

Name _____ Date _____ Per _____

AP REVIEW PACKET #1 INSTRUCTIONS – 30 points
Packet due by Friday, March 8 at 7:50

1) Read Topic One in your Barron's book (pp 35-54). As you read, answer the guided reading problems. It is also suggested that you review the multiple-choice and free-response review questions at the end of the section – they are fair game on the packet quiz, March 14th.

2) Read pp 55-86 Summarizing Distributions in the Baron's book and answer all reading questions. Again, end of section questions are fair game on the packet quiz, but are optional.

3) Answer the multiple choice questions in this packet, these will be graded for accuracy. Show all work and then copy your answers into the blanks below.

1. _____ 2. _____ 3. _____ 4. _____ 5. _____
6. _____ 7. _____ 8. _____ 9. _____ 10. _____

4) Answer the Free-Response Questions in this packet. Show all work and answer the questions completely

Optional Bonus Activity: Go to <http://learner.org/resources/series65.html>. Watch programs 1 & 2 and answer the questions provided. (Click the VoD icon on the right of the program description to access the video.) – credit will not be provided for student who do not complete numbers 1-4 above)

Guided Reading Questions for Packet #1:

1) What are the most common visual representations of data?

1. _____ 2. _____ 3. _____

2) The first thing to do with data is to _____ - always.

3) What type of variables are dotplots particularly useful for? _____

4) T/ F: A variable that has numerical values is always quantitative.

5) When are relative frequencies useful?

6) Define the following terms:

a) Center

b) Spread

c) Clusters

d) Gaps

e) Modes

Name _____ Date _____ Per _____

7) Draw example of each of the following types of distributions:

a) Symmetric

b) Skewed to the right

c) Skewed to the left

d) Bell-shaped

e) Uniform

8) What is an ogive?

9) Copy the summary bullet points on pg 44.

10) The word average is derived from the French _____, which refers to the _____ that shippers contributed to help compensate for losses suffered by other shippers whose _____. i.e. losses were _____.

11) The variance is determined by _____.

12) The standard deviation is the _____ of the _____.

13) What is the numerical rule used for designating outliers?

14) Percentile ranking indicates _____.

15) The z-score states very specifically _____.

16) The empirical rule (also called the ___ - ___ - ___ rule) applies specifically to _____.

17) For distributions that are skewed to the right, the mean is usually _____ than its median.

18) Read through the summary on pg 71: Pick 3 statements that you need to review.

1-

2-

3-

Packet Multiple Choice Questions:

1. An occupational safety administrator is interested in reducing repetitive motion injuries for office workers who use computers. The administrator divides 48 volunteers at random into two groups. The first group will use a standard keyboard for 12 months. The second group will use a new keyboard design for the same period of time. At the end of the study the frequencies of volunteers experiencing repetitive motion injuries with their respective keyboards will be compared. What would be the appropriate method of inference in this situation?
 - (A) One-sample t -test
 - (B) Paired t -test
 - (C) Two-sample t -test
 - (D) One-proportion z -test
 - (E) Two-proportion z -test

2. A simple random sample of size n is taken from a large population whose distribution of the variable under consideration is extremely right-skewed. Which of the following statements is true?
 - (A) As n increases, the sample mean is more likely to be within a given distance of the population mean.
 - (B) When $n > 30$, the distribution of the sample mean is normal.
 - (C) As n increases, the sample standard deviation decreases.
 - (D) As n increases, the distribution of the sample data becomes more normal.
 - (E) The sample standard deviation is equal to
$$\frac{\text{population standard deviation}}{\sqrt{n}}$$

3. The power of a significance test for a particular value of the parameter is computed to be 0.93. Which statement below is true?
 - (A) The probability of committing a Type I error is 0.07.
 - (B) The probability of committing a Type I error is 0.93.
 - (C) The probability of committing a Type II error is 0.07.
 - (D) The probability of committing a Type II error is 0.93.
 - (E) The probability of committing a Type II error is the same as the alpha level.

