

Cumulative AP[®] Practice Test 3

Section I: Multiple Choice Choose the best answer.

- AP3.1 Suppose the probability that a softball player gets a hit in any single at-bat is 0.300. Assuming that her chance of getting a hit on a particular time at bat is independent of her other times at bat, what is the probability that she will not get a hit until her fourth time at bat in a game?
- (a) $\binom{4}{3}(0.3)^1(0.7)^3$ (d) $(0.3)^3(0.7)^1$
 (b) $\binom{4}{3}(0.3)^3(0.7)^1$ (e) $(0.3)^1(0.7)^3$
 (c) $\binom{4}{1}(0.3)^3(0.7)^1$
- AP3.2 The probability that Color Me Dandy wins a horse race at Batavia Downs given good track conditions is 0.60. The probability of good track conditions on any given day is 0.85. What is the probability that Color Me Dandy wins or the track conditions are good?
- (a) 0.94 (b) 0.51 (c) 0.49 (d) 0.06
 (e) The answer cannot be determined from the given information.
- AP3.3 *Sports Illustrated* planned to ask a random sample of Division I college athletes, "Do you believe performance-enhancing drugs are a problem in college sports?" How many athletes must be interviewed to estimate the proportion concerned about use of drugs within $\pm 2\%$ with 90% confidence?
- (a) 17 (c) 1680 (e) 2401
 (b) 21 (d) 1702
- AP3.4 The distribution of grade point averages for a certain college is approximately Normal with a mean of 2.5 and a standard deviation of 0.6. Within which of the following intervals would we expect to find approximately 81.5% of all GPAs for students at this college?
- (a) (0.7, 3.1) (c) (1.9, 3.7) (e) (0.7, 4.3)
 (b) (1.3, 3.7) (d) (1.9, 4.3)
- AP3.5 Which of the following will increase the power of a significance test?
- (a) Increase the Type II error probability.
 (b) Decrease the sample size.
 (c) Reject the null hypothesis only if the P -value is smaller than the level of significance.
 (d) Increase the significance level α .
 (e) Select a value for the alternative hypothesis closer to the value of the null hypothesis.
- AP3.6 You can find some interesting polls online. Anyone can become part of the sample just by clicking on a response. One such poll asked, "Do you prefer watching first-run movies at a movie theater, or waiting until they are available to watch at home or on a digital device?" In all, 8896 people responded, with only 12% (1118 people) saying they preferred theaters. You can conclude that
- (a) American adults strongly prefer watching movies at home or on their digital devices.
 (b) the high nonresponse rate prevents us from drawing a conclusion.
 (c) the sample is too small to draw any conclusion.
 (d) the poll uses voluntary response, so the results tell us little about all American adults.
 (e) American adults strongly prefer seeing movies at a movie theater.
- AP3.7 A certain candy has different wrappers for various holidays. During Holiday 1, the candy wrappers are 30% silver, 30% red, and 40% pink. During Holiday 2, the wrappers are 50% silver and 50% blue. Forty pieces of candy are randomly selected from the Holiday 1 distribution, and 40 pieces are randomly selected from the Holiday 2 distribution. What are the expected value and standard deviation of the total number of silver wrappers?
- (a) 32, 18.4 (c) 32, 4.29 (e) 80, 4.29
 (b) 32, 6.06 (d) 80, 18.4
- AP3.8 A beef rancher randomly sampled 42 cattle from her large herd to obtain a 95% confidence interval to estimate the mean weight of the cows in the herd. The interval obtained was (1010, 1321). If the rancher had used a 98% confidence interval instead, the interval would have been
- (a) wider and would have less precision than the original estimate.
 (b) wider and would have more precision than the original estimate.
 (c) wider and would have the same precision as the original estimate.
 (d) narrower and would have less precision than the original estimate.
 (e) narrower and would have more precision than the original estimate.

AP3.9 School A has 400 students and School B has 2700 students. A local newspaper wants to compare the distributions of SAT scores for the two schools. Which of the following would be the most useful for making this comparison?

