

STAT303 Sec 507
Fall 2011
Exam #1
Form A

Instructor: Julie Hagen Carroll

Name: _____

1. Please PRINT your name in the blank above.
2. **Do not open this exam until you are told to do so.**
3. There are 20 multiple-choice questions on this exam, each worth the same. Please mark your answers **clearly** on a GRAY Scantron sheet. Multiple marks will be counted wrong.
4. You **must mark** your Scantron form with
 - (a) Your NAME and UID.
 - (b) Your correct SECTION (Thursday 11:10 is 507, 12:45 is 508, 2:20 is 509, and 3:55 is 510).
 - (c) This test FORM (A, B, C, or D).
5. You will have 60 minutes to finish this exam.
6. You may use the following.
 - (a) One $8\frac{1}{2} \times 11$ formula sheet (both sides) of your own making.
 - (b) A copy of the standard normal table and the Percent Curve handout.
 - (c) A stand-alone calculator, i.e., one that cannot communicate with the internet or anything outside itself.
7. You must put all possessions besides the materials listed under your table out of everyone else's way. This includes cell phones, which must be turned off.
8. If you have questions, please write out what you are thinking on the back of this test page so that we can discuss it after your results are returned to you.
9. If you are caught cheating or helping someone to cheat on this exam, you both will receive a grade of **zero** on the exam. You must work alone.
10. When you are finished please make sure you have marked your Section and Form and have an answer for every question, then turn in your Scantron form and show your ID. Turn your exam over(upside down) and let it at your computer. Please sit in the same spot next week so that you can get this back.
11. Good luck!

1. A researcher is interested in the effects of sunlight on elderly people (i.e., those over 65 years of age) living in Texas. Which of the following ways of sampling might tend to produce a *biased* view of the effect of sunlight on elderly Texans?

- A. Fifty people are randomly selected from a list of all elderly people in Texas.
- B. One hundred elderly people are randomly selected from a large nursing home in west Texas.
- C. The first thousand elderly people who respond to a phone survey make up the sample.
- D. All the above.
- E. Both B and C.

2. What is $P(1.06 < Z < 2.80)$?

- A. 0.142
- B. 0.1472
- C. 0.0409
- D. 0.9591
- E. 1.8528

3. If we had taken a sample of size 10 from the distribution above and found the probability for the sample mean, which of the following would be true?

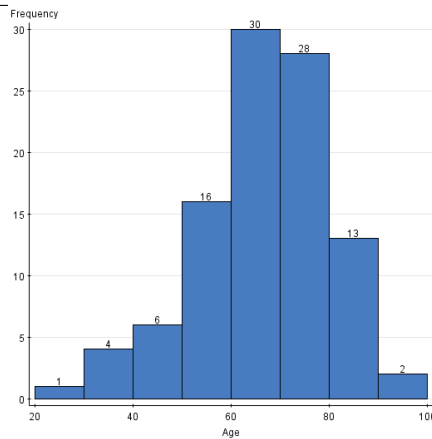
- A. The area(probability) would be the same since it's the same distribution.
- B. The area would be smaller since means vary less than individuals.
- C. The area would be smaller since this is a tail area and means have smaller (shorter) tail areas.
- D. The area would be larger since it's for 10, not just one.
- E. We don't know that the distribution is normal since n is only 10.

4. Sample size, n , affects

- A. the shape of the population distribution
- B. the shape of the sample distribution (distribution of the sample)
- C. the shape of the sampling distribution
- D. the variance of the population distribution
- E. the variance of the sample distribution (distribution of the sample)

5. An entrepreneurial student makes her large front yard available for parking on game day. She keeps meticulous records of how many trucks and how many cars park there. Over several seasons she has had an average of 10.2 trucks with a standard deviation of 2.1. Which row of the table below is most likely to be the 5 number summary of her truck counts?

