1. Don’t EVEN open this until you are told to do so.

2. There are 20 multiple-choice questions on this exam, each worth 5 points. There is partial credit. Please mark your answers clearly on the scantron. Multiple marks will be counted wrong.

3. Please turn in BOTH YOUR SCANTRON AND YOUR EXAM. Since you may not get your copy back, BE SURE AND MARK YOUR SCANTRON CORRECTLY.

4. You will have 60 minutes to finish this exam.

5. If you are caught cheating or helping someone to cheat on this exam or talking to someone after class about this exam, you both will receive a grade of zero on the exam. You must work alone.

6. This exam is worth 100 points, and will constitute 20% of your final grade.

7. Good luck!
1. Let's say we wanted to test if the sections had the same average on the last test or not. (Well, two of them.) Suppose that the sample averages were 73.2 and 67.5, and we ran a hypothesis test and got the output above. What can we conclude?

A. Of course, there's a difference, 73.2 is a C and 67.5 is a D.
B. We would fail to reject since the p-value is large and claim we have insufficient evidence to prove the true averages are different.
C. We would reject since the p-value is small and claim the true averages are different.
D. We would fail to reject since the p-value is small and claim the true averages are different.
E. We reject since the p-value is large and claim we have insufficient evidence to prove the true averages are different.

2. Which of the following is the best interpretation of the p-value above?

A. 32.7% of the students in either section passed the last test.
B. 32.7% of the time the section averages will differ even though the true mean is the same.
C. 32.7% of the time the section averages will be at least as different as 73.2 and 67.5 even though the true average is the same for both sections.
D. 32.7% of the time the true means will be the same even though the sample means are different.
E. 32.7% of the sections made between 67.5 and 73.2 on the last test.

3. What does power have to do with hypothesis testing?

A. The greater the power of the test, the better the test is at detecting false $H_0$'s.
B. The greater the sample size, $n$, the greater the power.
C. The smaller the $\alpha$, the smaller the power.
D. All of the above are true statements.
E. Only two of the above are true statements.

4. Which of the following is the largest?

A. $P(\bar{X}_{25} > 30)$, where $\bar{X}_{25} \sim N(25, 5^2)$
B. $P(\bar{X}_{50} > 20)$, where $\bar{X}_{50} \sim N(25, 2^2)$
C. $P(\bar{X}_{50} > 30)$, where $\bar{X}_{50} \sim N(25, 2^2)$
D. $P(\bar{X}_{50} > 20)$, where $\bar{X}_{50} \sim N(25, 5^2)$
E. $P(\bar{X}_{25} > 20)$, where $\bar{X}_{25} \sim N(25, 5^2)$

5. Using the output above is all based on the same data. What is the correct range of the p-value for testing $H_0 : \mu = 17$ vs. $H_A : \mu \neq 17$?

A. $p$-value > 0.10
B. 0.10 > $p$-value > 0.05
C. 0.05 > $p$-value > 0.01
D. 0.01 > $p$-value
E. Confidence intervals don’t give p-values.

6. Why do we need the assumptions under statistical inference (hypothesis testing and confidence intervals)?

A. The assumptions are what guarantee that our test statistic, e.g., $X$ or $p$, follows the correct distribution (so our percents and probabilities are valid).
B. The assumptions are what guarantee that our test statistic, e.g., $X$ or $p$, follows the normal distribution (so our percents and probabilities are valid).
C. The assumptions are what guarantee that our data follows the correct distribution.
D. The assumptions are what guarantee that our data follows the normal distribution (as long as $n$ is large enough).
E. The assumptions are necessary because Statistics is a rigid discipline.
7. The government requires that water be sampled and tested for safety every day. If the worker responsible for this testing uses a null hypothesis that means the water is safe, which of the following statements correctly describes a Type I and II error and the appropriate $\alpha$ to use?

A. Both errors would cause unnecessary panic since a mistake means something went wrong, so we should use $\alpha = 0.05$.
B. A Type I error would mean that the citizens were not warned of contaminated water, but a Type II error would mean that the citizens would panic although the water was ok, so $\alpha = 0.01$ should be used.
C. A Type II error would mean that the citizens were not warned of contaminated water, but a Type I error would mean that the citizens would panic although the water was ok, so $\alpha = 0.10$ should be used.
D. A Type II error would mean that the citizens were not warned of contaminated water, but a Type I error would mean that the citizens would panic although the water was ok, so $\alpha = 0.01$ should be used.
E. A Type I error would mean that the citizens were not warned of contaminated water, but a Type II error would mean that the citizens would panic although the water was ok, so $\alpha = 0.10$ should be used.

