flawed. How would one describe the error or Dr. Tyson’s argument?
2. It is sound but not valid
3. It is valid but not sound
4. **It is neither valid nor sound**
5. It is both valid and sound
6. None of the above

*A valid argument follows logically from premise to conclusion. A sound argument is valid with true premises. We assume the premise—the size of the budget—is true (and I believe it is). However, his argument is not valid: just because a budget for something valuable is small does not mean it is worth increasing (see opportunity cost and diminishing marginal returns). And because it is not valid, it cannot be sound (a sound argument is both valid and has true premises). Thus the answer is C.*

1. Which of the following is an example of a pair of variables that are likely positively correlated?
2. Insurance premiums and accidents
3. Income and spending
4. Weight and level of exercise
5. **A & B**
6. None of the above

*While you’d expect people with lots of accidents would have higher premiums and people with greater income to spend more, you wouldn’t expect people who exercise a lot to weigh a lot (most people, especially in the Western world, weigh a lot because of fat, not muscle).*

1. Which of the following have a normal distribution with zero skewness?
	1. The result of die roll
	2. Income
	3. **Height**
	4. A & C
	5. None of the above

*A die roll result has a uniform distribution, income is skewed to the right, but only height resembles a normal distribution with no skewness.*

1. What is the difference between a set of panel data and a time series?
	1. Panel data tends to be larger
	2. The data of a time series is continuous
	3. Panel data aren’t made up of cross-sectionals
	4. A & C
	5. **None of the above**

*None of these are true; panel data is made up of variables with some categorical difference between them but in the same time period (e.g. 2005 earnings across different industrial sectors) while a time series is made up data in the same category across different time periods (e.g. financial sector earnings from 1900 to 2010).*

1. The percent of families who own their own home is positively correlated with average income across different U.S. states. Suppose a governor attempts to increase average income by subsidizing homeownership. If this is a mistake, what is the most likely reason?
	1. **Reverse causation: home ownership is the result of high incomes.**
	2. Reverse causation: something else is causing both variables.
	3. Confounding variable: home ownership is the result of high incomes.
	4. Confounding variable: something else is causing both variables.
	5. None of the above / The governor has the correct interpretation

*Options (B) and (C) are interpreting reverse causation and confounding variables incorrectly. While (D) is possible, it is unlikely…what would that underlying variable possibly be? Interesting, E isn’t a terrible choice—perhaps homeownership allows people to borrow more because the house is collateral and that might mean they go to college which means, later, they get more income. Maybe, but there’s a lot of ifs, maybes, and conditions in that statement. Option (A) is the most likely answer.*

1. Which of the following is a weakness of presenting data in a pie chart?
	1. It is intuitively difficult to tell what the chart represents
	2. People have trouble interpreting round objects
	3. **It’s hard to tell which section is largest**
	4. A & C
	5. None of the above

*Pie charts are common because people intuitively know what they represent—each wedge is a percent of total something—but it’s difficult to tell which section is largest. Option (B) is kind of correct—what matters for charts is the share of each section’s circumference but the chart gives you an area—but because of how it is worded, it isn’t correct.*

1. Suppose time spent playing video games and non-violent criminal activity are negatively correlated. While there might be a causation story (video games offer a safe outlet for criminal urges), a confounding variable could also be an explanation. Which of the following is the most likely reasonable confounding variable?
	1. Frequency of police patrols
	2. Frequency of new video game releases
	3. **Weather**
	4. A & B
	5. None of the above

*While you could argue all these of these are confounding variables, only option C is really strong. Both A and B suggest criminals might stop thieving in order to play games—thus the negative correlation—but that relies on the same people switching tasks.*

*Weather, however, doesn’t require that assumption. If the weather is good, fewer people will play video games and more criminals will engage in illegal activity but these don’t have to be the same people. The people could be shopping, surfing, going to bars, etc. And because the criminals are outside, they are not sleeping or watching TV and, yes, playing video games. Because C doesn’t require that the gamers are the same folk as the criminals, C is the best option.*

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

10 points each.

1. An example of inductive reasoning would be “According to Jerry, he was at Mallory’s Bakery at 10:55 AM and then walked to Tuba St, arriving five minutes later. Since the only way to get to Tuba St. from Mallory’s that quickly is to cross the Jenkins Bridge, he must have crossed that bridge. Or he’s lying. ”

*False. This is deductive reasoning, not inductive. Assuming Jerry is telling the truth, this is what must be happening. (I added that “Or he’s lying” sentence to highlight the idea that you are making a logical argument based on premises. Note you can tell this is deduction because you can verify the premise by testing the conclusion: suppose Jenkins Bridge has a surveillance camera and you don’t see Jerry crossing it during the time in question. Given this evidence, you would hypothesize that Jerry was lying (your premise was false). In other words, you’re testing the premise indirectly (“if what you’re saying is true, this should also be true”) rather than directly (such as hooking up Jerry to a lie detector).*

1. It takes five and only five numbers to construct a box-and-whisker plot.

*True. You only need the maximum, minimum, median, first quartile, and third quartile to construct such a plot.*

1. If the kurtosis value of a data set is negative, the distribution skews to the right.

*False. Kurtosis is the measure of “peakness” of the data (negative is a sharper peak and positive is a flatter peak). If the distribution is skewed to the right, then skewness value would be positive.*

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

16 points each.

1. Sam grows potatoes and he’s curious if his crop is infected with a disease. He knows nationally, 1% of crops have the disease. His test is 90% sensitive and 98% specific. If the crop tests positive, what is the likelihood the crop has the disease?

$$\frac{\left(0.90\right)\left(0.01\right)}{\left(0.90\right)\left(0.01\right)+\left(0.02\right)\left(0.99\right)}=\frac{0.009}{0.009+0.0198}$$

$$=\frac{0.009}{0.0288}=0.3125$$

*There is a 31.25% chance of actually having a diseased crop.*

1. Suppose you do a taste test of a new fruit drink and people rate the drink on a scale of one to four. One means the new drink was much worse than a pre-existing drink they had, two means it was slightly worse, three means it was slightly better, and four means it was much better. Below is the data from the survey. Calculate the standard deviation.

|  |  |
| --- | --- |
| *Score* | *Number of people who indicated that score* |
| 1 | 40 |
| 2 | 80 |
| 3 | 60 |
| 4 | 20 |

*First, we have to find the mean which is the same as the weighted average.*

$$\frac{1\left(40\right)+2\left(80\right)+3\left(60\right)+4(20)}{40+80+60+20}$$

*So our mean is 2.3. Now we subtract the mean from the observation value and that difference is squared. We will do this multiple times per score (for example, we will do this forty times for the value of one).*

$$40(1-2.3)^{2}+80(2-2.3)^{2}+60(3-2.3)^{2}+20(4-2.3)^{2}$$

$$40\left(1.69\right)+80\left(0.09\right)+60\left(0.49\right)+20\left(2.89\right)=162$$

*Finally we divide by N (which is 200) and take the square root.*

$$σ=\sqrt{\frac{162}{200}}=\sqrt{0.81}=0.9$$