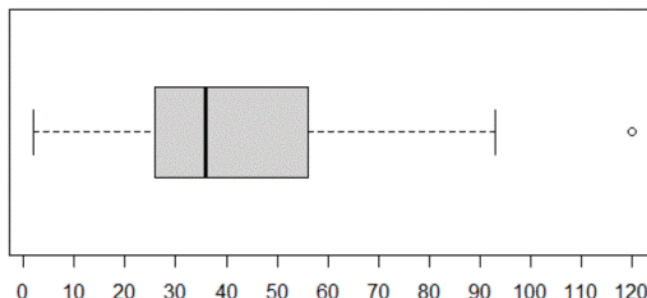


## Statistics Exam - University of Houston Math Contest January 29, 2022

1) The following graph is a boxplot of stopping distance in feet of 50 cars.



Which of the following is the correct statement?

- a) The maximum value is at 92 feet.
- b) The stopping distance is skewed left.
- c) About 25 of the cars had a stopping distance between 25 feet and 55 feet.
- d) The mean stopping distance is at 35 feet.
- e) None of the answer choices provided.

2) Which is a FALSE statement about the following stem-and-leaf plot?

**The decimal point is 1 digit(s) to the right of the |**

```

0 | 5
1 | 3567
2 | 08888
3 | 000123448
4 | 3
    
```

- a) The range is 38.
- b) The interquartile range is 14.
- c) The first quartile is 18.5
- d) The mean is 29.
- e) The maximum is 43.

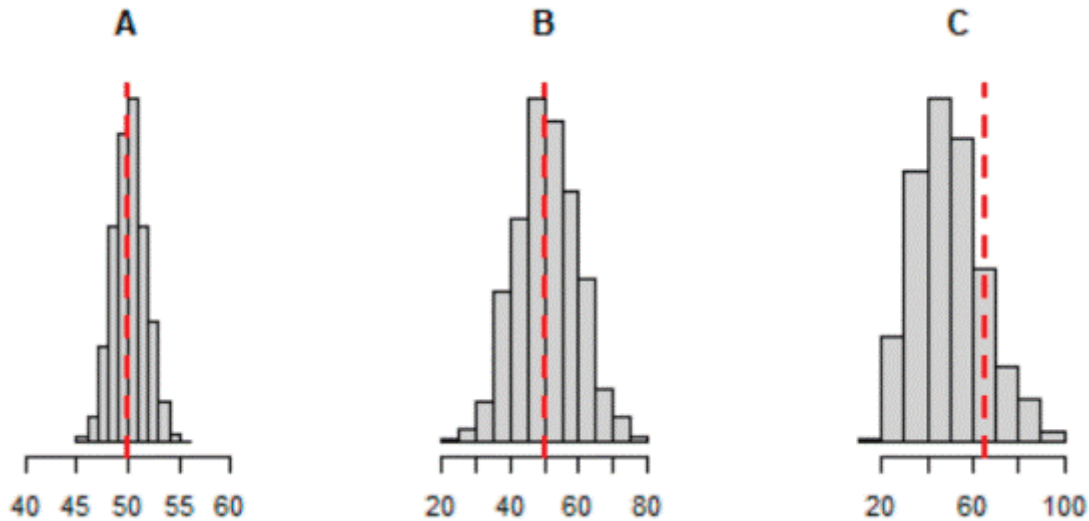
3) Which list of values has the largest standard deviation?

- a) 100,100,100,100,100
- b) 0,0,5,10,10
- c) 0,1,2,3,4
- d) 52,54,55,56,58
- e) 0,2,5,7,10

4) Suppose a random variable  $X$  has a mean of 50 and variance 25. Suppose a random variable  $Y = 3X - 10$  is created. Determine the mean and standard deviation of  $Y$ .

- a) Mean=140,standard deviation =75
- b) Mean=140,standard deviation =5
- c) Mean=140,standard deviation =15
- d) Mean=150,standard deviation =5
- e) Mean=150,standard deviation =65

5) The histograms below describe the values taken by three sample statistics in 1000 samples from the same population. The true value of the population parameter is marked on each histogram by a dotted red line.



