

Name: Adrianna Ramirez

Date: 10/10/22

Course: PSY 337

NYACK COLLEGE

Exam #1: Chapters 1-3

Multiple Choice (5 points for questions 1-30)

Directions: Make one selection for each question (unless otherwise stated).

1. Under descriptive statistics, we study

- A) The description of decision making tricks
- B) The methods for organizing, displaying, and describing data
- C) How to describe the probability distribution
- D) Samples to assist in decision making

2. Under inferential statistics, we study

- A) The methods to make decisions about one or more populations based on sample results
- B) How to make decisions about a mean, median, or mode
- C) How a sample is taken from a population
- D) Tables composed of summary measures

3. In statistics, a population consists of:

- A) All people living in a country
- B) All people living in the area under study
- C) All subjects or objects whose characteristics are being studied

D) A selection of a limited number of elements

4. In statistics, conducting a census means:

A) Making decisions based on sample results

B) Checking if a variable is qualitative or quantitative

C) Collecting information from all members of the population

D) Collecting a sample with replacement

5. In statistics, a representative sample is a sample that:

A) Contains the characteristics of the population as closely as possible

B) Represents the results of a sample exactly

C) Contains all people living in an area

D) Contains elements collected with replacement

6. A random sample is a sample drawn in such a way that:

A) Each member of the population has a 0.10 chance of being included in the sample

B) All elements of a population are included

C) Some members of the population have no chance of being included in the sample

D) Each member of the population has some chance of being included in the sample

7. A simple random sample is a sample drawn in such a way that:

A) Each member of the population has some chance of being included in the sample

- B) Every tenth element of an arranged population is included
- C) Each sample of the same size has an equal chance of being selected
- D) Each member of the population has a 0.10 chance for being included in the sample

8. A data set is a:

- A) Set of decisions made about the population
- B) Set of graphs and pictures
- C) Collection of observations on one or more variables
- D) Score collected from an element of the population

9. An observation is a:

- A) Graph observed for a data set
- B) Value of a variable for a single element
- C) Table prepared for a data set
- D) Sample observed from the population

10. A quantitative variable is the only type of variable that can:

- A) Assume numeric values for which arithmetic operations make sense
- B) Be graphed
- C) Be used to prepare tables
- D) Have no intermediate values

11. A discrete variable is a variable that can assume:

A) Categorical values only C) An uncountable set of values

B) A countable set of values only D) Non-numerical values

12. Raw data are the data that:

A) Are presented in the form of a frequency table

B) Give information on each individual sample member separately

C) Are arranged in increasing order

D) Are arranged in a random order

13. We obtain the relative frequency of a category by:

A) Dividing the frequency of that category by the sum of all frequencies

B) Multiplying the frequency of that category by 100

C) Dividing the frequency of that category by 100

D) Dividing the sum of all frequencies by the frequency of that category

14. We obtain the percentage of a category by:

A) Multiplying the frequency of that category by 100

B) Multiplying the relative frequency of that category by 100

C) Dividing the frequency of that category by 100

D) Dividing the sum of all frequencies by the frequency of that category

15. The mean of a data set is the:

- A) Value of the middle term in a ranked data set
- B) Sum of all values divided by the number of values
- C) Difference between the maximum and minimum values
- D) Average of the deviations of values from the average

16. The median of a data set is the:

- A) Value of the middle term in a ranked data set
- B) Value that occurs with maximum frequency
- C) Sum of all values divided by the number of values
- D) Average of the deviations of values from the average

17. You just dropped an outlier from a data set. The value of the mean:

- A) Is now more than the value of the median
- B) Is now less than the value of the median
- C) Is now less than the value of the mode
- D) Can't be determined from the given information

18. An outlier influences which of the following summary measures the most?

- A) Mean
- B) Median
- C) Mode
- D) Median and Mode

19. Which of the following is the only measure that can be calculated for qualitative data?

- A) Mean B) Range C) Mode D) Median

20. If a data set is right-skewed with one peak in the histogram, then which of the following is true?

- A) The values of the mean, median, and mode are the same
 B) The mean is greater than the median, which is greater than the mode
C) The mean and median are equal, but the mode is different
D) The mode is greater than the median, which is greater than the mean

21. If a distribution is symmetric with one peak, then:

- A) The values of the mean, median, and mode are identical
B) The mean is greater than the median, which is greater than the mode
C) The values of the mean and median are equal but the mode is different
D) The mode is greater than the median, which is greater than the mean

22. Which is not a measure of center?

- A) Median
B) Mode
C) Range
D) Mean

23. Which is the least reliable measure of dispersion?

- A) range
- B) standard deviation
- C) coefficient of variation
- D) variance

24. Which is the least reliable measure of center?

- A) Mean
- B) Median
- C) Mode
- D) Midpoint

25. Which definition most accurately describes an ogive plot?

- A) Is a curve drawn using the cumulative frequency distribution
- B) Is a curve drawn using a normal distribution
- C) Is a curve drawn based on a multi-modal distribution
- D) Is a curve derived from skewed-right distribution

