

Biostatistics Exam 1

SECTION I: MULTIPLE CHOICE QUESTIONS (20 MARKS)

1. Fill in the missing words to the quote: "Statistical methods may be described as methods for drawing conclusions about....."

- A. samples
- B. parameters
- C. statistics
- D. mean
- E. populations

2. The..... is the part that is actually examined in order to gather information.

- A. population
- B. explanatory variables
- C. subgroup
- D. sample

3. In an experiment to determine if nutrient X increases the weight of under fives, the following were measured on each of the under fives in the study. One-hundred under fives were involved in the study.

sex, initial weight, weight gain, age group.

Where age group is recorded as (0-28 days, 1-11 months, 12-23 months, 24-35 months, ? 36 months). The scale of measurement of these variables are:

- A. Nominal, ratio, interval, nominal
- B. Nominal, ratio, ratio, nominal
- C. Nominal, ratio, ratio, ordinal
- D. Ordinal, ratio, ratio, ordinal

4. In an immunization study, the following variables were measured on each child:
sex, initial weight(kg), body temperature (0C), weight gain(kg).

The scale of these four variables (in order) are:

- A. Nominal, ratio, ratio, ratio
- B. Nominal, ratio, interval, interval
- C. Ordinal, ratio, interval, ratio
- D. Nominal, ratio, interval, ratio

5. A study was conducted to investigate the effect of a radiation on malignant tumour in cancer patients. As part of the evaluation study, measurement of the size of the tumour was taken. The

following information was recorded for each patient's tumour:

sex (0=female, 1=male),
size of tumour (cm),
age (0=young, 1=adult),
weight (g).

The scale of these variables is:

- A. nominal, ratio, nominal, ratio
- B. nominal, interval, ordinal, ratio
- C. nominal, ratio, ordinal, ratio
- D. ordinal, ratio, nominal, ratio

6. An experiment was performed upon school children to investigate the effect of heamoglobin booster on children's weight increase. The following variables were measured:

gender (0=female, 1=male);
weight (g);
dose of drug X (nil, low, high);
number of children with a certain weight

The typical hb is about 12 g/dl and the weights were rounded to the nearest gram. The number of children was 400. Which of the following is FALSE?

- A. Gender is nominal scale; dose is ordinal scale
- B. Gender is discrete; weight is continuous
- C. Dose is ordinal scale and discrete
- D. Number of children is discrete and is interval scale

7. What is the median of the following set of scores?

18, 6, 12, 10, 14 ?

- A. 10
- B. 14
- C. 18
- D. 12

8. The standard deviation is:

- A. The square root of the variance
- B. A measure of central tendency
- C. An approximate indicator of how numbers vary from the median
- D. The most frequently occurring observation

9. A graph that uses vertical bars to represent data is called a _____.

- A. Line graph
- B. Bar graph
- C. Scatterplot
- D. Vertical graph

10. The goal of _____ is to focus on summarizing and explaining a specific set of data.

- A. Inferential statistics
 - B. Descriptive statistics
 - C. Generalization of results
 - D. Inductive reasoning
11. The most frequently occurring number in a set of values is called the ____.
- A. Mean
 - B. Median
 - C. Mode
 - D. Range
12. As a general rule, the _____ is the best measure of central tendency because it is more precise.
- A. Mean
 - B. Median
 - C. Mode
 - D. Range
13. Focusing on describing or explaining data versus going beyond immediate data and making inferences is the difference between _____.
- A. Central tendency and common tendency
 - B. Mutually exclusive and mutually exhaustive properties
 - C. Descriptive and inferential
 - D. Positive skew and negative skew
14. Why are variance and standard deviation the most popular measures of variability?
- A. They are the most stable and are foundations for more advanced statistical analysis
 - B. They are the most simple to calculate with large data sets
 - C. They provide nominally scaled data
 - D. They provide interval scaled data
15. _____ are used when you want to visually examine the relationship between two quantitative variables.
- A. Bar graphs
 - B. Pie graphs
 - C. Line graphs
 - D. Scatterplots
16. The _____ is often the preferred measure of central tendency if the data are severely skewed.
- A. Mean
 - B. Median
 - C. Mode
 - D. Range
17. Which of the following is a measure of central tendency?
- A. Median
 - B. Variance
 - C. Standard deviation

D. Range

18. What is the median of this set of numbers: 4, 6, 7, 8, 9, 2000?

- A. 7.5
- B. 6
- C. 7
- D. 4

19. Which of the following is interpreted as the percentage of scores in a reference group that falls below a particular raw score?