The statistic that has the largest bias among these three is:

- a) A and B have similar bias, and is larger than the bias of C.      b) A      c) C      d) B  
 e) All three have similar bias.

6) The name for the pattern of values that a statistic takes when we sample repeatedly from the same population is

- a) The bias of the statistic      b) The variability of the statistic      c) The distribution of the statistic.  
 d) The scale of measurement of the statistic      e) The sampling error.

7) Sale of eggs that are contaminated with salmonella can cause food poisoning among consumers. A large egg producer takes a simple random sample of 200 eggs from all the eggs shipped in one day. The laboratory reports that 9 of these eggs had salmonella contamination. Unknown to the producer, 0.1% (one-tenth of one percent) of all eggs shipped had salmonella. In this situation,

- a) 9 is a parameter and 0.1% is a statistic.      b) Both 0.1% and 9 are parameters.  
 c) 0.1% is an estimate and 9 is a margin of error.      d) 0.1% is a parameter and 9 is a statistic.  
 e) Both 0.1% and 9 are statistics.

8) An ad for a new mouthwash says that it "reduces plaque on teeth by 300 percent." What does this mean?

- a) It's nonsense, because plaque is a categorical variable, so percents do not make sense.  
 b) It means that three-tenths of the plaque is removed, because  $300/1000 = 0.3$ , or three-tenths  
 c) It means that there is 3 times as much plaque before using the mouthwash as there is after using it.  
 d) It's nonsense, because removing 100 percent of the plaque already removes all of it.  
 e) It's nonsense, because percent only make sense for counts, and amount of plaque is not a count.

9) When we take a census, we attempt to collect data from

- a) Every individual in the population
- b) A stratified random sample
- c) Every individual selected in a simple random sample
- d) A convenience sample
- e) A voluntary response sample

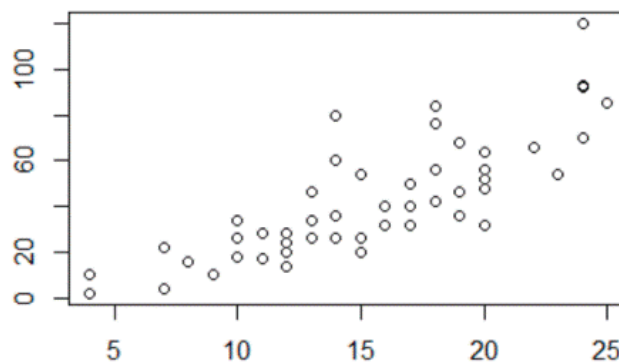
10) A personnel director at a large company studied the eating habits of employees by watching the movements of a selected group of employees at lunchtime. The purpose of the study was to determine the proportion of employees who buy lunch in the cafeteria, bring their own lunches, or go out to lunch. The study could best be categorized as:

- a) a survey sample
- b) an observational study
- c) a designed experiment
- d) a voluntary response study
- e) a census

11) After we take a sample, we want to estimate a population \_\_\_\_\_ by a sample \_\_\_\_\_. Because samples rarely give exactly correct results, we also give a margin of error. The margin of error includes \_\_\_\_\_, but it does not include the effects of \_\_\_\_\_. The sequence of terms that makes this statement correct is:

- a) Statistic, parameter, chance variation, nonsampling error.
- b) Statistic, parameter, nonsampling error, chance variation.
- c) Parameter, statistic, chance variation, probability.
- d) Parameter, statistic, chance variation, nonsampling error.
- e) Parameter, statistic, nonsampling error, chance variation.

12) Determine the value of the correlation coefficient based on the following scatterplot below.



- a) 0.1
- b) -0.1
- c) 1.0
- d) 0.8
- e) -0.8

13) The five-number summary of the distribution of exam scores was as follows:

18, 39, 62, 76, 100

A total of 416 students took the exam. About how many students had a score above 39?