- (a) Back-to-back stemplots for A and B
- (b) A scatterplot of A versus B
- (c) Dotplots for A and B drawn on the same scale
- (d) Two relative frequency histograms of A and B drawn on the same scale
- (e) Two bar graphs for A and B drawn on the same scale

AP3.10 Let X represent the outcome when a fair six-sided die is rolled. For this random variable, $\mu_X = 3.5$ and $\sigma_X = 1.71$. If the die is rolled 100 times, what is the approximate probability that the total score is at least 375?

- (a) 0.0000 (c) 0.0721 (e) 0.9279
- (b) 0.0017 (d) 0.4420

AP3.11 An agricultural station is testing the yields for six different varieties of seed corn. The station has four large fields available, which are located in four distinctly different parts of the county. The agricultural researchers consider the climatic and soil conditions in the four parts of the county as being unequal but are reasonably confident that the conditions within each field are fairly similar throughout. The researchers divide each field into six sections and then randomly assign one variety of corn seed to each section in that field. This procedure is done for each field. At the end of the growing season, the corn will be harvested, and the yield, measured in tons per acre, will be compared. Which one of the following statements about the design is correct?

- (a) This is an observational study because the researchers are watching the corn grow.
- (b) This is a randomized block design with fields as blocks and seed types as treatments.
- (c) This is a randomized block design with seed types as blocks and fields as treatments.
- (d) This is a completely randomized design because the six seed types were randomly assigned to the four fields.
- (e) This is a completely randomized design with 24 treatments—6 seed types and 4 fields.

AP3.12 The correlation between the heights of fathers and the heights of their (grownup) sons is $r = 0.52$, both measured in inches. If fathers' heights were measured in feet instead, the correlation between heights of fathers and heights of sons would be

- (a) much smaller than 0.52.
- (b) slightly smaller than 0.52.
- (c) unchanged; equal to 0.52.
- (d) slightly larger than 0.52.
- (e) much larger than 0.52.

AP3.13 A random sample of 200 New York State voters included 88 Republicans, while a random sample of 300 California voters produced 141 Republicans. Which of the following represents the 95% confidence interval that should be used to estimate the true difference in the proportions of Republicans in New York State and California?

- (a) $(0.44 - 0.47) \pm 1.96 \frac{(0.44)(0.56) + (0.47)(0.53)}{\sqrt{200 + 300}}$
- (b) $(0.44 - 0.47) \pm 1.96 \frac{(0.44)(0.56)}{\sqrt{200}} + \frac{(0.47)(0.53)}{\sqrt{300}}$
- (c) $(0.44 - 0.47) \pm 1.96 \sqrt{\frac{(0.44)(0.56)}{200} + \frac{(0.47)(0.53)}{300}}$
- (d) $(0.44 - 0.47) \pm 1.96 \sqrt{\frac{(0.44)(0.56) + (0.47)(0.53)}{200 + 300}}$
- (e) $(0.44 - 0.47) \pm 1.96 \sqrt{\frac{(0.45)(0.55)}{200} + \frac{(0.45)(0.55)}{300}}$

AP3.14 Which of the following is *not* a property of a binomial setting?

- (a) Outcomes of different trials are independent.
- (b) The chance process consists of a fixed number of trials, n .
- (c) The probability of success is the same for each trial.
- (d) Trials are repeated until a success occurs.
- (e) Each trial can result in either a success or a failure.

