

Excursions in Mathematics
Review for Test on Statistics
Fall, 2004

Name: _____

Problems 1 through 6 refer to the following: As a part of a statistics project, a 6th grade teacher brings to class a container with 200 red marbles and 800 white marbles which are thoroughly mixed. To figure out how many red marbles are in the container without actually counting them all, a student randomly draws 150 marbles from the container. Of the 150 marbles drawn, 33 are red.

1. The data collection method in this example can best be described as

2. The population consists of

3. The N -value for this population is

4. The sample consists of

5. The sampling rate is

6. Suppose that the student is given the N -value. What is a reasonable estimate for the number of red marbles in the container?

Problems 7 through 10 refer to the following: In order to determine the effectiveness of a new vaccine that is alleged to cure “math anxiety,” an experiment was conducted. One thousand volunteer college students enrolled in math courses across the United States were chosen. The 1,000 students were broken up into two groups. Those enrolled in calculus courses or higher were given the real vaccine. The students in remedial and basic math courses were given a fake vaccine consisting of sugared water. None of the students knew whether they were being given the real or the fake vaccine, but the researcher conducting the experiment knew. At the end of the semester the students were given a test that measured their level of math anxiety. The students in the treatment group showed significantly lower levels of math anxiety than those in the control group. On the basis of this experiment the vaccine was advertised as being highly effective in fighting math anxiety.

7. The control group in this experiment consists of

8. This experiment can best be described as a
 - (a) double blind randomized controlled experiment.
 - (b) double blind controlled placebo experiment.
 - (c) blind randomized controlled experiment.
 - (d) blind controlled placebo experiment.

9. This results of this experiment should be considered unreliable because

10. What is the most likely confounding variable for this experiment?

Problems 11 and 12 refer to the capture-recapture method: n_1 denotes the size of the tagged (captured) sample, n_2 denotes the size of the second (recaptured) sample, and k denotes the number of tagged individuals in the second sample.

11. If $n_1 = 250$, $n_2 = 150$, and $k = 25$, a reasonable estimate for the size of the population is

12. If $n_1 = 32$, $n_2 = 50$, and $k = 8$, the N -value of the population is approximately

13. A researcher interested in Cleansburg citizens' attitudes toward a revitalized downtown area surveys a randomly selected group of 200 downtown office workers. 66% of those surveyed indicated that they are in favor of revitalizing the downtown area. The researcher concluded that "about two thirds of the people in Cleansburg are in favor of revitalizing the downtown area." This conclusion might be invalid because
 - (a) there was no control group.
 - (b) 66% is not exactly two thirds.
 - (c) the sample is not representative of the population.
 - (d) the size of the sample is too small.

Problems 14 through 18 refer to the following example. The scores of a group of students on a 30-point multiple-choice exam are given in the following frequency table.

Exam Score	23	24	25	26	27	28	29	30
Frequency	6	6	4	1	6	6	4	1

14. The average score on the exam is
15. The median score on the exam is
16. The third quartile on the exam is
17. The range on the exam is
18. The inter-quartile range on the exam is

Problems 19 through 21 refer to the following data set: $\{-2, -3, 1, 8\}$.

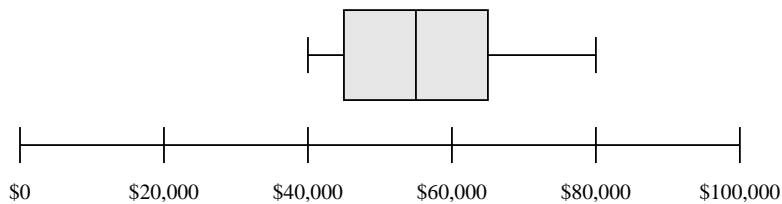
19. The average of the four numbers is
20. The median of the four numbers is
21. The standard deviation of the four numbers is

Problems 22 through 24 refer to a data set consisting of 251 numbers.

22. After sorting the data set (in increasing order from left to right), the median is
 - (a) the number in the 125th position.
 - (b) the number in the 126th position.
 - (c) the average of the numbers in the 125th and 126th positions.
 - (d) the sum of all the numbers divided by 251.

23. After sorting the data set (in increasing order from left to right), the first quartile is
- (a) the number in the 63rd position.
 - (b) the number in the 64th position.
 - (c) the average of the numbers in the 63rd and 64th positions.
 - (d) the average of the numbers in the 62nd and 63rd positions.
24. After sorting the data set (in increasing order from left to right), the third quartile is
- (a) the number in the 63rd position counting backwards from the right.
 - (b) the number in the 64th position counting backwards from the right.
 - (c) the average of the numbers in the 63rd and 64th positions counting backwards from the right.
 - (d) the sum of the median plus the first quartile.
25. Repeat problems 22 through 24, but now use a data set consisting of 252 numbers. (The answers might not be listed above.)
26. Repeat problems 22 through 24, but now use a data set consisting of 253 numbers. (The answers might not be listed above.)
27. Repeat problems 22 through 24, but now use a data set consisting of 254 numbers. (The answers might not be listed above.)

Problems 28 and 29 refer to the following figure. The box plot represents the annual salaries of a group of opera singers.



28. What is the five-number summary for the salaries of these opera singers?
29. The interquartile range of salaries for these opera singers is