- a) 328
- b) 416
- c) 312
- d) 104
- e) 400

14) A manufacturer purchased 200 units of a product for \$5.50 each, 100 units of another product for \$7.20 each, and 150 units of a third product for \$3.70 each. What is the mean cost of the purchases?

- a) \$5.28      b) \$5.47      c) \$7.92      d) \$3.75      e) \$16.40

15) For categorical variables, it only makes sense to speak of \_\_\_\_\_. It makes no sense to speak of the \_\_\_\_\_ or the \_\_\_\_\_ for the variable. The sequence of words to correctly complete the statement is

- a) Mean, standardize, median                      b) Median, mean, standard deviation  
c) Correlation, mean, standard deviation        d) Standard deviation, mean, median  
e) Counts or percents, mean, median

16) A crate contains 12 light bulbs of which 5 are defective. A Quality Control Inspector randomly samples 4 bulbs without replacement. What is the probability that at least one light bulb is defective?

- a) 0.3308      b) 0.0141      c) 0.9293      d) 0.8842      e) 0.0101

17) Suppose that 12% of drivers do not have insurance. A police officer stops 25 people for violations. What is the probability that at least 3 drivers do not have insurance?

- a) 0.2174      b) 0.6475      c) 0.7826      d) 0.5912      e) 0.3525

18) Suppose it is known that 5% of adults who take a certain medication experience negative side effects. A doctor prescribes this medicine to 35 patients. How many patients does the doctor expect to experience negative side effects? (Round to the nearest whole number)

- a) 1      b) 17      c) 20      d) 2      e) 5

19) A service organization in a large school organizes a raffle. One thousand raffle tickets are sold for \$1 each. Each has an equal chance of winning. First prize is \$300, second prize is \$200, and third prize is \$100. Let  $X$  denote the net gain from the purchase of one ticket. What is the expected net gain from the purchase of one ticket?

- a) 0.6      b) -0.4      c) 200      d) 300      e) -1

20) A discrete random variable,  $X$ , has the following probability distribution

$X$	-1	0	1	2
$P(X)$	0.2	0.1	0.3	?

Determine the standard deviation of  $X$ .

- a) 0.9      b) 2.1      c) 0.7351      d) 1.1358      e) 1.29

21) Let A and B be two events. Suppose that  $P(A) = 0.4$ ,  $P(B) = 0.3$ , and  $P(A \cup B) = 0.8$ . Which of the following statements are true?

- i. A and B are mutually exclusive,
- ii. A and B are independent,
- iii. The probabilities of A and B are impossible.

a) both i and ii are true      b) ii is true      c) iii is true      d) i is true      e) neither i, ii, nor iii are true

22) Let A and B be two events. Suppose that  $P(A) = 0.2$ ,  $P(B) = 0.3$ , and  $P(A \cup B) = 0.44$ . Which of the following statements are true?

- i. A and B are mutually exclusive,
- ii. A and B are independent,
- iii. The probabilities of A and B are impossible.

a) ii is true      b) iii is true      c) i is true      d) both i and ii are true      e) neither i, ii, nor iii are true

23) Which of the following terms means the difference in the results did not occur by random chance?

- a) ANOVA      b) Sample      c) Inferential statistics      d) Statistical significance
- e) None of the options are correct

24) Given:  $P(E) = 0.35$ ,  $P(F) = 0.53$ , and  $P(E \cup F) = 0.71$ . Find  $P(E \cap F)$  and  $P(F | E)$ .

- a)  $P(E \cap F) = 0.17$  and  $P(F | E) = 0.53$
- b)  $P(E \cap F) = 0.17$  and  $P(F | E) = 0.49$
- c)  $P(E \cap F) = 0.19$  and  $P(F | E) = 0.49$
- d)  $P(E \cap F) = 0.29$  and  $P(F | E) = 0.32$
- e)  $P(E \cap F) = 0.19$  and  $P(F | E) = 0.53$

25) A blood test indicates the presence of a particular disease 95% of the time when the disease is actually present. The same test indicates the presence of the disease 0.5% of the time when the disease is not present. One percent of the population actually has the disease. Determine the probability that a person has the disease given that the test indicates the presence of the disease.