26. What is the mode for the population of scores presented in the frequency distribution table below?

X	f
5	1
4	4
3	3
2	4
1	2

- | | |
|-------------------------------------|---|
| a. | 3 |
| <input checked="" type="radio"/> b. | 4 |

c. 2
d. 2 and 4

27. A set of scores ranges from a high of $X = 63$ to a low of $X = 28$. If these scores were put in a grouped frequency distribution table, what would be the best choice for the interval width?

a. 2 points
<input checked="" type="radio"/> b. 5 points
c. 7 points
d. 10 points

$$63 - 28 = 35$$

28. Which statement below is correct regarding a grouped frequency distribution table?

a. The $\sum f$ cannot be determined.
<input checked="" type="radio"/> b. The $\sum X$ cannot be determined.
c. Interval widths should be restricted to either 10 or 20.
d. The bottom score in each class interval should be divisible by 5.

29. For the following grouped frequency distribution table, how many people had scores less than $X = 14$?

X	f
30-34	3
25-29	2
20-24	2
15-19	5
10-14	4
5-9	1

$$\angle X = 14$$

$$\angle 14$$

$$+ = 5$$

<input checked="" type="radio"/> a. 5
b. 1
c. 12
d. Cannot be determined

30. The average verbal SAT score for the entire class of incoming college freshmen in the United States is 530. However, if a sample of 20 incoming college freshmen is randomly selected from the United States, it is likely that this sample's average verbal SAT score will not be exactly 530. This is consistent with the concept of _____.

a. statistical error
b. inferential error
<input checked="" type="radio"/> c. sampling error

Short Response (15 points for questions 31-34)

31. The mean time for 100-meter race was 9.5 seconds for seven contestants (i.e. runners). A contestant, who finished the race with a time of 9.75 seconds, was disqualified from the race. What is new mean time for the remaining six contestants?

$$\begin{aligned}
 \text{total} &= 9.5 \times 7 = 66.5 \text{ seconds} \\
 &66.5 - 9.75 \text{ (DQ)} \\
 &= 56.75 / 6 \text{ contestants} \\
 \text{new mean} &= \boxed{9.45 \text{ seconds}}
 \end{aligned}$$

32. According to Chebyshev's theorem, what is the percentage of employees who travel a distance from home to the office between 25 to 75 miles, if the mean travel distance is 50 miles with standard deviation 15 miles?

$$\begin{aligned}
 P(25 \leq x \leq 75) &= P(|x - 50| \leq 25) \geq 1 - \frac{v^2}{(25)^2} \\
 P(25 \leq x \leq 75) &\geq \underline{0.64} \qquad = 1 - \left(\frac{15}{25}\right)^2 = \underline{0.64}
 \end{aligned}$$

$$\boxed{64\%}$$

33. Use the following to answer questions (a)-(g):

Consider the following five pairs of m and f values:

m	f
6	3

6	3
9	5
7	5
13	6
7	8

$$S. (6-3)^2 \cdot 3 + (9-3)^2 \cdot 5 + (7-3)^2 \cdot 5 + (13-3)^2 \cdot 6 + (7-3)^2 \cdot 8$$

$$\sum (m-3)^2 f = 1015$$

- a. The value of $\sum m$ is: $6+9+7+13+7 = 42$
- b. The value of $\sum mf$ is: $6 \cdot 3 + 9 \cdot 5 + 7 \cdot 5 + 13 \cdot 6 + 7 \cdot 8 = 232$
- c. The value of $\sum m^2$ is: $6^2 + 9^2 + 7^2 + 13^2 + 7^2 = 384$
- d. The value of $\sum f^2$ is: $3^2 + 5^2 + 5^2 + 6^2 + 8^2 = 159$
- e. The value of $\sum m^2 f$ is: $6^2 \cdot 3 + 9^2 \cdot 5 + 7^2 \cdot 5 + 13^2 \cdot 6 + 7^2 \cdot 8 = 2164$
- f. The value of $\sum mf^2$ is: $6 \cdot 3^2 + 9 \cdot 5^2 + 7 \cdot 5^2 + 13 \cdot 6^2 + 7 \cdot 8^2 = 1370$
- g. The value of $\sum (m-3)^2 f$ is: 1015

34. Use the following to answer questions (a)-(f):

The following table gives the cumulative frequency distribution of the commuting time (in minutes) from home to work for a sample of 400 persons selected from a city.

