Statistics Exam - University of Houston 2023 Math Contest
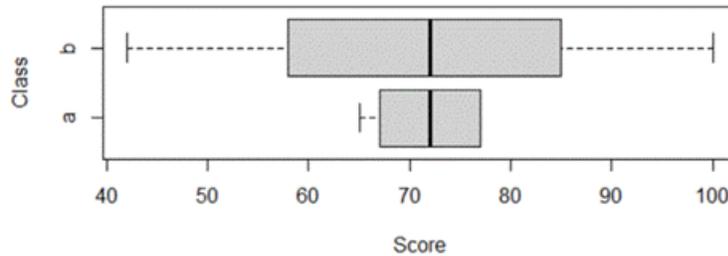## January 28, 2023

1) The test scores of 20 students have a mean of 71.6 and the test scores of another class of 14 students have a mean of 78.4. Determine the mean of the combined classes.
a) 70.8　　　　　　b) 74.4　　　　　　c) 75　　　　　　d) 78.4　　　　　　e) 71.6

2) The following is a boxplot of two classes tests scores, class a and b. Which statement cannot be determined by this graph?

a) The median score is the same for both classes.
b) The maximum score in class a is about 77.
c) The range for the scores in class b is about 60.
d) Class b has more students than class a.
e) Class b scores are more variable than class a.

3) The following is a sample data set of 10 points with sample mean 8.28. One data point is missing.

$$7.25, 4.9, 10.5, 7.35, 15.9, 12.6, 6.5, 4.8, 11.6, \underline{\quad\quad}$$

Determine the sample standard deviation (round to the nearest thousandths).
a) 3.809　　　　　b) 4.329　　　　　c) 4.107　　　　　d) 3.591　　　　　e) Cannot be determined.

4) Fill in the blanks in order: _____ try to gather data without influencing the responses. _____, on the other hand, impose some _____ in order to observe the response.
a) Experiments, Observational studies, Randomization
b) Observational studies, Experiments, Treatment
c) Response, Treatment, Randomization
d) Sample surveys, Response, Treatment
e) Explanatory variable, Response, Randomization

5) An online poll about parking on campus was conducted. Of the 136 people who responded, 79% said that there was not enough parking on campus. The number 79% is a
a) Margin of error　　　b) Parameter　　　c) Probability　　　d) Statistic　　　e) Reliability

6) Do doctors in managed care plans give less charity care? Researchers chose 60 communities at random, then chose doctors at random in each community. In all, they interviewed 10,881 doctors. This is what type of study design?
a) Block experiment design
b) Multistage random sample
c) Stratified random sample
d) Simple random sample
e) Completely randomized experimental design

7) When I set my alarm to go off at 6:30 am, it rings 10 minutes late every day. My alarm is
a) Biased   b) Insufficient   c) Imprecise   d) Invalid   e) Unreliable

8) I flipped a coin several times and computed a 95% confidence interval for the proportion of heads up to be [0.55, 0.61]. How many times did I flip the coin?
a) 500   b) 1000   c) 1040   d) 580   e) 250

9) A 99% confidence interval for the mean amount home deliveries per week has a margin of error of ± 0.05. Determine the standard error of the mean, $SE(\overline{X})$, round to the nearest hundredths (assume a large sample size).
a) 0.13   b) 2.576   c) 0.02   d) 0.05   e) Cannot be determined.

10) A 98% confidence interval for the mean is [10, 20]. If we would change the confidence level to 95% the confidence interval is (assume that the population standard deviation, σ, is known):
a) [10.8, 19.2]   b) [10, 20]   c) [5, 15]   d) [12.7, 17.3]   e) Cannot be determined.

11) An EPA researcher wants to design a study to estimate the mean lead level of fish in a lake located near an industrial area. Based on past sample data, the researcher estimates that σ for the lead level in the fish population is approximately 0.016 mg/g. He wants to use a 98% confidence interval having a margin of error no greater than 0.005 mg/g. How many fish does he need to catch?
a) 8   b) 40   c) 5542   d) 56   e) 7

12) The following are prices of regular unleaded gasoline at random gas stations. Determine a 97% confidence interval for the mean price of gasoline.

$$2.70, 3.10, 3.16, 2.68, 2.92, 2.53, 2.97, 2.89, 2.85, 3.14$$

a) [2.75, 3.04]   b) [2.35, 3.44]   c) [2.72, 3.07]   d) [2.76, 3.02]   e) [2.53, 3.16]

