

AP Statistics Midterm Review

Multiple Choice: *Identify the choice that best completes the statement or answers the question.*

- _____ 1. Each of the 30 major league baseball teams carries a 40-person roster. A sample of 60 players (5 percent of all 1,200 players) is to be randomly selected to undergo drug tests. To do this, each team is instructed to put their 40 names in a hat and randomly draw two names. Will this method result in a simple random sample of the 1,200 baseball players?
- a. Yes, because each player has the same chance of being selected.
 - b. Yes, because each team is equally represented.
 - c. Yes, because this is an example of stratified sampling, which is a special case of simple random sampling
 - d. No, because the teams are not chosen randomly
 - e. No, because not each group of 60 players has the same chance of being selected.
- _____ 2. Advantage(s) to using surveys as opposed to experiments is (are) that
- I. Surveys are generally cheaper to conduct.
 - II. It is generally easier to conclude cause and effect from surveys.
 - III. Surveys are generally not subject to bias
- a. I only
 - b. II only
 - c. III only
 - d. I and II
 - e. II and III
- _____ 3. A company wishes to survey what people think about a new product it plans to market. They decide to randomly sample from their customer database as this includes phone numbers and addresses. This procedure is an example of which type of sampling?
- a. Cluster
 - b. Convenience
 - c. Simple Random
 - d. Stratified
 - e. Systematic
- _____ 4. Two studies are run to compare the experiences of low-income families receiving food stamps to those receiving cash subsidies. The first study interviews 50 families who have been in each government program for at least 2 years, while the second randomly assigns 50 families to each program and interviews them after 2 years. Which of the following is a true statement?
- a. Both studies are observational studies because of the time period involved.
 - b. Both studies are observational studies because there are no control groups.
 - c. The first study is an observational study; the second is an experiment.
 - d. The first study is an experiment; the second is an observational study.
 - e. Both studies are experiments, because in each, families are receiving treatments (food stamps or cash).
- _____ 5. In a study of Parkinson's disease, 100 volunteers had incisions made through their skulls. The patients were randomly sorted into two groups, one of which had a new drug inserted into the brain. In the other group, the skulls were closed with no treatment given. The patients did not know who received the drug. In the weeks to follow all 100 volunteers showed similar improvement in physical condition. What is this an example of?
- a. The effect of a treatment unit.
 - b. The placebo effect.
 - c. The control group effect
 - d. Sampling error
 - e. Voluntary response bias

- _____ 6. Fifty migraine patients are randomly selected from hospital records. Half the patients are told to drink ice water and sit in the dark when they next experience a migraine; the remaining patients are told to use neither of these possible remedies. Participants then report back as to relief, if any. Serious faults of this experimental design include which of the following?
- I. Lack of randomization.
 - II. Probable confounding variables.
 - III. Lack of blinding
- a. I only
 - b. II only
 - c. III only
 - d. I and II
 - e. II and III
- _____ 7. Which of the following is most useful in establishing cause-and-effect relationships?
- a. A complete census
 - b. A least squares regression line showing high correlation
 - c. A simple random sample (SRS)
 - d. A well-designed, well-conducted survey incorporating chance to ensure a representative sample
 - e. A controlled experiment.
- _____ 8. A sales representative wishes to survey her client base of 47 companies. She has 47 business cards, all of identical size, from her contacts in the companies, and decides to drop them all in a small box, shake them up, and reach in to pick 5 cards for her sample. This procedure is an example of which type of sampling?
- a. Cluster
 - b. Convenience
 - c. Simple random
 - d. Stratified
 - e. Systematic
- _____ 9. A newspaper advice columnist asks her readers if they would have married their current spouse if they had to do it over again. Of the 25,000 or so responses, 80 percent said no. What does this show?
- a. The survey is meaningless because of voluntary response bias.
 - b. No meaningful conclusion is possible without knowing something more about the characteristics of her readers.
 - c. The survey would have been more meaningful if she had picked a random sample of the 25,000 readers who responded.
 - d. The survey would have been more meaningful if she had used a control group.
 - e. This was a legitimate sample, randomly drawn from her readers, and of sufficient size to allow the conclusion that most of her readers who are married would have second thoughts about marrying their current spouse.
- _____ 10. Which of the following is most important in minimizing the placebo effect?
- a. Replication and randomization
 - b. Replication and blinding
 - c. Randomization and blinding
 - d. Randomization and a control
 - e. Blinding and a control
- _____ 11. A bank wishes to survey its customers. The decision is made to randomly pick ten customers who just have checking accounts, ten customers who just have savings accounts, and ten customers who have both checking and savings accounts. This procedure is an example of which type of sampling?
- a. Cluster
 - b. Convenience
 - c. Simple random
 - d. Stratified
 - e. Systematic

