Statistics 151 Practice Midterm 1 – Multiple Choice (50 minutes)

Instructions:

1. This is a closed book exam.
2. You may use the STAT 151 formula sheets and tables provided and a NON-PROGRAMMABLE calculator only.
3. You have 50 minutes to complete the exam.
4. The exam consists of 25 multiple choice questions worth 1 mark each.
5. Topics covered: Chapters 1 – 8 in Agresti/Franklin 2.

---

1. The fraction of the variation in the values of a response *y* that is explained by the least-squares regression of *y* on *x* is
   A) the correlation coefficient.
   B) the slope of the least-squares regression line.
   C) the square of the correlation coefficient.
   D) the intercept of the least-squares regression line.

Use the following to answer the questions 2 - 3.
A sample was taken of the verbal GRE scores of 20 applicants to graduate school at a large midwestern university. Below are the scores. For convenience, the data are ordered.

280	310	340	350	370	410	420	420	420	470
490	510	520	520	600	610	670	720	750	770

2. The first quartile for the applicant scores is
   A) 340.
   B) 390.
   C) 480.
   D) 605.

3. If 25 points were added to each score, then the standard deviation of the new scores would
   A) be increased by 5.
   B) be increased by 25.
   C) be increased by 625.
   D) remain unchanged.

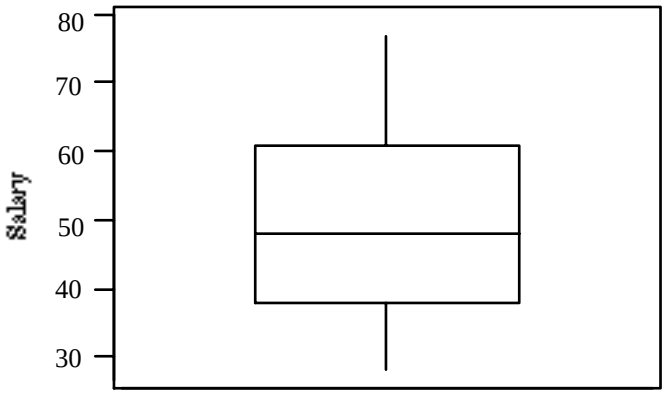
4. A newspaper conducted a statewide survey concerning a proposal to raise taxes in order to prevent budget cuts to education. The newspaper took a random sample (assume it is an SRS) of 1200 registered voters and found that 580 would vote to raise taxes. Let *p* represent the proportion of registered voters in the state that would vote to raise taxes. A 90% confidence interval for *p* is
   A) 0.483 ± 0.014.
   B) 0.483 ± 0.024.
   C) 0.483 ± 0.028.
   D) 0.483 ± 0.249.

5. In a certain town 60% of the households own mutual funds, 40% own individual stocks, and 20% own both mutual funds and individual stocks. The proportion of households that own neither mutual funds nor individual stocks is
   A) 20%.
   B) 30%.
   C) 40%.
   D) 50%.

6. A researcher reports that, on average, the participants in his study lost 10.4 lbs. after two months on his new diet. A friend of yours comments that she tried the diet for two months and lost no weight, so clearly the report must be a fraud. Which of the following statements is correct?
   A) Your friend must not have followed the diet correctly because she did not lose weight.
   B) Because your friend did not lose weight, the report must not be correct.
   C) The report gives only the average. This does not imply that all participants in the study lost 10.4 lbs. or even that all participants lost weight. Your friend's experience does not necessarily contradict the study results.
   D) In order for the study to be correct, we must now add your friend's results to those of the study and re-compute the new average.

7. In a large population of adults, the mean IQ is 112 with a standard deviation of 20. Suppose 200 adults are randomly selected for a market research campaign. The probability that the sample mean IQ is greater than 110 is
   A) 0.079.
   B) 0.421.
   C) 0.921.
   D) 0.579.

Use the following to answer the questions 8 - 9.
A sample was taken of the salaries of 20 employees of a large company. The following is a boxplot of the salaries (in thousands of dollars) for this year.



8. Based on this boxplot, the five-number summary is
   A) 28, 39, 48, 60.5, 77.
   B) 28, 41, 48, 58, 77.
   C) 28, 39, 51, 58, 77.
   D) 28, 41, 51, 60.5, 77.

9. Based on this boxplot, which of the following statements is true?
   A) The salary distribution is fairly symmetric.
   B) About 10 employees make over $50,000.
   C) Nobody makes over $80,000.
   D) All of the above.

10. I collect a random sample of size n from a population and from the data collected compute a 95% confidence interval for the mean of the population. Which of the following would produce a new confidence interval with larger width (larger margin of error) based on these same data?
    A) Use a larger confidence level.
    B) Use a smaller confidence level.
    C) Use the same confidence level, but compute the interval n times. Approximately 5% of these intervals will be larger.
    D) Nothing can guarantee absolutely that you will get a larger interval. One can only say the chance of obtaining a larger interval is 0.05.

11. The heights (in inches) of males in the United States are believed to be normally distributed with mean μ. The average height of a random sample of 50 American adult males is found to be $\overline{x}$ = 69.72 inches, and the standard deviation of the 50 heights is found to be s = 4.23. The standard error of $\overline{x}$ is
    A) 0.084.
    B) 0.357.
    C) 0.598.
    D) 0.731.