13) A survey is conducted of college bound and non-college bound high school seniors to compare the proportion who have at least one parent who attended college. This study is
a) two independent sample observational studies
b) two independent sample experiment
c) A matched pairs design experiment
d) A single independent sample observational study
e) A matched pairs design observational study

14) Suppose we want to know whether or not the mean weight, in pounds, between two different species (we will call species 1 and 2) of turtles is equal. We determined a 95% confidence interval for the difference in the mean weight between the two species is [-3.08, 13.08]. Determine the true statement.
a) There is no evidence at 5% significance of a difference between the mean weight of the two different species of turtles.
b) We are 95% sure that species 1 has a higher weight.
c) We are 95% sure that species 2 has a higher weight.
d) Species 1 has a mean 13.08 pounds more than species 2.
e) Species 1 has a mean 3.08 pounds less than species 2.

15) The following is a baseball team's record for the season.

	Home Games	Away Games
Won	45	35
Loss	35	45

Based on a 5% level of significance, what is the best statement about the relationship between where the team played and how well it did?

a) There is very strong evidence that the team plays better at home.
b) There is strong evidence that the team plays better at away games.
c) The team wins the same amount of games no matter if they are home or away.
d) There is no evidence that there is an association between where the team played and how well it did.
e) There is a 15% chance that the team loses away games.

16) The following is temperature readings from an office space that is supposed to be 72°F.

$$71.95, 72.30, 71.26, 71.26, 71.59, 72.07, 72.00, 71.29, 72.37, 71.30, 71.95, 72.30, 71.48, 71.24$$

Based on a 5% level of significance, create a hypothesis test to determine if the mean temperature is significantly less than 72°F. Give the test statistic, p-value and decision of this test.
a) T = -2.2247, p-value = 0.9778, fail to reject the null hypothesis
b) T = -2.2247, p-value = 0.0222, reject the null hypothesis
c) Z = -2.2247, p-value = 0.013, reject the null hypothesis
d) T = -2.2247, p-value = 0.0444, fail to reject the null hypothesis
e) Z = -2.2247, p-value = 0.013, fail to reject the null hypothesis

17) To study the effectiveness of wall insulation in saving energy for home heating, the energy consumption (in MWh) for 10 houses was recorded for two winters; the first winter was before insulation and the second winter was after insulation.

House	1	2	3	4	5	6	7	8	9	10
Before Insulation	12.1	11.0	14.1	13.8	15.5	12.2	12.8	9.9	10.8	12.7
After Insulation	12.0	10.6	13.4	11.2	15.3	13.6	12.6	8.8	9.6	12.4

We want to test if the insulation has reduced the average energy consumption. Give the calculated test statistic for this hypothesis test.
a) 0.54   b) 0.2585   c) 1.681   d) 0.662   e) 0.0635

18) The following is a scatterplot. What would be the best value of the correlation coefficient between x and y?
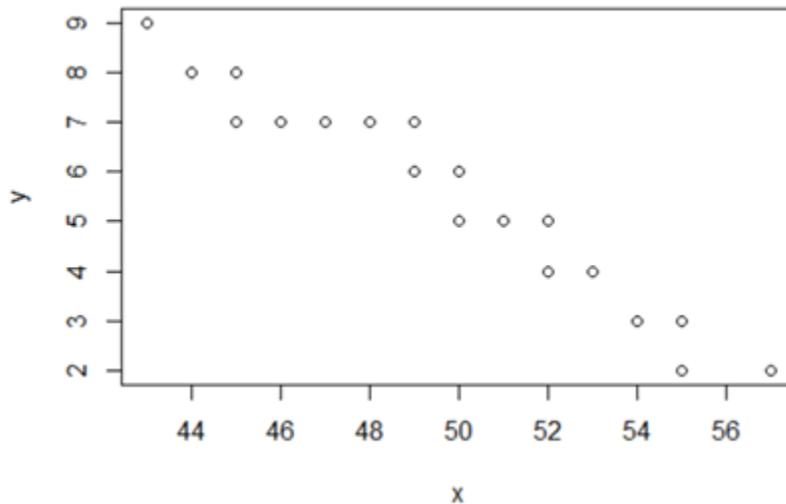
a) -0.965   b) -1   c) -0.55   d) 0.965   e) 0.55

19) A simple linear regression equation was calculated to be: $y = 30 - 0.5x$. Given that $s_x = 3.7$ and $s_y = 2$, determine the correlation coefficient, r.
a) 0.925   b) -0.925   c) -0.5   d) 0.5   e) -0.27

20) The following data set has five college students, their GPA and SAT scores. We want to determine the GPA by their SAT scores.

