

Test 4D (Cumulative) AP Statistics Name: _____

Part 1: Multiple Choice. *Circle the letter corresponding to the best answer.*

1. A political action committee sends out a questionnaire to randomly-selected mailing addresses, asking people to rate the importance of a variety of economic and social issues facing the country. At the end of the questionnaire is an (optional) invitation to donate money to the organization. Which of the following statements about this survey is true?
 - (a) The survey results are invalid because it's impossible to word questions about politics without creating bias.
 - (b) The survey results are invalid because only surveys conducted by phone are free from bias.
 - (c) The survey results will underestimate support for this political action committee in the entire population because people who respond are less likely to support an organization that asks for money.
 - (d) The survey results will overestimate support for this political action committee in the entire population because people more likely to respond if they are prepared to donate money.
 - (e) Since the survey was sent to randomly-selected households, it should be free of bias.

2. Which of the following is the best description of a simple random sample of size n from a population of size N .
 - (a) Any method of sampling in which individuals are selected in a completely haphazard fashion.
 - (b) Any method of sampling in which every group of individuals of size n is equally likely to be selected.
 - (c) Any method of sampling in which every individual is equally likely to be selected.
 - (d) Any method of sampling in which each individual has a probability of n/N of being selected.
 - (e) Any method of sampling that involves the use of a random digits table.

3. An airline has 10 daily flights from Philadelphia to Denver. To assess customer satisfaction, a random sample of 15 passengers from each flight on a single day are asked to fill out a survey about their experience on the flight. What type of sample is this?
 - (a) Simple random sample
 - (b) Stratified random sample
 - (c) Multistage sample
 - (d) Cluster sample
 - (e) Convenience sample

4. Some studies of the relationship between car color and frequency of accidents have found that red cars are more likely to be in accidents than black cars, despite how visible they are. Some experts warn that we should not conclude that red cars are less safe than black cars, because of possible confounding. Which of the following best describes what this means?
- (a) It can be hard to determine whether red cars or black cars are more visible on modern highways.
 - (b) Much of the data on car color and accidents is contradictory, thus making it hard to draw definitive conclusions.
 - (c) Much of the data on car color and accidents is subject to bias, thus making it hard to draw definitive conclusions.
 - (d) Studies of this type are all observational, and it is not possible to separate the effect of car color from the type of people who choose to drive red cars.
 - (e) There are too many variables involved in traffic accidents to isolate the effect of car color.
5. A track coach wants to test the effectiveness of a new training program for distance runners. He selects his two fastest runners and, using a coin flip, randomly chooses one to participate in the new program. The other will continue with the standard training regimen. He then picks the next two fastest runners and randomly assigns one to each program. He repeats this process with all his runners, choosing the two fastest remaining each time. He then compares race times within each group of two runners. What is the name for this kind of experimental design?
- (a) Completely randomized design.
 - (b) Randomized block design, but not matched pairs.
 - (c) Stratified random design.
 - (d) Non-random pairs design.
 - (e) Matched pairs design.
6. For the track coach's study described in the previous exercise, which of the following best describes a conclusion that can be drawn?
- (a) We can determine whether race times improve for runners in the new program, but we can't establish cause and effect.
 - (b) We can determine whether the new training program improves race times more than the standard program for any distance runner.
 - (c) We can determine whether the new training program improves race times more than the standard program for the distance runners in this study.
 - (d) We cannot draw any conclusions, since all the volunteers were already members of this coach's team.
 - (e) We cannot draw any conclusions, because there was no control group.
7. What is the principle purpose of control in a randomized comparative experiment?
- (a) Eliminating bias.
 - (b) Isolating the effect of the treatment variable by keeping all other variables the same in the treatment groups.
 - (c) Creating treatment groups that are as similar as possible.
 - (d) Reducing the impact of random variation arising from random assignment to groups.
 - (e) Ensuring that an experiment is conducted in a double blind fashion.

