Name: _____

Period: _____

AP Stats Review Take Home Test

Multiple Choice Questions: Choose the best answer. Bubble in your answer on the scantron.

1. Here are the IQ scores of 10 randomly chosen fifth-grade students:

   145 139 126 122 125 130 96 110 118 118

   Which of the following statements about this data set is *not true*?

   (a) The student with an IQ of 96 is considered an outlier by the 1.5 X IQR Rule
   (b) The five number summary of the 10 IQ scores is 96, 118, 123.5, 130, 145
   (c) If the value 96 were removed from the data set, the mean of the remaining 9 IQ scores would be greater than the mean of all 10 IQ scores
   (d) If the value 96 were removed from the data set, the standard deviation of the remaining 9 IQ scores would be less than the standard deviation of all 10 IQ scores.
   (e) If the value 96 were removed from the data set, the IQR of the remaining 9 IQ scores would be less than the IQR of all 10 IQ scores.

2. You work for an advertising agency that is preparing a new television commercial to appeal to women. You have been asked to design an experiment to compare the effectiveness of three versions of the commercial. Each subject will be shown one of the three versions and then asked about her attitude toward the product. You think there may be large differences between women who are employed and those who are not. Because of these differences, you should use

   (a) A block design, but not a matched pairs design
   (b) A completely randomized design
   (c) A matched pairs design
   (d) A simple random sample
   (e) A stratified random sample.

3. The correlation between the age and height of children under the age of 12 is found to be r=.60. Suppose we use the age x of a child to predict the height y of the child. What can we conclude?

(a) The height is generally 60% of a child's weight.
(b) About 60% of the time, age will accurately predict height
(c) Thirty-six percent of the variation in height is accounted for by the linear model relating height to age.
(d) For every 1 year older a child is, the regression line predicts an increase of .6 feet in height.
(e) Thirty six percent of the time, the least-squares regression line accurately predicts height from age.

4. An agronomist wants to test three different types of fertilizer (A, B, and C) on the yield of a new variety of wheat. The yield will be measured in bushels per acre. Six 1-acre plots of land were randomly assigned to each of the three fertilizers. The treatment, experimental unit, and response variable are, respectively,

(a) A specific fertilizer, bushels per acre, a plot of land.
(b) A plot of land, bushels per acre, a specific fertilizer.
(c) Random assignment, a plot of land, wheat yield
(d) A specific fertilizer, a plot of land, wheat yield
(e) A specific fertilizer, the agronomist, wheat yield.

5. Which one of the following would be a correct interpretation if you have a z-score of +2.0 on an exam?

(a) It means you missed two questions on the exam
(b) It means that you got twice as many questions correct as the average student
(c) It means that your grade was 2 points higher than the mean grade on this exam.
(d) It means that your grade was in the upper 2% of all grades on this exam.
(e) It means that your grade is 2 standard deviations above the mean for this exam.

6. 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
(e) Select a value for the alternative hypothesis closer to the value of the null hypothesis.

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
   (b) 32, 6.06
   (c) 32, 4.29
   (d) 80, 18.4
   (e) 80, 4.29

8. A 96% confidence interval for the proportion of the labor force that is unemployed in a certain city is (.07,.10). Which of the following statements about this interval is true?

   (a) The probability is .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 .07 and .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.

9. 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

10. 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.

11. 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.

12. 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) $(.44-.47) \pm 1.96\frac{(.44)(.56)+(.47)(.53)}{\sqrt{200+300}}$
(b) $(.44-.47) \pm 1.96\frac{(.44)(.56)}{\sqrt{200}} + \frac{(.47)(.53)}{\sqrt{300}}$
(c) $(.44-.47) \pm 1.96\sqrt{\frac{(.44)(.56)}{200} + \frac{(.47)(.53)}{300}}$
(d) $(.44-.47) \pm 1.96\sqrt{\frac{(.44)(.56)+(.47)(.53)}{200+300}}$
(e) $(.44-.47) \pm 1.96\sqrt{\frac{(.45)(.55)}{200} + \frac{(.45)(.55)}{300}}$

13. 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?

(a) It remains the same
(b) It is multiplied by 2
(c) It is multiplied by 4
(d) It is divided by 2
(e) It is divided by 4

14. The school board in a certain school district obtained a random sample of 200 residents and asked if they were in favor of raising property taxes to fund the hiring of more statistics teachers. The resulting confidence interval for the true proportion of residents in favor of raising taxes was (.183,.257). The margin of error for this confidence interval is

(a) .037
(b) .183

(c) .220
(d) .257
(e) .740

15. A large machine is filled with thousands of small pieces of candy, 40% of which are orange. When money is deposited, the machine dispenses 60 randomly selected pieces of candy. The machine will be recalibrated if a group of 60 candies contains fewer than 18 that are orange. What is the approximate probability that this will happen if the machine is working correctly?