4. John recently scored 113 on a particular standardized achievement test. The scores on the test are distributed with a mean of 100 and a standard deviation of 10. His cousin, Brandon, took a different standardized test and scored 263. The scores on Brandon's test have a mean of 250 and a standard deviation of 25. Which student did relatively better on his particular test?
- (A) John did better on his test.
 - (B) Brandon did better on his test.
 - (C) They both performed equally as well on their respective tests.
 - (D) It is impossible to tell since they did not take the same test.
 - (E) It is impossible to tell since the number of students taking the test is unknown.
5. Who makes more mistakes on their income tax forms: accountants or taxpayers who prepare the forms themselves? A random sample of income tax forms that were prepared by accountants was drawn from IRS records.

An equal number of forms that were self-prepared by taxpayers was also drawn. The average number of errors per form was compared to determine if one group tends to make more mistakes than the other. What type of study is this?

- (A) Census
 - (B) Experiment
 - (C) Voluntary response survey
 - (D) Observational study
 - (E) Matched-pairs study
6. A study was conducted to estimate the proportion of American families who owned a video tape player (VCR) and/or a digital video disc player (DVD). A random sample of a large group of Americans was taken. The 95% confidence interval created from the data produced the interval (0.784, 0.844). Which of the following is the correct interpretation of 95% confidence?
- (A) We are 95% confident that the true proportion of American families that own a VCR or DVD player is between 0.784 and 0.844.
 - (B) Ninety-five percent of the time, a sample such as this one will produce a sample proportion between 0.784 and 0.844.
 - (C) There is a 95% chance that the sample proportion from the data is between 0.784 and 0.844.
 - (D) Ninety-five percent of all intervals created like this one will contain the true population proportion.
 - (E) There is a 95% chance that the population proportion is between 0.784 and 0.844.

7. The baggage handling services of On-Time Airlines is interested in how many baggage handlers they need on duty at various times of the day to ensure that passengers do not wait an unreasonable amount of time for their baggage. An airport executive performed a study and found that there is a correlation between the number of passengers arriving at given times and the number of baggage handlers needed. She sampled various times during the day and different days of the week including weekend. She recorded the number of passengers arriving within any 1-hour time block.

The computer output from the regression equation analysis is shown below.

Predicted Baggage Handlers = $2.86 + 0.00408$
(number of passengers)

Predictor	Coef	StDev	T	P
Constant	2.860	1.324	2.16	0.083
Passengers	0.004081	0.001168	3.49	0.017

S = 1.562 R-sq = 70.9% R-Sq(adj) = 65.1%

What is the value of the correlation coefficient for the number of baggage handlers and number of arriving passengers?

- (A) -0.842
(B) 0.651
(C) 0.709
(D) 0.842
(E) 1.562
8. A dance club holds a raffle at the end of each dance. Five dancers are selected at random to each draw one numbered tag from a hat without replacement. There are 50 tags in the hat numbered from 1 to 50. Drawing a tag from 1 through 5 wins \$20, tags from 6 through 25 wins \$10, and tags from 26 through 50 wins \$5. In order to determine the average amount of money paid out, a simulation will be conducted using a random number table. Which of the following assignments of random numbers to tag values is most appropriate for the simulation?
- (A) Using single-digit numbers, assign 0 to represent a \$20 prize, 1–4 to represent a \$10 prize, and 5–9 to represent a \$5 prize.
(B) Using single-digit numbers, assign 0 to represent a \$20 prize, 1 to represent a \$10 prize, and 2 to represent a \$5 prize. Numbers 3–9 are ignored.
(C) Using two-digit numbers, assign 20 to represent a \$20 prize, 10 to represent a \$10 prize, and 05 to represent a \$5 prize. Numbers 00–04, 06–09, 11–19, 21–99 are ignored.
(D) Using two-digit numbers, assign 01–05 to represent a \$20 prize, 06–25 to represent a \$10 prize, and 26–50 to represent a \$5 prize. Numbers 51–99 and 00 are ignored.
(E) Using two-digit numbers, assign 01–10 to represent a \$20 prize, 11–40 to represent a \$10 prize, and 41–99 and 00 to represent a \$5 prize.