- A. Standard scores
- B. Percentile rank
- C. Reference group
- D. Nominal scale

20. Which measure of central tendency takes into account the magnitude of scores?

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

SECTION II: LONG ANSWER QUESTIONS

Table below shows six variables of interest on which information was collected during a study on 40 males. Use the table to answer the following questions 1- 5:

First name	age	marital	division	Level of education	weight
Peter	16	1	3	2	31
Denis	17	1	1	2	20
Paul	20	1	3	2	37
Frederick	23	1	1	2	19
Richard	23	1	1	2	23
Henry	24	2	1	2	47
Ssamson	25	3	1	2	25
Clement	26	1	3	2	25
John	26	3	4	2	33
Maurice	27	2	2	2	27
Titus	27	1	2	2	45
Shadrack	27	1	3	2	46
Peter	28	1	1	2	45
Lawrence	30	2	1	2	34
Meshack	30	3	1	2	41
Antipers	30	2	4	1	48
Johnam	31	2	1	2	29
Tom	31	2	1	1	35

Charles	32	1	3	2	30
Joshua	32	2	1	2	60
Maurice	33	2	4	2	30
Moses	33	2	1	2	37
Philip	34	2	2	2	51
John	35	1	3	1	33
Fredrick	36	1	3	2	19
Barrack	37	2	1	2	45
Mickael	38	2	2	2	22
Ssamson	38	4	1	2	43
John	40	2	1	2	26
Joseph	41	3	1	3	13
Ssamwel	41	2	3	3	45
Daniel	42	2	1	2	18
Charles	46	3	2	2	28
Michael	50	3	3	2	30
Paul	50	2	3	2	36
Washington	50	2	3	1	43
Jackson	50	2	2	2	40
Elly	51	2	3	2	35
Lawrence	52	2	2	3	34
Jectone	54	2	1	2	25
Joseph	54	2	3	1	45

1. With each of the variables listed here, two graphical methods are mentioned. Indicate which method is more appropriate. State why one method is more appropriate than the other: (8 marks)

1. Marital status: ogive, bar graph
2. Gender: pie chart; frequency polygon
3. Age: bar graph; histogram
4. Weight: bar graph; histogram

2. Construct a frequency table using a class width of 5 for the age variable (5 marks)

3. What was the most common age and weight? (2 mark)

4. Classify the following variables as either quantitative or qualitative. (5 marks)

1. names of the study participants
2. gender
3. age
4. weight

5. division

5. Find the following (use class width of 5):

1. grouped mean age, grouped median age, grouped modal age, grouped standard deviation
(16 points each)

6. For the grouped weight frequency distribution **5-9; 10-14; 15-19; 20-24; 25-29; 30-34; 35-39; 40-44; 45-49; 50-54; 55-59; 60-64**. Answer the following questions: (8 marks)

1. What is the class interval?
2. What is the class width?
3. What are the lower and upper true class limits for class interval **35-39**?
4. What are the lower and upper class boundaries for class interval: **35-39**?

7. The following are the temperature readings of 10 patients suffering from malaria: 36.5; 36.8; 37.0; 37.5; 37.2; 37.8; 38.5; 39.2; 38.1; 37.3. Using your calculator, calculate the following:

(i) the range (5 pts)

(ii) the mean (5 pts)

(iii) the standard deviation (5 pts)

(iv) the mean temperature lies between what range of standard deviations? (5 pts)

8. A nurse gave the following report in her presentation based on temperatures of 10 patients suffering from malaria in her ward: $37.20C \pm 11.4$.

(i) Compared with group in **question 7 above**, are the temperatures in the second group homogenous or heterogenous?

(ii) Compared with group **question 7 above**, explain the nature of the variation in terms of the standard deviation between **question 7** and **question 8** patient groups. (10 pts)

1. An investigator interested in the weight in kilograms of 18 secondary school students recorded the observations below. The investigator summarized the observations using frequency distribution table shown below. Use the data to calculate grouped mode, median, 25th Percentile and the mean (6 marks).

53 35 67 48 63 42 48 55 33 50 46 45 59 40 47 51 66 53

Weight (kg) of 18 secondary school students

Weight (kg)	Frequency
Class interval	f
30-34	1
35-39	1
40-44	2
45-49	5
50-54	4
55-59	2
60-64	1
65-69	2
Total	18

1. Calculate the median group.
2. Calculate the 25th Percentile and interpret your findings
3. Calculate the grouped mean and interpret your findings using the mean and range.