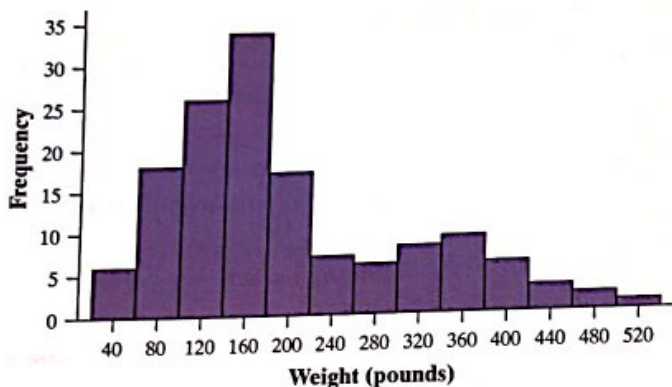
AP3.15 Mrs. Woods and Mrs. Bryan are avid vegetable gardeners. They use different fertilizers, and each claims that hers is the best fertilizer to use when growing tomatoes. Both agree to do a study using the weight of their tomatoes as the response variable. They had each planted the same varieties of tomatoes on the same day and fertilized the plants on the same schedule throughout the growing season. At harvest time, they each randomly select 15 tomatoes from their respective gardens and weigh them. After performing a two-sample t test on the difference in mean weights of tomatoes, they get $t = 5.24$ and $P = 0.0008$. Can the gardener with the larger mean claim that her fertilizer caused her tomatoes to be heavier?

- (a) Yes, because a different fertilizer was used on each garden.



- (b) Yes, because random samples were taken from each garden.
- (c) Yes, because the P -value is so small.
- (d) No, because the soil conditions in the two gardens is a potential confounding variable.
- (e) No, because there was no replication.
- AP3.16** The Environmental Protection Agency is charged with monitoring industrial emissions that pollute the atmosphere and water. So long as emission levels stay within specified guidelines, the EPA does not take action against the polluter. If the polluter is in violation of the regulations, the offender can be fined, forced to clean up the problem, or possibly closed. Suppose that for a particular industry the acceptable emission level has been set at no more than 5 parts per million (5 ppm). The null and alternative hypotheses are $H_0: \mu = 5$ versus $H_a: \mu > 5$. Which of the following describes a Type II error?
- (a) The EPA fails to find convincing evidence that emissions exceed acceptable limits when, in fact, they are within acceptable limits.
- (b) The EPA finds convincing evidence that emissions exceed acceptable limits when, in fact, they are within acceptable limits.
- (c) The EPA finds convincing evidence that emissions exceed acceptable limits when, in fact, they do exceed acceptable limits.
- (d) The EPA takes more samples to ensure that they make the correct decision.
- (e) The EPA fails to find convincing evidence that emissions exceed acceptable limits when, in fact, they do exceed acceptable limits.
- AP3.17** Which of the following is *false*?
- (a) A measure of center alone does not completely describe the characteristics of a set of data. Some measure of spread is also needed.
- (b) If the original measurements are in inches, converting them to centimeters will not change the mean or standard deviation.
- (c) One of the disadvantages of a histogram is that it doesn't show each data value.
- (d) Between the range and the interquartile range, the IQR is a better measure of spread if there are outliers.
- (e) If a distribution is skewed, the median and interquartile range should be reported rather than the mean and standard deviation.
- AP3.18** A 96% confidence interval for the proportion of the labor force that is unemployed in a certain city is (0.07, 0.10). Which of the following statements about this interval is true?
- (a) The probability is 0.96 that between 7% and 10% of the labor force is unemployed.
- (b) About 96% of the intervals constructed by this method will contain the true proportion of unemployed in the city.
- (c) In repeated samples of the same size, there is a 96% chance that the sample proportion will fall between 0.07 and 0.10.
- (d) The true rate of unemployment lies within this interval 96% of the time.
- (e) Between 7% and 10% of the labor force is unemployed 96% of the time.
- AP3.19** A large toy company introduces a lot of new toys to its product line each year. The company wants to predict the demand as measured by y , first-year sales (in millions of dollars) using x , awareness of the product (as measured by the percent of customers who had heard of the product by the end of the second month after its introduction). A random sample of 65 new products was taken, and a correlation of 0.96 was computed. Which of the following is a correct interpretation of this value?
- (a) Ninety-six percent of the time, the least-squares regression line accurately predicts first-year sales.
- (b) About 92% of the time, the percent of people who have heard of the product by the end of the second month will correctly predict first-year sales.
- (c) About 92% of first-year sales can be accounted for by the percent of people who have heard of the product by the end of the second month.
- (d) For each increase of 1% in awareness of the new product, the predicted sales will go up by 0.96 million dollars.
- (e) About 92% of the variation in first-year sales can be accounted for by the least-squares regression line with percent of people who have heard of the product by the end of the second month as the explanatory variable.
- AP3.20** Final grades for a class are approximately Normally distributed with a mean of 76 and a standard deviation of 8. A professor says that the top 10% of the class will receive an A, the next 20% a B, the next 40% a C, the next 20% a D, and the bottom 10% an F. What is the approximate maximum grade a student could attain and still receive an F for the course?
- (a) 70 (c) 65.75 (e) 57
 (b) 69.27 (d) 62.84