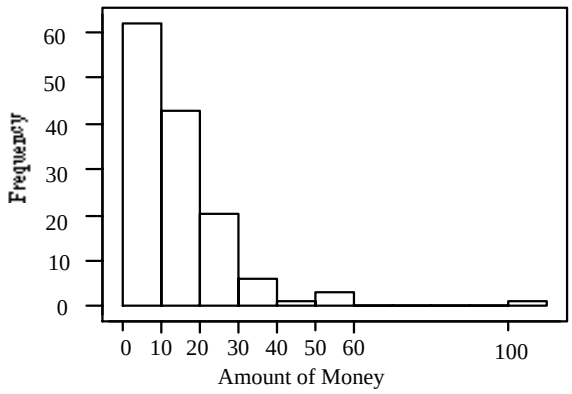
12. A study found a correlation of $r = -0.61$ between the gender of a worker and his or her income. You may correctly conclude
    A) women earn more than men on the average.
    B) women earn less than men on the average.
    C) an arithmetic mistake was made. Correlation must be positive.
    D) this is incorrect because r makes no sense here.

13. Event A occurs with probability 0.3 and event B occurs with probability 0.4. If A and B are independent, we may conclude
    A) $P(A \text{ and } B) = 0.12$.
    B) $P(A|B) = 0.3$.
    C) $P(B|A) = 0.4$.
    D) all of the above.

14. Birthweights at a local hospital have a normal distribution with a mean of 110 oz. and a standard deviation of 15 oz. The proportion of infants with birthweights above 125 oz. is
    A) 0.500.
    B) 0.159.
    C) 0.341.
    D) 0.841.

15. An event A will occur with probability 0.5. An event B will occur with probability 0.6. The probability that both A and B will occur is 0.1. We may conclude
    A) events A and B are independent.
    B) events A and B are disjoint.
    C) either A or B always occurs.
    D) none of the above.

Use the following to answer the question 16.
In a Business Statistics class with 136 students, the professor records how much money each student has in his or her possession during the first class of the semester. The histogram below is of the data collected.



.

16. From the histogram, which of the following is true?
    A) The mean is much larger than the median.
    B) The mean is much smaller than the median.
    C) The mean and median are approximately equal.
    D) It is impossible to compare the mean and median for these data.

17. A roulette wheel has 38 slots in which the ball can land. Two of the slots are green, 18 are red, and 18 are black. The ball is equally likely to land in any slot. The roulette wheel is going to be spun twice, and the outcomes of the two spins are independent. The probability that it lands once on red and once on black is
    A) 0.2244.
    B) 0.2770.
    C) 0.4488.
    D) 0.9474.

18. The time to complete a standardized exam is approximately normal with a mean of 70 minutes and a standard deviation of 10 minutes. How much time should be given to complete the exam so that 80% of the students will complete the exam in the time given?
    A) 84 minutes
    B) 78.4 minutes
    C) 92.8 minutes
    D) 79.8 minutes

19. Eighty rats whose mothers were exposed to high levels of tobacco smoke during pregnancy were put through a simple maze. The maze required the rats to make a choice between going left or going right at the outset. Sixty of the rats went right when running the maze for the first time. Assume that the 80 rats can be considered an SRS from the population all rats born to mothers exposed to high levels of tobacco smoke during pregnancy (note that this assumption may or may not be reasonable, but researchers often assume lab rats are representative of such larger populations because lab rats are often bred to have very uniform characteristics). The standard error for the proportion $\hat{p}$ of those who went right the first time when running the maze is
    A) 0.0559.
    B) 0.0023.
    C) 0.0484.
    D) 0.0548.

20. A level 0.90 confidence interval is
A) any interval with margin of error ± 0.90.
B) an interval computed from sample data by a method that has probability 0.90 of producing an interval containing the true value of the parameter of interest.
C) an interval with margin of error ± 0.90, which is also correct 90% of the time.
D) an interval computed from sample data by a method that guarantees that the probability the interval computed contains the parameter of interest is 0.90.

21. The density curve for a continuous random variable *X* has which of the following properties?
A) The probability of any event is the area under the density curve and above the values of *X* that make up the event.
B) The total area under the density curve for *X* must be exactly 1.
C) The probability of any event of the form *X* = constant is 0.
D) All of the above.

22. A radio talk show host with a large audience is interested in the proportion $\hat{p}$ of adults in his listening area that think the drinking age should be lowered to 18. To find this out he poses the following question to his listeners: "Do you think that the drinking age should be reduced to 18 in light of the fact that 18-year-olds are eligible for military service?" He asks listeners to phone in and vote "yes" if they agree the drinking age should be lowered and "no" if they do not. You are told that the proportion $\hat{p}$ of those who phoned in and answered yes is $\hat{p}$ = 0.70 and the standard error $\mathrm{SE}_{\hat{p}}$ of the proportion is 0.0459. The number of people who phoned in
A) is 50.
B) is 100.
C) is 200.
D) cannot be determined from the information given.

23. A SRS of 20 third grade children is selected in Chicago and each is given a test to measure his or her reading ability. In the sample, the mean score is 64 points and the standard deviation is 12 points. We are interested in a 95% confidence interval for the population mean score. The margin of error associated with the confidence interval is
A) 2.68 points.
B) 4.64 points.
C) 5.62 points.
D) 6.84 points.

24. Which of the following is correct?
A) The correlation *r* is the slope of the least-squares regression line.
B) The square of the correlation is the slope of the least-squares regression line.
C) The square of the correlation is the proportion of the data lying on the least-squares regression line.
D) The mean of the residuals from least-squares regression is 0.

25. Can one predict a student's score on the midterm exam in a statistics course from the number of hours the student spent studying for the exam? To explore this, the teacher of the course asks students how many hours they spent studying for the exam and then makes a scatterplot of the time students spent studying and their scores on the exam. In making the scatterplot, the teacher should
A) plot the score on the exam on the horizontal axis.
B) plot time spent studying for the exam on the horizontal axis.
C) first determine if the scores on the exam approximately follow a normal distribution.
D) use a plotting scale that makes the overall trend roughly linear.