8. Which of the following is not true about standard deviation?
- (a) Standard deviation is always non-negative.
 - (b) Standard deviation measure the typical distance between values in a data set and the mean value.
 - (c) If the units for the data are in inches, standard deviation is in square inches.
 - (d) Standard deviation is very sensitive to outliers.
 - (e) Adding a value to the data set that is equal to the mean will decrease the standard deviation.
9. The time it takes people to finish a certain kind of puzzle is strongly skewed right. Which of the following statements is true about the mean and median of this distribution?
- (a) The mean and median will be exactly equal.
 - (b) The mean and median will be approximately equal, but probably not exactly equal.
 - (c) The mean will be less than the median.
 - (d) The mean will be greater than the median.
 - (e) Which measure of center is greater depends on how strongly skewed the data is.
10. Forty pound bags of bird seed do not have exactly 40 pounds of seed, but the weight of seeds is approximately Normally distributed with a mean of 40 pounds and a standard deviation of 0.25 pounds. Which of the following is the best estimate of the 90th percentile for the weight of these bags of seed?
- (a) 40.3 pounds
 - (b) 40.4 pounds
 - (c) 40.5 pounds
 - (d) 41.0 pounds
 - (e) 43.2 pounds

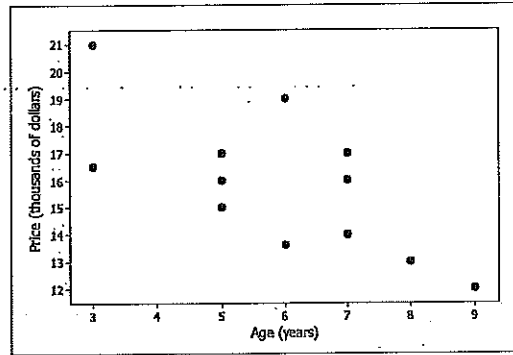
11. Tim Wilson, a psychology professor at the University of Virginia, studies the effect of “personal narratives” on people’s behavior. For example, if the “story” you have about your academic ability is positive, you do better in school than if your “story” is negative. Something as simple as hearing older students describe overcoming challenges similar to yours can help you change your story and improve your performance.

(a) Suppose you have 40 college freshman who have volunteered to be part of a study. Design a completely randomized experiment that tests the hypothesis that hearing older students talk about overcoming challenges improves academic performance. Be sure your design addresses how randomization will be incorporated.

(b) Suppose you have reason to believe that students with a history of strong academic performance respond differently than those with a history of modest performance. Describe how you would incorporate blocking to address this difference in response.

(c) Can this experimental design be carried out in a double blind manner? Explain.

12. The scatterplot at right describes the relationship between the age (in years) and price (in thousands of dollars) for used Ford Explorers.



(a) Describe what the scatterplot reveals about the relationship between the age of cars and their price.

(b) Below is a computer printout of basic statistics for the two variables.

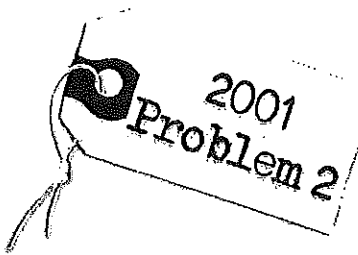
Variable	N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Age	13	5.846	1.772	3.000	5.000	6.000	7.000	9.000
Price	13	15.854	2.445	12.000	13.800	16.000	17.000	21.000

One three-year-old car sells for \$21 thousand. Use the standard scores for this car's age and price to explain how you know that its contribution to correlation is negative.

(c) Below is a computer printout for a least-squares linear regression of Price on Age. Use it to predict the price of a Ford Explorer that is 5 years old.

Predictor	Coef	SE Coef	T	P
Constant	21.405	1.837	11.65	0.000
Age	-0.9496	0.3017	-3.15	0.009

S = 1.85237 R-Sq = 47.4% R-Sq(adj) = 42.6%



Chapter 7.1 "FRAPPY" {Free Response AP Problem...Yay!}

The following problem is taken from an actual Advanced Placement Statistics Examination. Your task is to generate a complete, concise statistical response in 15 minutes. You will be graded based on the AP rubric and will earn a score of 0-4. After grading, keep this problem in your binder for your AP Exam preparation.

A department supervisor is considering purchasing one of two comparable photocopy machines, A or B. Machine A costs \$10,000 and Machine B costs \$10,500. This department replaces photocopy machines every three years. The repair contract for Machine A costs \$50 per month and covers an unlimited number of repairs. The repair contract for Machine B costs \$200 per repair. Based on past performance, the distribution of the number of repairs needed over any one-year period for Machine B is shown below.

Number of Repairs	0	1	2	3
Probability	0.50	0.25	0.15	0.10

Scoring:

You are asked to give a recommendation based on overall cost as to which machine, A or B, along with its repair contract, should be purchased. What would your recommendation be? Give statistical justification to support your recommendation.

E P I

E P I

E P I

E P I

Total: ___/4