Multiple-Choice Test – Introduction to Statistics

Instructions: Choose ONE option per question. Do not mark answers in this copy. Suggested time:

30–40 minutes.

1. What is the goal of descriptive statistics?

A) Use a sample to infer population parameters

B) Organize, summarize, and display data using tables, graphs, and summary measures

C) Establish causal relations between variables

2. What is the goal of inferential statistics?

A) Describe observed data with graphs

B) Estimate population parameters and test hypotheses using a sample

C) Compute the median and percentiles

3. Which statement about population and sample is correct?

A) The population is always smaller than the sample

B) The sample is the set of all elements of interest

C) The population is the reference set; the sample is an observed subset

4. What is a representative sample?

A) A convenience sample

B) A sample that reflects population characteristics as closely as possible

C) A sample including only average units

5. Simple random sampling (SRS) is defined as:

A) Each possible sample of the same size has the same chance of selection

B) Each element is chosen from homogeneous strata

C) One unit is selected every k units on the list

6. When is stratified sampling appropriate?

A) When the population is divided into internally heterogeneous groups

B) When entire groups representative of the population must be sampled

C) When homogeneous subgroups exist and coverage of each subgroup is desired

7. Systematic sampling proceeds by:

A) Dividing the population into homogeneous strata

B) Randomly selecting one unit in the first group, then taking every k-th unit

C) Selecting clusters and sampling all units

8. Cluster sampling: which statement is correct?

A) Entire clusters are selected and all (or some) units in chosen clusters are observed

B) Units are independent draws from the same distribution

C) Clusters must be internally homogeneous by design

9. Time-series vs cross-section data: which is true?

A) Time-series observe many units at one point in time

B) Cross-section follows a single unit over time

C) Time-series observe the same variable over time; cross-section observes many units at a

10. Which chart is most appropriate for categorical variables?

A) Histogram

B) Ogive

C) Bar chart or pie chart (and Pareto)

11. What is an ogive?

A) A chart of frequencies by class

B) A plot of cumulative frequencies (or cumulative percentages)

C) A chart for linear relationships

12. Which statement about a histogram is correct?

A) It shows nominal categories with separated bars

B) It represents classed numerical data with adjacent bars

C) It is identical to a bar chart

13. Contingency (cross) tables:

A) Show the relationship between two quantitative variables

B) List frequencies for each combination of two categorical/ordinal variables

C) Are charts for time-series data

14. A scatter plot is used to:

A) Show the trend of a variable over time

B) Show the distribution of one variable into classes

C) Visualize the relationship between two quantitative variables

15. A Pareto chart:

A) Is a histogram of cumulative frequencies

B) Orders categories in descending (or ascending) frequency and often includes a cumulative line

16. When building classes for a frequency distribution:

A) Classes may overlap to avoid gaps

B) Class endpoints must avoid overlap

C) Class widths must always differ across classes

17. A line chart is appropriate when:

A) Comparing frequencies across unordered categories

B) Displaying a variable over time with time on the horizontal axis

C) Showing cumulative percentages by class