Student	1	2	3	4	5
GPA	3.1	3.4	3.7	3.2	3.7
SAT	650	680	770	710	730

Suppose we want to predict a student's GPA who has an SAT score of 788. The best predicted value is:
a) 3.72   b) 3.70   c) 3.75   d) 3.82   e) 3.79

21) Assume that IQ scores have a Normal distribution with mean 110 and standard deviation 15. What is the interquartile range of the IQ scores?
a) 120.12   b) 99.88   c) 20.23   d) 15   e) Do not have enough information to determine.

22) The lifetime of car brakes varies according to a Normal distribution with mean $\mu =$ 50,000 miles and standard deviation $\sigma =$ 3000 miles. Suppose a sample of nine car brakes are tested, what is the probability that the sample mean will be less than 47,000 miles?
a) 0.8413   b) 0.9987   c) 0.0013   d) 0.1587   e) The sample size is not large enough to determine.

23) A grocery store has 15 check out lanes. During a busy hour the probability that any given lane is occupied (has at least one customer) is 0.8. Assume that the lanes are occupied or not occupied independently of each other. What is the probability that a customer will find at least one lane unoccupied?
a) 0.0352   b) 0.1352   c) 0.8329   d) 0.9648   e) 0.1671

24) An office supply warehouse receives an order for three laptops. The warehouse has 25 laptops in stock, two of which are defective. The order is filled by randomly drawing from the laptops in stock. What is the probability that at least one laptop will be defective?
a) 0.27   b) 0.22   c) 0.23   d) 0.77   e) 0.88

25) The five-number summary of the distribution of scores on a Math exam was

$$18, 39, 62, 76, 100$$

A total of 416 took the exam. About how many students had a score above 39?
a) 312   b) 416   c) 104   d) 208   e) 400

26) A store runs a prize game by giving each customer a ticket that may win a prize when a box is scratched. Printed on the ticket are the following probabilities for a customer who shops once a week:

Amount Won	Probability
$0	0.74
$10	0.20
$50	0.05
$500	0.01

Suppose the store decided to give everyone 10 times plus 5 dollars more than what is on the ticket. What is the expected value of a customer's winnings in this game?
a) $1405   b) $100   c) $255   d) $9.50   e) $140

27) Let A and B be two events such that $P(A) = 0.8$, $P(B) = 0.6$, and $P(A \cap B) = 0.4$. Which statement is correct?
a) Events A and B are the entire sample space.
b) Events A and B are independent.
c) Events A and B are both mutually exclusive and independent.
d) Events A and B are mutually exclusive (disjoint).
e) None of these statements are correct.

28) A machine fills cans of cola which should have a mean volume of 12 ounces. We need to determine if the machine is underfilling the cans. A hypothesis test is conducted, to test $H_0 : \mu \geq 12$ against $H_\alpha : \mu < 12$. A test of a random sample of 40 cans of cola filled from this machine determined that the mean amount is statistically less than 12 ounces. However, after checking all of the cans, the mean amount was at least 12 ounces. This represents
a) Type 3 error   b) Machine error   c) Sampling error
d) Type 1 error   e) Type 2 error

29) Hospital records show that 12% of all patients are admitted for heart disease, 20% are admitted for cancer (oncology) treatment, and 8% receive both coronary and oncology care. What is the probability that a randomly selected patient is admitted for coronary care, oncology or both? (Note that heart disease is a coronary care issue.)
a) 0.16   b) 0.32   c) 0.68   d) 0.24   e) 0.40

30) A large-scale study conducted over a one-year period has shown that break-ins at home occur about 7% of the time in the population. The study also shows home security alarms went off 3% of the time when no one was breaking into the home. The security alarm failed to go off 4% of the time when someone was really breaking into the home. If an alarm is going off, what is the probability that the house was broken into?
a) 0.2066   b) 0.9600   c) 0.7066   d) 0.0672   e) 0.9328

31) Suppose two events A and B are independent with $P(A \cup B) = 0.61$ and $P(A) = 0.25$. Determine $P(B)$.
a) 0.36   b) 0.48   c) 0.41   d) 0.25   e) Cannot be determined.