9. The student council wants to survey students at the school to see what brands of soda pop they want in the school machines. They randomly sampled 30 freshmen, 30 sophomores, 30 juniors, and 30 seniors. The sampling method they used is a
- (A) simple random sample
 - (B) stratified random sample
 - (C) cluster sample
 - (D) systematic random sample
 - (E) convenience sample

10. Which of the following is a legitimate probability distribution?

(A)

x	3	4	5	6
$P(x)$	0.1	0.1	0.1	0.1

(B)

x	0	1	2	3	4	5
$P(x)$	0.2	0.1	0.1	0.2	0.2	0.3

(C)

x	-5	0	5	10	15
$P(x)$	0.6	0.1	0.2	0.1	0

(D)

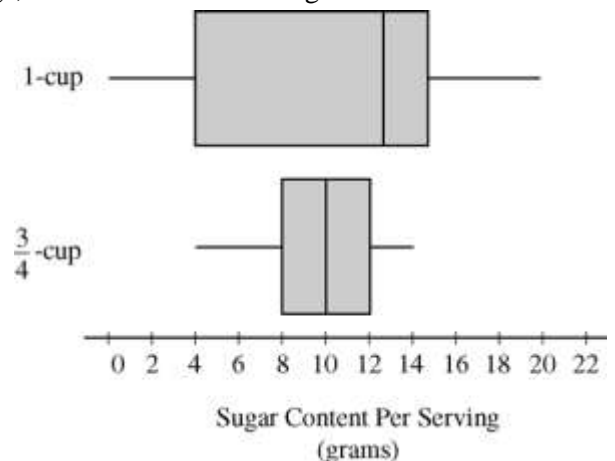
x	1	3	5	7
$P(x)$	0.3	0.2	0.2	0.4

(E)

x	1	1.5	2	2.5	3	3.5	4	4.5
$P(x)$	0.05	0.08	0.07	0.20	0.30	0.10	0.21	-0.01

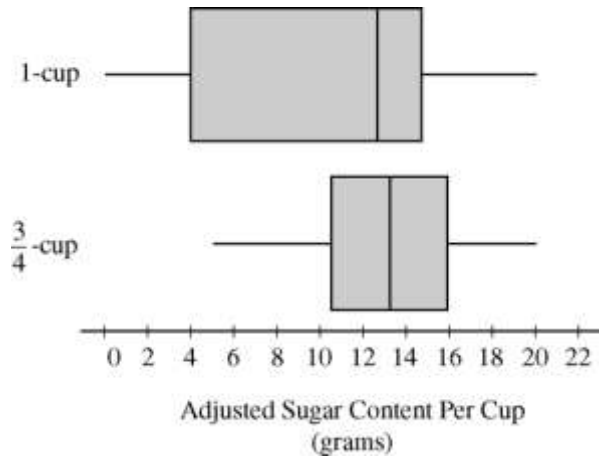
Free Response Questions for Packet #1: (For #1 use the blank space to the right, 2-3 the space below each question)

2. To determine the amount of sugar in a typical serving of breakfast cereal, a student randomly selected 60 boxes of different types of cereal from the shelves of a large grocery store. The student noticed that the side panels of some of the cereal boxes showed sugar content based on one-cup servings, while others showed sugar content based on three-quarter-cup servings. Many of the cereal boxes with side panels that showed three-quarter-cup servings were ones that appealed to young children, and the student wondered whether there might be some difference in the sugar content of the cereals that showed different-size servings on their side panels. To investigate the question, the data were separated into two groups. One group consisted of 29 cereals that showed one-cup serving sizes; the other group consisted of 31 cereals that showed three-quarter-cup serving sizes. The boxplots shown right display sugar content (in grams) per serving of the cereals for each of the two serving sizes.



- (a) Write a few sentences to compare the distributions of sugar content per serving for the two serving sizes of cereals.

After analyzing the boxplots on the preceding page, the student decided that instead of a comparison of sugar content per recommended serving, it might be more appropriate to compare sugar content for equal-size servings. To compare the amount of sugar in serving sizes of one cup each, the amount of sugar in each of the cereals showing three-quarter-cup servings on their side panels was multiplied by $\frac{4}{3}$. The bottom boxplot shown right displays sugar content (in grams) per cup for those cereals that showed a serving size of three-quarter-cup on their side panels.