8. Remember the TAAS tests? This time the sixth grade teacher wants to know how much her students remember from fifth grade, so she readministers the fifth grade test at the beginning of the sixth grade year. What case should she use?

A. Case 11, compare the proportion who passed in fifth grade with the proportion who passed now.
B. Case 9, compare the class average from fifth grade with the new class average.
C. Case 10, compare the difference in their scores now and last year (fifth grade) with zero.
D. Case 3, compare the new average with the school's overall average from last year.
E. Case 6, compare the proportion who pass now with the last year's pass rate.

9. In the scenario above, in any of the cases listed, the null hypothesis would say that the students did the same as last year, so the alternative would be that the students did worse (it'd be great if they did better, but we're only concerned about what or how much they forgot over the summer). Which of the following would be considered a Type I error?

A. We conclude that the students did worse when they really know all the same stuff.
B. We conclude that, although we couldn’t prove they know all the same stuff, they do.
C. We conclude that, although we couldn’t prove they know less, they do (they forgot alot of stuff).
D. We conclude that they forgot alot when they really did forget everything!
E. If the students don’t make the same scores, of course, they forgot stuff.

10. Suppose we calculated a 95% confidence interval using the scores from the sixth grade class: (62.4, 91.5). Which of the following would be true?

A. Less than 95% of the class made below 70 since 70 is below the center of the interval.
B. At least 95% of the class passed (assuming a 60 is passing) since the entire interval is above 60.
C. About 95% of the class passed (assuming a 60 is passing) since the entire interval is above 60.
D. There’s a 95% passing rate for this class since the entire interval is above 60.
E. None of the above are correct.
11. Remember the owner of ‘Lotsaburp’? Well, the regulating agency that controls ‘truth in advertising’ made him up the amount he puts in his cans (since he obviously didn’t have enough before). He wants to check to make sure he didn’t up it too much, so again he tests $H_0 : \mu = 12$ vs. $H_A : \mu > 12$. He sampled 50 cans again and got a p-value of 0.42. What does this mean?

A. He’s putting in 12.42oz in each can, on average.
B. Although his sample average was above 12 oz, it wasn’t enough for him to conclude that he’s putting too much in the cans on average.
C. Although his sample average was above 12 oz, it was only 0.42oz over, so he can’t justify reducing the amount.
D. His sample was obviously over 12 oz this time (a p-value less than 0.5 means his sample mean was greater than 12), so he should reduce the amount he puts in his can.
E. He should increase the amount he puts in each can since his p-value was NOT statistically significant.

12. What conclusion should be made based on the output above? The test was $H_0 : \mu = 25$ vs. $H_A : \mu < 25$.

A. Since the p-value = 0.017, there’s only a 1.7% chance we would reject the null.
B. Since the p-value = 0.017, there’s only a 1.7% chance we would fail to reject the null.
C. Since the p-value = 0.017, we would conclude at the 5 and 10% levels that the true mean is not less than 25.
D. Since the p-value = 0.017, we would conclude at the 5 and 10% levels that the true mean is not less than 25.

E. Since the p-value = 0.017, we would conclude at the 5 and 10% levels that the true mean is more than 25.

13. Assuming that we rejected in the test above, (it doesn’t matter what you choose), and we later find out that the true mean is actually 20, what happened?

A. We made the correct decision since our sample mean was also 20.
B. We made a Type II error since we’re not supposed to know what the true mean really is.
C. We made a Type I error since we’re not supposed to know what the true mean really is.
D. We made a Type I error since that’s the only type of error possible with a reject.
E. We made the correct decision since the true mean WAS less than 25.