- a) 0.95      b) 0.0145      c) 0.1610      d) 0.6574      e) 0.0095

26) John scored a 90 on a test in Statistics where the average score was 75 points and standard deviation 10. He scored 85 on a test in English where the average score was 70 points and standard deviation was 7.5. Which class did John rank higher?

- a) Statistics      b) English      c) John's rank was the same in both classes.

27) Scores of adults on the Wechsler Adult Intelligence Scale (a common "IQ" test) follow a Normal distribution, with mean at 100. Approximately 68% of all scores fall between 85 and 115. What is the standard deviation of this test score?

- a) 225      b) 30      c) 15      d) 45      e) Cannot determine with this information.

28) The weight of a certain type of apple is normally distributed with a mean of 8 ounces and standard deviation of 0.96 ounces. What is the first quartile,  $Q_1$ , of the weight of this type of apple?

- a) 8.6475      b) 8      c) 8.24      d) 3.432      e) 7.3525

29) The average life a manufacturer's blender is 60 months, with a standard deviation of 12 months. Assuming that the lives of these blenders follow approximately a normal distribution, find the probability that the mean life a random sample of 9 such blenders falls between 4.4 and 5.2 years.

- a) 0.2743      b) 0.6898      c) 0.3050      d) 0.3102      e) 0

30) Fifty people (25 men and 25 women) were asked what type of movie they would prefer to watch and the following results are recorded in the following two-way table.

	Romance	Action	Horror
Men	5	14	6
Women	10	11	4

What percent of the women prefer to watch a Romance movie?

- a) 80%      b) 60%      c) 20%      d) 30%      e) 40%

31) Which of the following would be the most appropriate test of significance for the following situation? Does the temperature of what you drink affect your mouth temperature? Randomly selected subjects first had their mouth temperature taken. Then, the same subjects were randomly assigned to drink 1/4 cup of cold or hot water and hold it in their mouths for 10 seconds. Then, their mouth temperatures were measured again. After 20 minutes, the process was repeated using the other water temperature treatment.

- a) Two independent sample t-test for means  
b) Matched pairs t-test for mean  
c) One sample z-test for proportion  
d) One sample z-test for mean  
e) None of the choices are appropriate

32) If the 90% confidence limits for the population mean are 34 and 46, which of the following could be the 97% confidence limits

- a) (39,43)      b) (39,41)      c) (36,41)      d) (38,45)      e) (32,48)

33) An important problem in industry is shipment damage. A pottery producing company ships its product by truck and determines that it can meet its profit expectations if, on average, the number of damaged items per truckload is fewer than 11. A random sample of 19 departing truckloads is selected at the delivery point and the mean number of damaged items per truckload is calculated to be 9.4 with a calculated sample of variance of 0.49. Select a 95% confidence interval for the true mean of damaged items. Assume that the number of damaged items has a Normal distribution.

- a) (10.6626, 11.3374)      b) (9.0626, 9.7374)      c) (9.1638, 9.6362)  
d) (9.0852, 9.7148)      e) (9.1767, 9.6203)

34) An experimenter flips a coin 100 times and gets 43 heads. Test the claim that the coin is fair against the two-sided claim that it is not fair at the 1% significance level. Give the null and alternative hypothesis, test statics, and decision.

- a)  $H_0$  is  $p = 0.5$ ,  $H_a$  is  $p \neq 0.5$ ,  $z = -1.40$ , Reject  $H_0$  at the 1% significance level.  
b)  $H_0$  is  $p = 0.5$ ,  $H_a$  is  $p \neq 0.5$ ,  $z = -1.40$ , Fail to reject  $H_0$  at the 1% significance level.  
c)  $H_0$  is  $p = 0.5$ ,  $H_a$  is  $p < 0.5$ ,  $z = -1.41$ , Fail to reject  $H_0$  at the 1% significance level.  
d)  $H_0$  is  $p = 0.5$ ,  $H_a$  is  $p \neq 0.5$ ,  $z = -1.41$ , Reject  $H_0$  at the 1% significance level,  
e)  $H_0$  is  $p = 0.5$ ,  $H_a$  is  $p < 0.5$ ,  $z = -1.41$ , Reject  $H_0$  at the 1% significance level.