AP3.21 National Park rangers keep data on the bears that inhabit their park. Below is a histogram of the weights of 143 bears measured in a recent year.



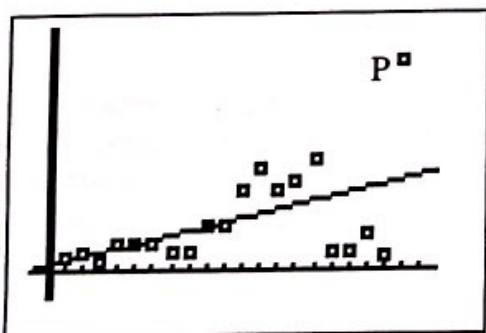
Which statement below is correct?

- The median will lie in the interval (140, 180), and the mean will lie in the interval (180, 220).
- The median will lie in the interval (140, 180), and the mean will lie in the interval (260, 300).
- The median will lie in the interval (100, 140), and the mean will lie in the interval (180, 220).
- The mean will lie in the interval (140, 180), and the median will lie in the interval (260, 300).
- The mean will lie in the interval (100, 140), and the median will lie in the interval (180, 220).

AP3.22 A random sample of size n will be selected from a population, and the proportion of those in the sample who have a Facebook page will be calculated. How would the margin of error for a 95% confidence interval be affected if the sample size were increased from 50 to 200?

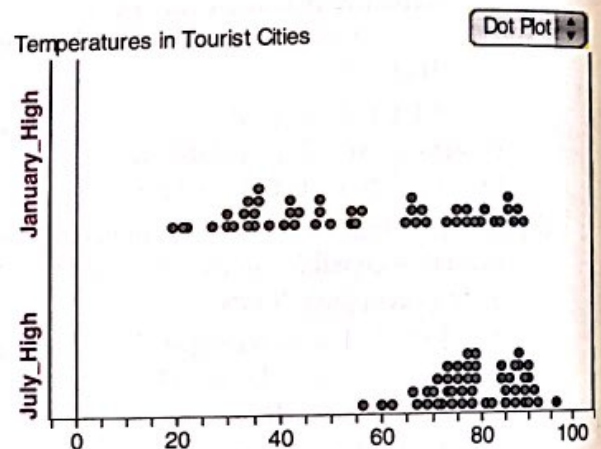
- It remains the same.
- It is multiplied by 2.
- It is multiplied by 4.
- It is divided by 2.
- It is divided by 4.

AP3.23 A scatterplot and a least-squares regression line are shown in the figure below. What effect does point P have on the slope of the regression line and the correlation?



- Point P increases the slope and increases the correlation.
- Point P increases the slope and decreases the correlation.
- Point P decreases the slope and decreases the correlation.
- Point P decreases the slope and increases the correlation.
- No conclusion can be drawn because the other coordinates are unknown.

AP3.24 The following dotplots show the average high temperatures (in degrees Fahrenheit) for a sample of tourist cities from around the world. Both the January and July average high temperatures are shown. What is one statement that can be made with certainty from an analysis of the graphical display?



- Every city has a larger average high temperature in July than in January.
 - The distribution of temperatures in July is skewed right, while the distribution of temperatures in January is skewed left.
 - The median average high temperature for January is higher than the median average high temperature for July.
 - There appear to be outliers in the average high temperatures for January and July.
 - There is more variability in average high temperatures in January than in July.
- AP3.25 Suppose the null and alternative hypotheses for a significance test are defined as

$$H_0: \mu = 40$$

$$H_a: \mu < 40$$

Which of the following specific values for H_a will give the highest power?