Time (minutes)	f
0 to less than 10	67
0 to less than 20	158
0 to less than 30	223
0 to less than 40	291
0 to less than 50	350
0 to less than 60	400

$$223/400 = .56$$

$$350/400$$

$$291/400 = 0.7275$$

$$.88$$

$$291/400 = 72.75\%$$

$$100 - 72.75 = 27.25$$

$$400 - 158 = 242$$

- a. The sample size is: 400
- b. The percentage of persons who commute for less than 30 minutes, rounded to two decimal places, is: 56%
- c. The cumulative relative frequency of the fourth class, rounded to four decimal places, is: 0.7275
- a. The percentage of persons who commute for 40 or more minutes, rounded to two decimal places, is: 27.25
- e. The percentage of persons who commute for less than 50 minutes, rounded to two decimal places, is: 88%
- f. The number of persons who commute for 20 or more minutes is: 242 ppl

Class Class Rel. Cum. Cum.
 Limit Bound. Freque. Freq. Freq.

3-7	2.5-7.5	2	$2/30 = 0.0667$	2
8-12	7.5-12.5	6	$6/30 = 0.2$	8
13-17	12.5-17.5	8	$8/30 = 0.2667$	16
18-22	17.5-22.5	8	$8/30 = 0.2667$	24
23-27	22.5-27.5	6	$6/30 = 0.2$	30

Total:

30 1.00

Extended Response (30 points for questions 35-36)

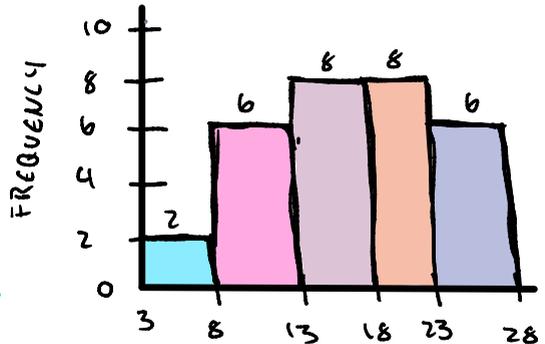
Frequency
 ↑
 Table

35. The following data give the total number of iPods® sold by a mail order company on each of 30 days. Construct a frequency distribution table, draw a histogram, and ogive curve.

(Note: The frequency table must contain the following columns: class limits, class boundaries, frequency, relative frequency, and cumulative frequency)

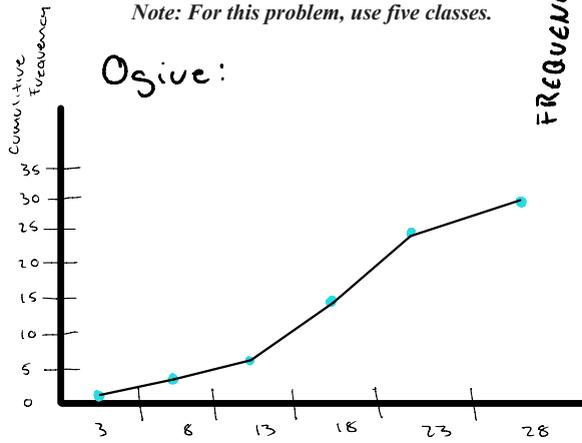
18	25	11	15	9	12	10	5	17	21
22	13	26	16	8	12	19	26	20	26
23	14	19	3	20	16	27	16	21	14

Histogram:



Note: For this problem, use five classes.

Ogive:



36. The following are the average points-per-game of the tallest thirty NBA players for the previous week:

5	32	20	36	40	39	11	17	13	25
19	22	42	7	26	26	13	8	31	31
23	14	37	24	21	35	12	12	37	28

(a) Construct a stem-and-leaf display.

Key: $1 | 1 = 11$

stem	leaf
0	5, 7, 8
1	1, 2, 2, 3, 3, 4, 7, 9
2	0, 1, 2, 3, 4, 5, 6, 6, 8
3	1, 1, 2, 5, 6, 7, 7, 9
4	0, 2

(b) Determine the median and mode of the aforementioned data.

{5, 7, 8, 11, 12, 12, 13, 13, 14, 17, 19, 20, 21, 22, 23, 24, 25, 26, 26, 28,
31, 31, 32, 35, 36, 37, 37, 39, 40, 42}

median: $23 + 24 / 2$
: 23.5

mode: no mode, multiple values have 2

(c) Calculate the mean and the range of the aforementioned data.

mean: $706 / 30 = 23.53$

range: $42 - 5 = 37$

(d) Calculate the standard deviation and variance of the aforementioned data.

11

x	x ²	x	x ²
5	25	24	576
7	49	25	625
8	64	26	676
11	121	26	676
12	144	28	784
12	144	31	961
13	169	31	961
13	169	32	1024
14	196	35	1225
17	289	36	1296
19	361	37	1369
20	400	37	1369
21	441	39	1521
22	484	40	1600
23	529	42	1764

$n = 30$
 $\sum x = 706$
 $\sum x^2 = 19051$

$$s = \sqrt{\frac{\sum x^2 - (\sum x)^2 / n}{n - 1}}$$

$$\sqrt{\frac{19051 - (706)^2 / 30}{30 - 1}}$$

$s = 9.17$

$CV = \frac{s}{\bar{x}} \cdot 100$

$CV = 1.79$