35) Suppose that prior to conducting a coin-flipping experiment, we suspect that the coin is fair. How many times would we have to flip the coin in order to obtain a 96.5% confidence interval of width of at most .16 for the probability of flipping a head? (note that the z-score was rounded to three decimal places in the calculation)

- a) 129      b) 178      c) 171      d) 174      e) 126

36) It has been observed that some persons who suffer colitis, again suffer colitis within one year of the first episode. This is due, in part, to damage from the first episode. The performance of a new drug designed to prevent a second episode is to be tested for its effectiveness in preventing a second episode. In order to do this two groups of people suffering a first episode are selected. There are 170 people in the first group and this group will be administered the new drug. There are 132 people in the second group and this group will be administered a placebo. After one year, 14% of the first group has a second episode and 18% of the second group has a second episode. Select a 90% confidence interval for the difference in true proportion of the two groups.

- a) (-0.610, 0.530)      b) (-0.030, -0.013)      c) (-0.044, 0.124)      d) (-0.124, 0.044)      e) (-0.110, 0.030)

37) Suppose you have the following data:

X	1	2	3	4	5	6
Y	43	37	29	21	20	13

Based on the linear regression line, determine the residual for  $X = 5$ .

- a) 20      b) -1.79      c) 1.79      d) 18.212      e) 5

38) If the least squares regression line relating the independent variable  $x$  and the dependent variable  $y$  for a given problem is  $\hat{y} = 2x + 5$ . Then an increase of 1 unit in  $x$  is associated with an increase in how many units in  $y$ ?

- a) 0      b) 5      c) 1      d) 2      e) Cannot be determined.

39) A can of Pepsi is supposed to have a mean volume of 12 ounces. Both overfilling and under-filling are undesirable. If either occurs, the machine that fills the cans has to be readjusted. A statistician conducted a significance test on a sample of 50 cans. Based on this test, the decision was to readjust the machine. However, it was discovered that the machine was filling the cans with a mean volume of 12 ounces. This is an example of:

- a) Type 1 error      b) Type 2 error      c) Correct decision      d) Type 3 error      e) None of these answers

40) A random sample of 100 rolls of a dice was conducted to test if the dice is fair, (that is each face of the dice comes up equally). The following was the result of the 100 rolls.

Face	1	2	3	4	5	6
Count	19	13	19	20	15	14

Construct the proper hypothesis test to determine if the dice is fair. Give the test statistic of this test.

- a)  $Z = 2.72$       b)  $X^2 = 2.72$       c)  $X^2 = 5$       d)  $X^2 = 100$       e)  $Z = 5.31$

41) The level of significance of a statistical test indicates

- a) The significant difference between the means  
b) The chance we were wrong in rejecting the null hypothesis  
c) The chance we are right in to rejecting the null hypothesis  
d) Whether to accept or reject the null hypothesis  
e) The chance that the alternative hypothesis is true

42) Which statement is the best, as a general rule, for interpreting the level of significance of a statistical test?

- a) If p-value is less than 0.10, there is statistical significance.  
b) If p-value is greater than 0.1, we accept the null hypothesis.  
c) If p-value is greater than 0.10, there is marginal significance.  
d) If p-value is less than 0.01, there is no statistical significance.  
e) If p-value is less than 0.05, there is practical significance.



43) A study reports the mean change in HDL (high-density lipoprotein, or "good" cholesterol) of adults eating raw garlic six days a week for six months. The margin of error for a 95% confidence interval is given as plus or minus 4 milligrams per deciliter of blood (mg/dl). This means that

- a) We are 95% confident that the true population mean is within 4 mg/dl of the calculated sample mean.
- b) 95% percent of the population has changed their HDL after eating raw garlic six days a week for six months.
- c) We could be certain that the study result is within 4 mg/dl of the truth about the population if the conditions for inferences were satisfied.
- d) We can be certain that the study results is within 4 mg/dl of the truth about the population.
- e) There is a 95% probability that the true population mean is within the calculated interval.