14. As it turned out, Florida was critical in deciding the winner based on the electoral college votes. There were 14,000 uncounted ballots in Florida. For Gore to win, it would be necessary for him to get 300 more votes than Bush in those 14,000, or 7150 votes (vs. Bush’s 6850). If we assume that these 14,000 votes are a random sample from a population with $\pi_{Gore} = 0.50$, how likely are we to see a ‘sample’ with at least 7150 out of the 14,000 votes for Gore? This probability, which is the p-value for testing $H_0 : \pi_{Gore} = 0.5$ vs. $H_A : \pi_{Gore} > 0.5$, is 0.0056. How should we interpret this p-value?

A. Only 0.56% of the time would we see Gore get at least this many votes if the election was truly tied ($\pi_{Gore} = 0.5$).
B. 99.44% of the states would have elected Bush.
C. Only 0.56% of the time would we see Gore get this many votes if the election was truly tied ($\pi_{Gore} = 0.5$).
D. Only 0.56% of the time would we see Gore get no more than this many votes if the election was truly tied ($\pi_{Gore} = 0.5$).
E. 99.44% of the time, if we continued to hold elections, Bush would win.
15. What conclusion should we draw from this?

A. It would have been practically impossible for Gore to win the election.
B. If Gore actually got this many votes, then he should have won the election.
C. 0.0056 is less than $\alpha = 0.01$, so we would reject the null and conclude Gore the winner ($\pi_{\text{Gore}} > 0.5$).
D. Gore only counted the votes for him and threw out the votes for Bush.
E. Who cares, Bush won.

99% CI for 0-1 proportion $\pi$ (approximate):
$n = 50$, $p = .43$
Lower Limit = 0.24965515
Upper Limit = 0.61034485

16. If you wanted to test whether the true proportion of red M&M’s is 25% or not, but the interval above is the only information you had, what would you conclude?

A. You’re out of luck since you don’t have a hypothesis test to help you decide.
B. 25% is a plausible value for the true proportion, so you would fail to reject and conclude that there’s insufficient evidence to claim the true proportion is not 25%.
C. 25% is not in the interval, so you’d reject 25% as a plausible value and claim the true proportion is not 25%.
D. 25% is not the center of the interval, so you’d reject 25% as a plausible value and claim the true proportion is not 25%.
E. 25% is not less than $\alpha = 1\%$, so you’d reject 25% as a plausible value and claim the true proportion is not 25%.

17. What all affects the sampling distribution of the sample mean, $\bar{X}$?

A. the distribution of the original population
B. the sample mean’s value
C. the size of the sample
D. All of the above affect the distribution of $\bar{X}$.
E. Only two of the above affect the distribution of $\bar{X}$.

18. Okay, so we decided to check more government projects since 5 wasn’t enough and here’s what resulted. Which of the following best describes what happened?

A. Although this shows that the true average waste on urban renewal projects is statistically less than $500,000$, there is really no practical significance between $500,000$ and $499,000$.
B. The p-value is 0.000 which is less than any possible $\alpha$-level, so the source is still wrong, it’s really less than $500,000$.
C. The p-value is 0.000, so there’s no (zero) difference between our sample average and $500,000$.
D. We can’t rerun a hypothesis test just because we didn’t like the answer. This test is invalid.
E. The sample size is only 30 which isn’t enough to say the sample mean is normally distributed.
19. Is there a difference in the true mean cholesterol level of children vs. adults? How would we find out?

A. Take a large random sample of children and their parents, take the difference in their cholesterol levels, and test whether they are the same or not using Case 10.
B. Take a large random sample of children, find the proportion with cholesterol levels above the ‘standard’ for adults, and test whether they are the same or not using Case 6.
C. Take a large random sample of both children and adults, calculate the average cholesterol for both groups, then test whether they are the same or not using Case 9.
D. Take a large random sample of both children and adults, find the proportion in each group with cholesterol levels above the ‘standard’, and test whether they are the same or not using Case 11.
E. Take a large random sample consisting of both children and adults, calculate the average cholesterol level, and test whether it’s above the average using Case 3.

20. Which of the following is true about the t distribution?

A. As the sample size increases, the center (the mean) of the curve increases.
B. As the sample size increases, the center (the mean) of the curve decreases.
C. As the sample size increases, the spread (the standard deviation) of the curve increases.
D. As the sample size increases, the spread (the standard deviation) of the curve decreases.
E. Exactly two of the above are true.

1B,2C,3D,4B,5C,6A,7C,8C,9A,10E,11B