c. Simple random

- _____ 12. Which of the following are true statements?
- I. If bias is present in a sampling procedure, it can be overcome by dramatically increasing the sample size.
 - II. There is no such thing as a 'bad sample'.
 - III. Sampling techniques that use probability techniques effectively eliminate bias.
- a. I only
 - b. II only
 - c. III only
 - d. none of these statements are true
 - e. none of these gives the complete set of true responses.
- _____ 13. Which of the following are true statements about blocking?
- I. Blocking is to experiment design as stratification is to sampling design.
 - II. By controlling certain variables, blocking can make conclusions more specific.
 - III. The paired comparison design is a special case of blocking.
- a. I and II
 - b. I and III
 - c. II and III
 - d. I, II, and III
 - e. None of these gives the complete set of true responses.
- _____ 14. What is *bias* in conducting surveys?
- a. An example of sampling error
 - b. Lack of a control group
 - c. Confounding variables
 - d. Difficulty in concluding cause and effect
 - e. A tendency to favor the selection of certain members of a population.
- _____ 15. A television network conducts a weekly survey to determine the proportion of viewers who watch various programs. For the coming year, they decide to double the sample size. The main benefit of this is to
- a. Reduce undercoverage bias
 - b. Reduce nonresponse bias
 - c. Eliminate sampling error
 - d. Decrease population variability
 - e. Decrease the standard deviation of the sampling distribution.
- _____ 16. Which of the following are true statements?
- I. Voluntary response samples often underrepresent people with strong opinions.
 - II. Convenience samples often lead to undercoverage bias.
 - III. Questionnaires with nonneutral wording are likely to have response bias.
- a. I and II
 - b. I and III
 - c. II and III
 - d. I, II, and III
 - e. None of these gives the complete set of true responses.
- _____ 17. A critical difference between experiments and observational studies is
- a. An experiment often suggests a causal relationship, whereas an observational study only suggests an association.
 - b. Observational studies make use of randomization, whereas experiments do not.
 - c. Experiments are free to choose subjects from an entire population, whereas an observational study only considers a random sample.
 - d. Tests of significance can be used on data collected from experiments but not on data from observational studies.
 - e. Experiments are free to choose subjects from an entire population, whereas an observational study only considers a random sample.

- c. Experiments are generally more cost and time effective than observational studies.