	Min	Q1	Med	Q3	Max
A	0	5	6	10	15
B	2	5	9	11	13
C	4	8	10	12	16
D	6	10	12	14	18
E	7	9	10	11	12



6. Assuming each bin above contains ONLY the value to the left, which of the following describe this histogram?

- A. The 5-Number-Summary is 20, 50, 60, 70, 100
- B. The mean is less than the median since it is skewed to the left.
- C. The mean is less than the median since it is skewed to the right.
- D. The mean is greater than the median since it is skewed to the left.
- E. Two of the above are correct descriptions.

7. In the histogram above, what would happen if the 20-year-olds became 40-year-olds?

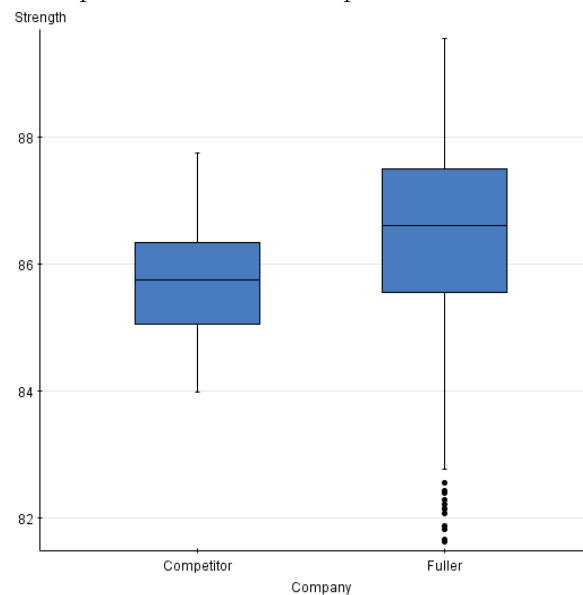
- A. Nothing since it would still be the same data.
- B. The mean and standard deviation would increase.
- C. The mean would increase but the standard deviation would decrease.
- D. The mean and the median would increase but the IQR and standard deviation would stay the same.
- E. The mean and the median would increase but the IQR and standard deviation would decrease.

8. Suppose that the serum cholesterol levels of males in a given age bracket have a mean of 223.9 and a standard deviation of 15.8. Two researchers take random samples from this population of males. Researcher A takes a random sample of 40 males, and researcher B takes a sample of 80 males. What can we say about the distributions of the sample means computed by the two researchers?

- A. They have the same mean and variance because they come from the same population.
- B. The distribution for A has a smaller mean and a smaller variance than for B because the sample size is smaller.
- C. The distribution for A has a larger mean and a larger variance because the sample size is smaller.
- D. Both distributions have the same mean. The distribution for A has a smaller variance than for B because the sample size is smaller.
- E. Both distributions have the same mean. The distribution for A has a larger variance than for B because the sample size is smaller.

9. A sampling distribution
- is a distribution of all the sample statistics that can be found for one sample.
 - of any statistic has an approximately normal distribution.
 - is a distribution of a sample.
 - of the mean is a distribution of the means computed from all possible samples of a given size n that could be taken from the population.
 - Two of the above are correct.
10. Which of the following are important principles in sound experimental design?
- Controlling the effect of lurking variables.
 - Hand-picking experimental subjects to avoid the poor choice of subjects produced by randomization.
 - Replicating each treatment on enough subjects to reduce chance variation in the results.
 - All the above.
 - Two of the above.
11. Which of the following is an appropriate graph or summary number for categorical data?
- a stem-and-leaf plot
 - a mean, \bar{x}
 - a proportion (or a probability)
 - a median, \tilde{x}
 - None of the above are appropriate
12. Which two values for Z have 96.52% of the data in between them?
- 0.0348 and 0.0348
 - 1.82 and 1.82
 - .0174 and 0.0174
 - 2.11 and 2.11
 - 3.48 and 3.48
13. Suppose that the volume of Pepsi in a "2 liters" bottle is normally distributed with a mean 2 liters and a standard deviation 0.08 liters. If I purchase 4 bottles of 2-liter Pepsi and find that the average volume is 2.1 liters. What is the probability that an average volume of 4 bottles of 2-liter Pepsi will exceed 2.1 liters?
- 0.9938
 - 0.0062
 - 0.8944
 - 0.1056
 - None of the Above.
14. The five number summary for total arsenic levels in eighty five 8 oz. boxes of apple juice was Min = 4.2ppb, $Q_1 = 6.1$ ppb, Med = 8.3ppb, $Q_3 = 11.6$ ppb, Max = 18.7ppb. Which of the following is most reasonable?
- The data have a normal distribution.
 - The data are approximately symmetric in distribution.
 - The data are skewed left.
 - The data are skewed right.
 - The data are bimodal.