44) Let X represent the hemoglobin count (HC) in grams per 100 milliliters of whole blood. The distribution for HC is approximately normal with  $\mu = 14$  for healthy adult women. Suppose that a female patient has taken 12 laboratory blood samples in the last year. The HC data sent to her doctor is listed as follows: 17,19,17,17,21,15,22,16,19,19,17,14. We would like to know if the data indicates this patient has significantly high HC compared to the population. Give the p-value and interpret the results.

- a) P-value = 0.0001; Based on 5% significance level, I will fail to reject the null hypothesis and conclude this patient does not have a significantly high HC level.
- b) P-value = 0.0002; Based on 5% significance level, I will reject the null hypothesis and conclude this patient has a significantly high HC level.
- c) P-value = 0.0562; Based on 5% significance level, I will fail to reject the null hypothesis and conclude this patient does not have a significantly high HC level.
- d) P-value = 0.1053; Based on 5% significance level, I will fail to reject the null hypothesis and conclude this patient does not have a significantly high HC level.
- e) P-value = 0.0001; Based on 5% significance level, I will reject the null hypothesis and conclude this patient has a significantly high HC level.

45) In a test of significance, assuming the null hypothesis is true, the probability of observing the test statistic extreme or more extreme than the observed test statistic (in the way of the alternative hypothesis) is

- a) The power of the test.
- b) The level of significance,  $\alpha$
- c) The probability the null hypothesis is true
- d) The probability the null hypothesis is false
- e) The p-value

46) Solid fats are more likely to raise blood cholesterol levels than liquid fats. Suppose a nutritionist analyzed the percentage of saturated fat for a sample of 6 brands of stick margarine (solid fat) and for a sample of 6 brands of liquid margarine and obtained the following results:

Stick:[25.5,26.7,26.5,26.6,26.3,26.4],

Liquid:[16.5,17.1,17.5,17.3,17.2,16.7]

We want to determine if there a significant difference in the average amount of saturated fat in solid and liquid fats. What is the test statistic?

- a) Z = 39.604
- b) T = 39.604
- c) T = 25.263
- d) Z = 39.104
- e) T = 39.104

47) In an experiment on relaxation techniques, subject's brain signals were measured before and after the relaxation exercises with the following results:

Person	1	2	3	4	5
Before	31	38	66	51	32
After	26	35	57	51	26

Is there sufficient evidence to suggest that the relaxation exercise slowed the brain waves? (Use  $\alpha=0.05$ ) Which is the correct conclusion?

- a) There is no evidence that the relaxation techniques slows the brain waves.
- b) There is not enough information to make a conclusion.
- c) The relaxation techniques do not slow the brain waves.
- d) There is strong evidence at 5% that the relations techniques increase the brain waves.
- e) There is strong evidence at 5% level that the relaxation techniques does slow the brain waves.

48) Suppose we have an ordinary linear regression between two variables  $x$  and  $y$  with correlation  $r = 0.81$ . Which statement is true?,

- a) About 81% of  $y$  came from  $x$ .
- b) About 90% of the variation in  $y$  is predictable from  $x$ .
- c) About 65.61% of the variation in  $y$  is predictable from  $x$ .
- d) About 81% of the variation in  $y$  is predictable from  $x$ .
- e) About 65.61% of  $y$  came from  $x$ .

49) The distribution of heights of adult men has a mean of 69 inches and standard deviation of 2.5 inches. How tall is a man whose standardized height is 0.3?

- a) 68.25 inches      b) 68.7 inches      c) 69.3 inches      d) 69.75 inches
- e) Cannot be determined because we do not know the distribution of the heights.

50) **TIEBREAKER QUESTION:** Suppose you were told that a 90% confidence interval for the population mean of mpg of a hybrid car was (27, 43). Determine the standard error for this population mean.

- a) 27      b) 8      c) 4.86      d) 1.645      e) 35