- ___ 18. Which of the following are true statements?
- I. In well-designed observational studies, responses are systematically influenced during the collection of data.
 - II. In well-designed experiments, the treatments result in responses that are as similar as possible.
 - III. A well-designed experiment always has a single treatment but may test that treatment at different levels.
- a. I only
 - b. II only
 - c. III only
 - d. II and III
 - e. None of these is true
- ___ 19. You wish to survey people who have brought in their cars for service during the past month. You decide to pick a random sample of gas stations in the city and then survey all customers from those stations who had work done during the past month. This procedure is an example of which type of sampling?
- a. Cluster
 - b. Convenience
 - c. Simple random
 - d. Stratified
 - e. Systematic
- ___ 20. Suppose you wish to compare the average height of math/science teachers to the average height of English/social studies teachers in your high school. Which is the most appropriate technique for gathering the needed data?
- a. Census
 - b. Sample survey
 - c. Experiment
 - d. Observational study
 - e. None of these methods is appropriate.
- ___ 21. Which of the following are true statements?
- I. Based on careful use of control groups, experiments can often indicate cause-and-effect relationships.
 - II. Observational studies may suggest relationships, but it would be very difficult to conclude cause and effect because of the lack of control over lurking variables.
 - III. A complete census is the only way to absolutely establish a cause-and-effect relationship.
- a. I and II
 - b. I and III
 - c. II and III
 - d. I, II, and III
 - e. None of these gives the complete set of true responses.
- ___ 22. Which of the following best explains why we try to guard against confounding when designing experiments?
- a. Confounding can lead to bias
 - b. Confounding can conflict with randomization
 - c. Confounding can lead to uncertainty as to which variable is causing an effect.
 - d. Confounding can make it more difficult to separate subjects into treatment and control groups.
 - e. Confounding can negate the benefits of blinding.
- ___ 23. Given that the median is 270 and the interquartile range is 20, which of the following statements are true?
- I. Fifty percent of the data are greater than or equal to 270.

- II. Fifty percent of the data are between 260 and 280.
- III. Seventy-five percent of the data are less than or equal to 280.

- a. I only
- b. II only
- c. III only
- d. I and II
- e. I, II, and III

_____ 24. Suppose the correlation is negative. Given two points from the scatterplot, which of the following is possible?

- I. The first point has a larger x-value and a smaller y-value than the second point.
- II. The first point has a larger x-value and a larger y-value than the second point.
- III. The first point has a smaller x-value and a larger y-value than the second point.

- a. I only
- b. II only
- c. III only
- d. I and III
- e. I, II, and III

_____ 25. Which of the following are true statements?

- I. Stemplots are useful both for quantitative and categorical data sets.
- II. Stemplots are equally useful for small and very large data sets.
- III. Stemplots can show symmetry, gaps, clusters, and outliers.

- a. I only
- b. II only
- c. III only
- d. I and II
- e. I and III

_____ 26. If ten executives have salaries of \$80,000, six have salaries of \$75,000, and three have salaries of \$70,000, what is the median salary?

- a. \$75,000
- b. \$76,842
- c. \$77,500
- d. \$80,000
- e. None of the above

_____ 27. Data on ages (in years) and prices (in \$100) for ten cars of a specific model result in the regression line: Price = 250 - 30(Age). Given that 64 percent of the variation in price is explainable by variation in age, what is the value of the correlation coefficient r ?

- a. -0.64
- b. -0.80
- c. 0.64
- d. 0.80
- e. There is insufficient information to answer this question.

_____ 28. A real estate agent, working entirely on commission, weekly makes an average of \$850 with a standard deviation of \$260 selling property in the city and an average of \$1,340 with a standard deviation of \$390 selling property in the suburbs. Assuming independence of what she sells in the two locations, what are the mean and standard deviation of her total weekly sales?

- a. Mean = \$1,095
Standard deviation = \$325
- b. Mean = \$1,095
Standard deviation = \$469
- c. Mean = \$2,190
Standard deviation = \$325
- d. Mean = \$2,190
Standard deviation = \$469
- e. Mean = \$2,190
Standard deviation = \$650