15. A parameter is
- a numerical characteristic of a sample.
 - a numerical characteristic of a population.
 - a special meter used to monitor the results of medical experiments.
 - a well-designed experiment.
 - a treatment applied to subjects that volunteer for an experiment.
16. In California, all high school students must pass a test, CAHSEE, to earn a high school diploma. The 2010-11 CAHSEE results show that the scores of English-Language Art exam are normally distributed with mean 373 and standard deviation 34. Furthermore, 25% of the exam takers did not pass the English-Language Art exam. What was the minimum score to pass the English-Language Art exam?
- 305
 - 339
 - 350
 - 396
 - None of the Above.
17. Fuller Bottle Company wants to compare their new bottle with their competitor's. The 120 impact strengths of Fuller bottles is shown in a boxplot below, alongside a sample of 80 from the competitor.

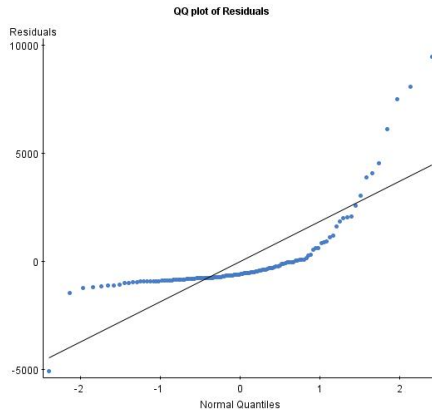


Which of the following best describes how the two companies compare?

- Fuller's bottles tend to be stronger than the competitor's.
- Fuller's bottles are less variable than the competitor's.
- Fuller's bottles are more symmetrically distributed than the competitor's.
- Fuller's bottles tend to be stronger but are more variable than the competitor's.
- Only 50% of Fuller's bottles have strength greater than the competitor's median.

18. Which of the following is/are true?

- A. Outliers affect all of the statistics we have talked about.
- B. Shift changes only affect measures of location.
- C. Scales changes only affect measures of spread.
- D. All of the above are true.
- E. Only two of the above are true.



19. For the dataset represented above, we can

- A. say the distribution is fairly normal since the points are bell-shaped.
- B. say the distribution is symmetry but not necessarily normal.
- C. use the 68-95-99.7% rule.
- D. Two of the above are true.
- E. None of the above are true.

20. Why are we not good randomizers?

- A. We allow prior knowledge to affect our choice.
- B. We allow our environment to affect our choice (think size and location).
- C. We know what our results should look like, so we pick our observations to fit.
- D. All of the above are reasons.
- E. Only two of the above are reasons.

21. Bonus: This is in the notes, but we didn't talk about it. What is z^* such that $P(-z^* < Z < z^*) = 0.65$ and $Z \sim N(0, 1^2)$

- A. 0.385
- B. 0.3711
- C. 0.7422
- D. 0.935
- E. 0.6368

1E,2A,3C,4C,5C,6B,7C,8E,9D,10E,11C,
12D,13B,14D,15B,16C,17D,18B,19E,20D,21D