(a) $P\left(z < \frac{.3-.4}{\sqrt{\frac{(.4)(.6)}{60}}}\right)$

(b) $P\left(z < \frac{.3-.4}{\sqrt{\frac{(.3)(.7)}{60}}}\right)$

(c) $P\left(z < \frac{.3-.4}{\frac{\sqrt{(.4)(.6)}}{60}}\right)$

(d) $P\left(z < \frac{.3-.4}{\frac{(.4)(.6)}{\sqrt{60}}}\right)$

(e) $P\left(z < \frac{.4-.3}{\frac{\sqrt{(.3)(.7)}}{60}}\right)$

16. Which sampling method was used in each of the following settings, in order from I to IV

I. A student chooses for a survey the first 20 students to arrive at school
II. The name of each student in a school is written on a card, the cards are well mixed, and 10 names are drawn.
III. A state agency randomly selects 50 people from each of the state's senatorial districts
IV. A city council randomly selects eight city blocks and then surveys all the voting age residents of those blocks.

(a) Voluntary response, SRS, stratified, cluster
(b) Convenience, SRS, stratified, cluster
(c) Convenience, cluster, SRS, stratified
(d) Convenience, SRS, cluster, stratified
(e) Cluster, SRS, stratified, convenience

17. Which of the following statements about the t distribution with degrees of freedom df is (are) true?

I. It is symmetric

II. It has more variability than the t distribution with df+1 degrees of freedom

III. As df increases, the t distribution approaches the standard normal distribution

(a) I only
(b) II only
(c) III only
(d) I and III
(e) I,II, and III

18. The table below provides data on the political affiliation and opinion about the death penalty of 850 randomly selected voters from a congressional district.

	Favor	Oppose	total
Republican	299	98	397
Democrat	77	171	248
Other	118	87	205
Total	494	356	850

Which of the following does not support the conclusion that being republican and favoring the death penalty are not independent?

(a) $\frac{299}{494} \neq \frac{98}{356}$
(b) $\frac{299}{494} \neq \frac{397}{850}$
(c) $\frac{494}{850} \neq \frac{299}{397}$
(d) $\frac{494}{850} \neq \frac{397}{850}$
(e) $\frac{(397)(494)}{850} \neq 299$

19. In USA today/Gallup Poll taken shortly after the game, 64% of respondents who described themselves as baseball fans said Major League Baseball should overturn an umpire's safe call that cost pitcher Armando Gallaraga a perfect game. The poll, based on 470 respondents, has a margin of error of 6%. Which of the following statements best describes what is meant by the 6% margin of error?

(a) About 6% of the baseball fans agreed with the umpire's call
(b) About 6% of those polled were not actually baseball fans
(c) The difference between the sample percentage and the population percentage is likely to be less than or equal to 6%.
(d) About 6% of the sample should not be included in the population
(e) The difference between the percentage of people in favor of overturning the call in this survey and in a second similar survey would be less than 6%

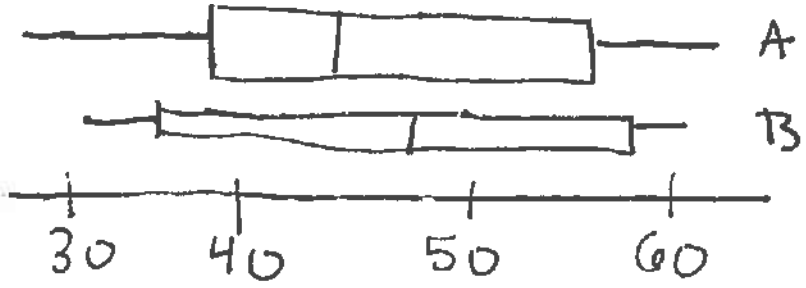
20. A fertilizer company is trying to convince corn farmers using a particular fertilizer that their new product cornplosion will lead to an increase in yield over the currently used fertilizer. To determine which fertilizer results in a higher yield of corn, 10 one-acre fields are split into halves. The current fertilizer is used on one-half of each field and cornplosion on the other half. The fertilizer that goes on the east half of the field is determined by coin flip. The yield of corn is measured in bushels per acre. The data are given below:

Field	Generic Fertilizer	Cornplosion
1	15	21
2	12	16
3	8	7
4	22	30
5	14	12
6	25	33
7	11	14
8	12	21
9	20	22
10	18	16

What is the number of degrees of freedom associated with the appropriate t test for testing to see if there is a corn difference between the mean corn yield for the two types of fertilizer?

(a) 9
(b) 10
(c) 13
(d) 18
(e) 20

21.