- _____ 35. Suppose the correlation between two variables is $r = 0.28$. What will the new correlation be if 0.17 is added to all values of the x-variable, every value of the y-variable is doubled, and the two variables are interchanged?
- a. 0.28
b. 0.45
c. 0.56
d. 0.90
e. -0.28
- _____ 36. Which of the following statements are true?
- I. Two students working with the same set of data may come up with histograms that look different.
II. Displaying outliers is less problematic when using histograms than when using stemplots.
III. Histograms are more widely used than stemplots or dotplots because histograms display the values of individual observations.
- a. I only
b. II only
c. III only
d. I and II
e. II and III
- _____ 37. Using the most commonly accepted definition of outliers, a set has five outliers. If every value of the set is increased by 20 percent, how many outliers will there now be?
- a. Fewer than five
b. Five
c. Six
d. More than six
e. It is impossible to determine without further information.
- _____ 38. Which of the following statements about the correlation r are true?
- I. When $r = 0$, there is no relationship between the variables.
II. When $r = .2$, 20 percent of the variables are closely related.
III. When $r = 1$, there is a perfect cause-and-effect relationship between the variables.
- a. I only
b. II only
c. III only
d. I, II, and III
e. All the statements are false
- _____ 39. Suppose a study finds that the correlation coefficient relating job satisfaction to salary is $r = +1$. Which of the following are proper conclusions?
- I. High salary causes high job satisfaction.
II. Low salary causes low job satisfaction.
III. There is a very strong association between salary and job satisfaction.
- a. I only
b. II only
c. III only
d. I and II
e. I, II, and III
- _____ 40. When a set of data has suspect outliers, which of the following are preferred measures of central tendency and of variability?
- a. Mean and standard deviation
b. Mean and variance
c. Mean and range
d. Median and range
e. Median and interquartile range

- _____ 41. Given two independent random variables, X with mean 12.3 and standard deviation 0.5, and Y with mean 9.1 and standard deviation 0.3, which of the following is a true statement?
- The mean of X - Y is 21.4
 - The median of X - Y is 3.2
 - The range of X - Y is 21.4
 - The standard deviation of X - Y is 0.8
 - The variance of X - Y is 0.34
- _____ 42. A data set includes two outliers, one at each end. If both these outliers are removed, which of the following is a possible result?
- Both the mean and standard deviation remain unchanged
 - Both the median and standard deviation remain unchanged
 - Both the standard deviation and variance remain unchanged
 - Both the mean and median remain unchanged.
 - Both the mean and standard deviation increase.
- _____ 43. A histogram of the educational level (in number of years of schooling) of the adult populations of the United States would have which of the following characteristics?
- Symmetry
 - Clusters
 - Skewness to the left
- II only
 - I and II
 - I and III
 - II and III
 - I, II, and III
- _____ 44. A doctor wishes to compare the resting heart rates of his younger patients (younger than 30 years old) versus his older patients (older than 30 years old). Which of the following graphical displays is *innapropriate*?
- Back-to-back stemplot
 - Parallel boxplots
 - Side by side histograms
 - Scatterplot
 - All of these displays are appropriate.
- _____ 45. Suppose X and Y are random variables with $\mu_x = 38$, $\sigma_x = 12$, $\mu_y = 35$, $\sigma_y = 9$. Given the X and Y are independent, what is the standard deviation of the random variable X - Y?
- $\sqrt{3}$
 - $\sqrt{21}$
 - 3
 - 15
 - 21
- _____ 46. Which of the following statements are true?
- All symmetric histograms have single peaks.
 - All symmetric bell-shaped curves are normal.
 - All normal curves are bell-shaped and symmetric.
- I only
 - II only
 - III only
 - I and II
 - None of these gives the complete set of true responses.
- _____ 47. Suppose the scores on an exam have a mean of 75 with a standard deviation of 8. If one student has a test result with a z-score of -1.5, and a second student has a result with a z-score of 2.0, how many points higher was the second student than that of the first?
- 3.5
 - 4
 - 16
 - 28