(b) What new information about sugar content do the boxplots above provide?

(c) Based on the boxplots shown above on this page, how would you expect the mean amounts of sugar per cup to compare for the different recommended serving sizes? Explain.

3. A professional sports team evaluates potential players for a certain position based on two main characteristics, speed and strength.

(a) Speed is measured by the time required to run a distance of 40 yards, with smaller times indicating desirable (faster) speeds. From previous speed data for all players in this position, the times to run 40 yards have a mean of 4.60 seconds and a standard deviation of 0.15 seconds, with a minimum time of 4.40 seconds, as shown in the table below.

	Mean	Standard Deviation	Minimum
Time to run 40 yards	4.60 seconds	0.15 seconds	4.40 seconds

Based on the relationship between the mean, standard deviation, and minimum time, is it reasonable to believe that the distribution of 40-yard running times is approximately normal? Explain.

Name _____ Date _____ Per _____

(b) Strength is measured by the amount of weight lifted, with more weight indicating more desirable (greater) strength. From previous strength data for all players in this position, the amount of weight lifted has a mean of 310 pounds and a standard deviation of 25 pounds, as shown in the table below.

	Mean	Standard Deviation
Amount of weight lifted	310 pounds	25 pounds

Calculate and interpret the z -score for a player in this position who can lift a weight of 370 pounds.

(c) The characteristics of speed and strength are considered to be of equal importance to the team in selecting a player for the position. Based on the information about the means and standard deviations of the speed and strength data for all players and the measurements listed in the table below for Players A and B, which player should the team select if the team can only select one of the two players? Justify your answer.

	Player A	Player B
Time to run 40 yards	4.42 seconds	4.57 seconds
Amount of weight lifted	370 pounds	375 pounds

Name _____ Date _____ Per _____

Video 1 Worksheet. WHAT IS STATISTICS?

1. What is the narrator's job? _____
2. How were the two groups of children different in the creativity experiment?

3. Tell one way that Domino's used statistics. _____
4. What are the three steps in using statistics? a. _____,
b. _____, c. _____.
5. Based on the study shown, when does lightning usually begin? _____
6. When did Sarah's growth rate start to fall below that which was considered to be normal? _____
7. What relationship did researchers find between the number of manatees and the number of boat registrations? _____
8. Is there a relationship between number of home runs and salaries? _____ If so, what is it?

9. What is a placebo? _____
10. Of the studies shown, which used a placebo? _____
11. Why was the study of potato chips being done? _____
12. What three things were done to ensure accurate results in the large survey shown?
 - a. _____
 - b. _____
 - c. _____
13. As a result of the space shuttle Challenger explosion, what statistical tool was instrumental in overhauling the entire program? _____

Name _____ Date _____ Per _____

14. Is running a casino a profitable business? Why? _____

15. List four other studies shown in this video.

a. _____ b. _____

c. _____ d. _____

16. What was the outcome of the creativity experiment? _____

17. Were the children who were rewarded with a prize for creativity **more** or **less** creative in the study shown?

_____. The conclusion was that competition _____ creativity.

Optional: Video 2 Worksheet - PICTURING DISTRIBUTIONS

1. What is the overall pattern in a set of observations called? _____
2. What type of picture did Raul Lopez use to plot the lightning data? _____
3. When do most lightning storms begin? _____
4. When is the maximum number of lightning flashes? _____
5. What are observations that stand apart from the overall pattern of the distribution called? _____
6. When looking for the big picture in a distribution, what should you look for first? _____
7. What term is used to describe a distribution whose two sides are mirror images of each other? _____
8. What is the second important aspect of a histogram? _____
9. What is the third important aspect of a histogram? _____
10. What term is used to describe a distribution where one side is more spread out than the other? _____
11. If a distribution trails off to the right, what phrase is used to describe the distribution? _____

Name _____ Date _____ Per _____

12. What determines which way we say a distribution is skewed? _____

13. When constructing a histogram, what is the most important rule? _____

14. Fill in the blank: When constructing histograms, classes that are too large are too small _____

15. What is the term used to describe how spread out the observations are? _____

16. What is an advantage of a stemplot over a histogram? _____

17. How are back-to-back stemplots used? _____