The boxplots above summarize two data sets, A and B. Which of the following must be true?

I. Set B has a higher median than set A

II. Set B contains more data than set A

III. The data in set B have a large range than the data in set A

(a) I only
(b) III only
(c) I and II only
(d) I and III only
(e) I, II, and III

22. The back-to-back stem and leaf plot below gives the growth in millimeters of 23 onion bulbs planted by students in Mrs. Shelton's third grade class.

After one week		After two weeks
9 8 8 7 6 6 6 5 4 4 2 1 1 0	0	
8 6 5 5 3 3 2	1	7
4	2	5 6 6 6 7 9
	3	0 0 0 0 1 2 3 3 5 6 9
	4	2 2 3 6 8
	5	

Which of the following statements is not justified by the data?

(a) The plants shown grew on average less than 25 mm per week
(b) The mean size of the plants increased from week one to week two
(c) The spread between the shortest and tallest plant did not change much between the two weeks
(d) The size of every plant increased from week one to week two
(e) The median size of the plants increased from week one to weak two.

23. Laws have been proposed in some border states that would allow police officers to ask motorists to prove that they are U.S. citizens during routine traffic stops. If it is determined that these motorists are not in the United States legally, then they will be deported. Thinking of this in terms of a hypothesis test with H_o is that the motorist is in the United States legally and H_a is that the motorist is not in the United States legally. Which of the following is an example of a type I error?

(a) A motorist is in the country legally but is thought to not be in the country legally
(b) A motorist is not in the country legally but is thought to be in the country legally
(c) A motorist is not in the country legally but when stopped gets deported.
(d) A motorist is not in the country legally but never gets stopped by the police
(e) A motorist is in the country legally and is not stopped by the police.

24. A window manufacturer claims that its new seal-tight system will save homeowners in cold climates $100 a month on average during the winter. A consumer group is skeptical of this claim and thinks that the manufacturer may be overstating the savings. If μ represents the true mean savings for this new window system, which of the following gives the null and alternative hypotheses that the consumer groups should test?

(a) $H_o: \mu \leq \$100, H_a: \mu > \$100$
(b) $H_o: \mu = \$100, H_a: \mu \neq \$100$
(c) $H_o: \mu < \$100, H_a: \mu \geq \$100$
(d) $H_o: \mu = \$100, H_a: \mu < \$100$
(e) $H_o: \mu = \$100, H_a: \mu > \$100$

25. Before going to college, Chelsea wanted to see if she could sell her old car online. She selected a sample of the prices on 25 similar cars sold online in the past month. Later, Chelsea realized that the lowest priced car in the sample was mistakenly recorded as half the price it actually sold for. However, after correcting the error, the corrected price was still less than or equal to any other car price in her sample. Which of the following sample statistics must have remained the same after the correction was made?

(a) Mean
(b) Median
(c) Mode
(d) Range
(e) Variance

26. The number of people in a gym at noon on each of 60 randomly selected days produces a mean of 34.7 people and a standard deviation of 8.1 people. Which of the following is an approximate 90% confidence interval for the mean number of people in the gym at noon?

(a) (26.70,42,70)
(b) (32.95,36.45)
(c) (32.65,36.75)
(d) (33.65, 35.75)
(e) (34.48, 34.92)

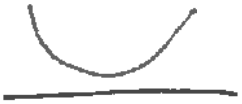
27. Which of the following represents a distribution where you would expect the mean is greater than the median?

(A) 

(B) 

(C) 

(D) 

(E) 

28. There is a linear relationship between the amount of fat in a sandwich and the amount of calories in a sandwich. A least squares regression line was fit using some data collected by a nutritionist, resulting in:

$$y = 217.3 + 35.2x$$

Where x is the grams of fat in the sandwich and y is the estimated caloric content of the sandwich. What is the estimated increase in calories that corresponds to an increase of 8 grams of fat?

(a) 35.2
(b) 217.3
(c) 281.6
(d) 498.9
(e) 1738.4

29. In a test of the null hypothesis $H_o: \mu = 23$ against the alternative hypothesis $H_a: \mu > 23$, a random sample from a normal population with a known standard deviation produces a mean of 28.9. The z statistic for the test is 2.48 and the p-Value is .0066. Based on these statistics, which of the following conclusions could be drawn?

(a) There is convincing evidence that $\mu > 23$
(b) Due to random fluctuation, 49.34% of the time a sample produces a mean larger than 23.
(c) .66% of the time, rejecting the alternative hypothesis is an error.
(d) .66% of the time, the mean is above 23.
(e) 99.34% of the time, the mean is below 23.