c. 12

- ___ 48. Which of the following statements about influential points are true?
- I. Looking at a residual plot is an excellent way of picking out influential points.
 - II. Removal of an influential point sharply affects the regression line.
 - III. Determining a regression model with and without a point is an excellent way of picking out influential points.
- a. I and II
 - b. I and III
 - c. II and III
 - d. I, II, and III
 - e. none of these gives the complete set of true responses.
- ___ 49. When there are multiple gaps and clusters, which of the following is the best choice to give an overall picture of a distribution?
- a. Mean and standard deviation
 - b. Median and interquartile range
 - c. Boxplot with its five-number summary
 - d. Stemplot or histogram
 - e. None of these are really helpful in showing gaps and clusters.
- ___ 50. If quartiles $Q_1 = 50$ and $Q_3 = 70$, which of the following must be true?
- I. The median is 60
 - II. The mean is between 50 and 70.
 - III. The standard deviation is at most 20.
- a. I only
 - b. II only
 - c. III only
 - d. All are true
 - e. None may be true.
- ___ 51. Many professional schools require applicants to take a standardized test. Suppose that 1000 students take such a test. Several weeks after the test, Joe receives his score report: he got a 63, which placed him at the 73rd percentile. This means that:
- a. Joe's score was below the median
 - b. Joe did worse than about 63% of all test takers.
 - c. Joe did worse than about 73% of all test takers.
 - d. Joe did better than about 63% of all test takers.
 - e. Joe did better than about 73% of all test takers.
- ___ 52. Which of the following is NOT correct about about a standard Normal distribution?
- a. The proportion of scores that satisfy $0 < z < 1.5$ is 0.4332
 - b. The proportion of scores that satisfy $z < -1.0$ is 0.1587
 - c. The proportion of scores that satisfy $z > 2.0$ is 0.0228
 - d. The proportion of scores that satisfy $z < 1.5$ is 0.9332
 - e. The proportion of scores that satisfy $z > -3.0$ is 0.9938
- ___ 53. If the heights of Canadian men follow a Normal Distribution, and 99.7% have heights between 5'0" and 7'0" what is your estimate of the standard deviation fo the height of Canadian men?
- a. 1"
 - b. 3"
 - c. 4"
 - d. 6"
 - e. 12"

Until the scale was changed in 1995, SAT scores were based on a scale set many years ago. For Math scores, the mean under the old scale in the 1990s was 470 and the standard deviation was 110. In 2009, the mean was 515 and the standard deviation was 116.

c. .3085

- ___ 62. If $P(A) = .25$ and $P(B) = .34$, what is $P(A \cup B)$ if A and B are independent?
- a. .085
b. .505
c. .590
d. .675
e. There is insufficient information to answer this question.
- ___ 63. A person has a 10 percent chance of winning the daily office lottery. What is the probability she first wins on the fourth day?
- a. $\binom{4}{1} (.10)^3 (.90)$
b. $\binom{4}{3} (.10)(.90)^3$
c. $(.10)^3 (.90)$
d. $(.10)(.90)^3$
e. None of these gives the correct probability.
- ___ 64. Which of the following are true statements?
- I. The area under a normal curve is always equal to 1, no matter what the mean and standard deviation are.
II. The smaller the standard deviation of a normal curve, the higher and narrower the graph.
III. Normal curves with different means are centered around different numbers.
- a. I and II
b. I and III
c. II and III
d. I, II, and III
e. None of these gives the complete set of true responses.
- ___ 65. A piece of clothing takes an average of 38 minutes to move through an assembly line. If the standard deviation is 4 minutes, and the distribution is normal, what is the probability that a piece of clothing will take over 45 minutes?
- a. .040
b. .080
c. .175
d. .227
e. .460
- ___ 66. There are five outcomes to an experiment and a student calculates the respective probabilities of the outcomes to be .34, .50, .42, 0, and -.26. The proper conclusion is that
- a. The sum of the individual probabilities is 1.
b. One of the outcomes will never occur
c. One of the outcomes will occur 50 percent of the time.
d. All of these are true
e. The student made an error.
- ___ 67. Box A has four \$10 bills and single \$100 bill, box B has 400 \$10 bills and 100 \$100 bills, and box C has 28 \$1 bills. You can have all of box C or blindly pick one bill out of either box A or box B. Which choice offers the greatest expected winning?
- a. Box A
b. Box B
c. Box C
d. Either A or B, but not C
e. All offer the same expected winning.