32) Let A and B be two events such that $P(A) = 0.2$, $P(B) = 0.3$, and $P(A \cup B) = 0.4$. Which statement is correct?
a) Events A and B are both mutually exclusive and independent.
b) Events A and B are the entire sample space.
c) None of the other statements are correct.
d) Events A and B are independent.
e) Events A and B are mutually exclusive (disjoint).

33) What can be said about the relationship between the mean and the median for the data represented in the histogram below?
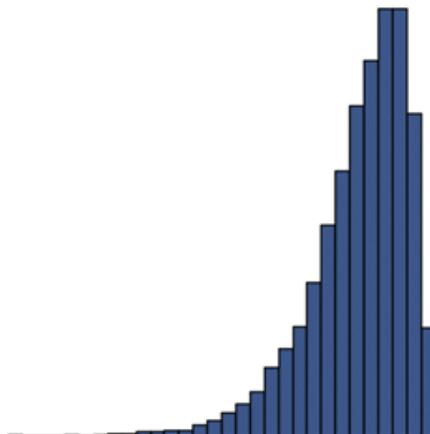
a) The mean is greater than the median.
b) The mean is less than the median.
c) The mean and median are equal
d) The mean is 2 times greater than the median.
e) We cannot determine because we do not have the actual numbers.

34) The following is a probability distribution of grades in a class where A = 4, B = 3, C = 2, D = 1 and F = 0.

Grade	0	1	2	3	4
Probability	0.15	0.15	0.3	0.3	0.1

Given that a student passed with at least a C, what is the expected grade of that student?
a) 2.50   b) 2.71   c) 2.00   d) 2.05   e) 1.90

35) 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 8 milligrams per deciliter of blood (mg/dl). This means that
a) We are 95% confident that the true population mean is within 8 mg/dl of the calculated sample mean.
b) We could be certain that the study result is within 8 mg/dl of the truth about the population if the conditions for inferences were satisfied.
c) We can be certain that the study results is within 8 mg/dl of the truth about the population.
d) 95% percent of the population has changed their HDL after eating raw garlic six days a week for six months.
e) There is a 95% probability that the true population mean is within the calculated interval.

36) The probability that the test statistics will be at least as extreme as what is actually observed assuming the null hypothesis, H_0, is true is the
a) power of the test
b) p-value
c) critical value
d) probability of type 1 error
e) probability of type 2 error

37) Suppose that a fair, 6-sided die is rolled. Let X indicate the event that an even number is rolled (in other words, Let $X = 1$ if an even number is rolled and Let $X = 0$ otherwise). Let Y indicate the event that 3, 4, or 5 is rolled (in other words, Let $Y = 1$ if 3, 4, or 5 is rolled and Let $Y = 0$ otherwise). Find $P(X = 0, Y = 1)$.
a) 5/6   b) 1/3   c) 1/6   d) 1/2   e) 2/3

38) A six-sided die is thrown 50 times. The numbers of occurrences of each face are shown below.

Face	1	2	3	4	5	6
Count	5	5	12	12	9	7

Can you conclude that the die is not fair? Determine the type of test should be used in this situation and the test statistic.
a) Chi squared Test for Independence, $\chi^2 = 3.4727$
b) One sample z-test for proportions, z = 1.265
c) Chi squared Goodness of Fit, $\chi^2 = 6.16$.
d) Two sample z-test for proportions, z = 2.656
e) Chi squared Goodness of Fit, $\chi^2 = 2.551$

39) Suppose a random variable X has a mean of 300 and variance of 169. We create a new variable by $Y = 5X + 30$. What is the standard deviation of the new variable, Y?
a) 875   b) 95   c) 4225   d) 65   e) 845

40) The probability that a student correctly answers on the first try (the event A) is $P(A) = 0.3$. If the student answers incorrectly on the first try, the student is allowed a second try to correctly answer the question (the event B). The probability that the student answers correctly on the second try given that he answered incorrectly on the first try is 0.4. Find the probability that the student correctly answers the question on the first or second try.

a) 0.51   b) 0.58   c) 0.54   d) 0.12   e) 0.70