30. The equation of the least-squares regression line for a set of data is y=.68+1.21x. What is the residual for the point (3,4)

(a) -.31
(b) -.68
(c) -1.52
(d) -3.63
(e) -4.31

31. Which of the following is least likely to be a potential confounding variable in a study of the effect of diet on weight loss in pounds?

(a) Gender
(b) Beginning weight
(c) Time spent exercising
(d) Name
(e) Time spent watching tv

32. The periodic table provides an atomic mass for each element, but these figures do not prove accurate for certain isotopes of some elements. Because isotopes are likely to contain a different mix of neutrons and electrons, the weights of some isotopes can be significantly more or less than the weight of the general element. Such isotopes occur much more rarely than the general element. If you were to plot the weights of the element and its isotopes, the isotopes could be clearly identified as falling far away from the accepted mean. The best word to describe the isotopes as they relate to such a plot in statistics would be:

(a) Outlier
(b) Residual
(c) Deviation
(d) Quartile
(e) minimum

33. A study was conducted by psychologists who wanted to determine at what age children develop their first crush. They asked college freshman to select from one of four age categories, and the results are summarized below:

	Age 0-8	Age 9-12	Age 13-16	Age 16+
Male	2	11	35	10
Female	7	19	23	4

According to the two way table, what percentage of males had their first crush before the age of 13?

(a) 3.4
(b) 18.9
(c) 22.4
(d) 34.5
(e) 49.1

34. The 23 members of the math club decide to hold the "random awards" in which they use a random number table to award 1st, 2nd, and 3rd prizes. Each member is give a two digit number from 01 to 23, then the prizes are drawn in reverse order with 3rd prize being the first member who appears in the random number table below when digits are read two at a time:

92646 90110 79365 04891 39174 39823

Which numbered members of the math club won the three prizes?

(a) 11, 04, 17
(b) 9, 2, 6
(c) 01, 10, 17
(d) 01, 11, 10
(e) 23, 17, 10

35. A church choir wants to attract more members between the ages of 16 and 30. The current choir members believe that early Sunday morning practices may be a problem for this group, so they decided to take a simple random sample of 30 church members in the targeted age group to determine a convenient time to hold practices. Each person in the sample was sent a survey. Ten people responded to the survey, and half of the responses identified Friday nights as their best practice time. If the choir used this information to make a decision to change the practice time to Friday nights, which type of bias might be impacting their decision?

(a) Selection bias
(b) Measurement or response bias
(c) Nonresponse bias
(d) Sampling frame bias
(e) Replacement bias

36. Suppose that the number of dandelions in a one-acre grassland plot is approximately normally distributed with a mean of 750 and a standard deviation of 80. Approximately what percentage of one-acre grassland plots have between 590 and 910 dandelions?

(a) 50%
(b) 68%
(c) 75%
(d) 95%
(e) 99.7%

37. A game is played in which two players (player A and player B) each toss a fair coin. If both tosses result in heads or both tosses result in tails, the players will toss again. If the tosses result in one head and one tail, the player whose toss was the head is the winner. What is the probability that each player must toss exactly three times before a winner declared and that player A is the winner?

(a) ¼
(b) ½
(c) $(\frac{1}{2})^3$
(d) $\frac{1}{2}^2\left(\frac{1}{4}\right)$
(e) $1/4^3$

38. A vaccine has been developed that is designed to reduce the chance of catching a cold during the winter months. In an experiment to evaluate the effectiveness of this vaccine, 500 volunteers were randomly assigned to one of two groups. Those in one group received the vaccine and those in the other group received an injection with no active ingredients. The no active ingredient injection is an example of

(a) Blinding
(b) A confounding variable
(c) A blocking factor
(d) A volunteer treatment
(e) A placebo

39. Suppose that in 2010, the proportion of all cars for sale in California that were hybrids were .08. Suppose that a new government program provided incentives for production of hybrids. In 2011, a random sample of cars for sale in California included 32 hybrids. In trying to establish whether or not the government program was effective in increasing the availability of hybrid vehicles, statisticians are asked to use the sample to determine if this program has led to a statistically significant increase in the proportion of hybrids. Assuming the conditions for inference are met, what is the z test statistic for this test?

(a) .11
(b) .27
(c) .68
(d) 1.70
(e) 5.10

40. It is sometimes stated that the American judicial system is built on the idea that it is better to let 100 guilty men go free than to lock up one innocent man. It is assumed from the saying "presumed innocent" that the null hypothesis of a criminal trial is that the accused is not guilty, which type of error is represented in the following two situations?

Situation 1: A guilty man is found not guilty
Situation 2: An innocent man is found guilty

(a) Both represent Type 1 error
(b) Both represent type II error
(c) Situation 1 represents type I error and situation 2 represents a type II error
(d) Situation 1 represents type II error and situation 2 represents a type I error
(e) Both represent both types of error