- ___ 68. Given that 49.0 percent of the U.S. population is male, and 12.1 percent of the population are over 65 years of age, can we conclude that $(.490)(.121) = 5.93$ percent of the population are men older than 65?
- Yes, by the multiplication rule
 - Yes, by conditional probabilities
 - Yes, by the Law of Large Numbers
 - No, because the events are not independent
 - No, because the events are not mutually exclusive.
- ___ 69. According to one poll, only 8 percent of the public say the “trust Congress.” In a simple random sample of ten people, what is the probability that at least one person “trusts Congress”?
- .188
 - .378
 - .434
 - .566
 - .622
- ___ 70. Suppose you toss a fair coin ten times and it comes up heads every time. Which of the following is a true statement?
- By the Law of Large Numbers, the next toss is more likely to be tails than another heads.
 - By the properties of conditional probability, the next toss is more likely to be heads given that ten tosses in a row have been heads.
 - Coins actually do have memories, and thus what comes up on the next toss is influenced by the past tosses.
 - The Law of Large Numbers tells how many tosses will be necessary before the percentages of heads and tails are again in balance.
 - None of these are true statements.
- ___ 71. A mortgage company advertises that 85 percent of applications are approved. In a random sample of 30 applications, what is the expected number that will be turned down?
- $30(.85)$
 - $30(.15)$
 - $30(.85)(.15)$
 - $\sqrt{30(.85)(.15)}$
 - $\sqrt{\frac{(.85)(.15)}{30}}$
- ___ 72. Suppose $P(X) = .25$ and $P(Y) = .40$. If $P(X|Y) = .20$, what is $P(Y|X)$?
- .10
 - .125
 - .32
 - .45
 - .50
- ___ 73. Given the probabilities $P(A) = .3$ and $P(A \cup B) = .7$, what is the probability $P(B)$ is A and B are mutually exclusive? If A and B are independent?
- .4, .3
 - .4, $4/7$
 - $4/7$, .4
 - .7, $4/7$
 - .7, .3
- ___ 74. The average noise level in a bar is 36 decibels with a standard deviation of 5 decibels. Assuming a normal distribution, what is the probability the noise level is between 30 and 40 decibels?
- .327
 - .337
 - .381
 - .673
 - .683

___ 75. Given a random variable X taking three possible values x_1, x_2, x_3 , which of the following statements must be true?

I. $x_1 + x_2 + x_3 = 1$

II. $E(X) = \frac{1}{3} \sum x_i$

III. $\text{var}(X) = \frac{1}{3} \sum (x_i - \bar{x})^2$

- a. I only
- b. III only
- c. II and III
- d. I, II, and III
- e. None of the statements are true.

___ 76. There are 8,253 men and 10,327 women at a state university. If 43 percent of the men and 27 percent of the women are business majors, what is the expected number of business majors in a random sample of 200 students?

- a. 31.7
- b. 34.1
- c. 63.4
- d. 68.2
- e. 70.0

___ 77. Suppose we have a binomial random variable where the probability of exactly four successes is $\binom{n}{4} p^4 (.37)^{n-4}$.

What is the mean of the distribution?

- a. 2.52
- b. 2.59
- c. 4.07
- d. 4.41
- e. 6.93

___ 78. Suppose that 62 percent of the graduates from your high school go on to four-year colleges, 15 percent go on to two-year colleges, 18 percent find employment, and the remaining graduates search for a job. If a randomly selected student is not going on to a four-year college, what is the probability he or she will find employment?

- a. .440
- b. .474
- c. .526
- d. .545
- e. .560

The weights of laboratory cockroaches follow a Normal distribution with mean 80 grams and standard deviation 2 grams.

___ 79. About what percent of the cockroaches have weights between 76 and 84 grams?

- a. 99.7%
- b. 95%
- c. 68%
- d. 47.5%
- e. 34%

___ 80. About what percent of the cockroaches have weights less than 78 grams?

- a. 34%
- b. 32%
- c. 16%
- d. 2.5%
- e. none of these