- $\mu = 38$
- $\mu = 39$
- $\mu = 40$
- $\mu = 41$
- $\mu = 42$

AP3.26 A large university is considering the establishment of a schoolwide recycling program. To gauge interest in the program by means of a questionnaire, the university takes separate random samples of undergraduate students, graduate students, faculty, and staff. This is an example of what type of sampling design?

- (a) Simple random sample
- (b) Stratified random sample
- (c) Convenience sample
- (d) Cluster sample
- (e) Randomized block design

AP3.27 Suppose the true proportion of people who use public transportation to get to work in the Washington, D.C., area is 0.45. In a simple random sample of 250 people who work in Washington, about how far do you expect the sample proportion to be from the true proportion?

- (a) 0.4975 (c) 0.0315 (e) 0
- (b) 0.2475 (d) 0.0009

Questions 28 and 29 refer to the following setting. According to sleep researchers, if you are between the ages of 12 and 18 years old, you need 9 hours of sleep to be fully functional. A simple random sample of 28 students was chosen from a large high school, and these students were asked how much sleep they got the previous night. The mean of the responses was 7.9 hours, with a standard deviation of 2.1 hours.

AP3.28 If we are interested in whether students at this high school are getting too little sleep, which of the following represents the appropriate null and alternative hypotheses?

- (a) $H_0: \mu = 7.9$ and $H_a: \mu < 7.9$
- (b) $H_0: \mu = 7.9$ and $H_a: \mu \neq 7.9$
- (c) $H_0: \mu = 9$ and $H_a: \mu \neq 9$
- (d) $H_0: \mu = 9$ and $H_a: \mu < 9$
- (e) $H_0: \mu \leq 9$ and $H_a: \mu \geq 9$

AP3.29 Which of the following is the test statistic for the hypothesis test?

- (a) $t = \frac{7.9 - 9}{\frac{2.1}{\sqrt{28}}}$
- (b) $t = \frac{9 - 7.9}{\frac{2.1}{\sqrt{28}}}$
- (c) $t = \frac{7.9 - 9}{\sqrt{\frac{2.1}{28}}}$
- (d) $t = \frac{7.9 - 9}{\frac{2.1}{\sqrt{27}}}$
- (e) $t = \frac{9 - 7.9}{\frac{2.1}{\sqrt{27}}}$

AP3.30 Shortly before the 2012 presidential election, a survey was taken by the school newspaper at a very large state university. Randomly selected students were asked, "Whom do you plan to vote for in the upcoming presidential election?" Here is a two-way table of the responses by political persuasion for 1850 students:

Candidate of choice	Political Persuasion			Total
	Democrat	Republican	Independent	
Obama	925	78	26	1029
Romney	78	598	19	695
Other	2	8	11	21
Undecided	32	28	45	105
Total	1037	712	101	1850

Which of the following statements about these data is true?

- (a) The percent of Republicans among the respondents is 41%.
- (b) The marginal distribution of the variable choice of candidate is given by Obama: 55.6%; Romney: 37.6%; Other: 1.1%; Undecided: 5.7%.
- (c) About 11.2% of Democrats reported that they planned to vote for Romney.
- (d) About 44.6% of those who are undecided are Independents.
- (e) The conditional distribution of political persuasion among those for whom Romney is the candidate of choice is Democrat: 7.5%; Republican: 84.0%; Independent: 18.8%

Section II: Free Response Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

AP3.31 A researcher wants to determine whether or not a five-week crash diet is effective over a long period of time. A random sample of 15 dieters is selected. Each person's weight is recorded before starting the diet and one year after it is concluded. Based on the data shown at right (weight in pounds), can we conclude that the diet has a long-term effect, that is, that dieters manage to not regain the weight they lose? Include appropriate statistical evidence to justify your answer.

	1	2	3	4	5	6	7	8
Before	158	185	176	172	164	234	258	200
After	163	182	188	150	161	220	235	191
	9	10	11	12	13	14	15	
Before	228	246	198	221	236	255	231	
After	228	237	209	220	222	268	234	