

BIO401 - BIOSTATISTICS

FINAL TERM MOST IMPORTANT QUESTIONS

PREPARED BY: ADMIN, BS ZOOLOGY ANNOUNCEMENT

COMPILED BY: SIR MUGHAL

CONTACT# 03132310940

GROUP LINK: [HTTPS://CHAT.WHATSAPP.COM/BJC6WOJZX8O9YAC2ISHIGU](https://chat.whatsapp.com/BJC6WOJZX8O9YAC2ISHIGU)

Q.1 What is Random error?

An error is said to be random when the observed value is consistently and constantly higher or lower than the true value. Biased errors arise from the personal limitations of the observer, the imperfection in the instruments used or some other conditions which control the measurements.

Q.2 What are types of sampling ?

There are two types of sampling methods:

- Probability sampling involves random selection, allowing you to make strong statistical inferences about the whole group.
- Simple random sampling is a completely random method of selecting subjects. These can include assigning numbers to all subjects and then using a random number generator to choose random numbers. The members whose numbers are chosen are included in the sample.
- Stratified Random Sampling involves splitting subjects into mutually exclusive groups and then using simple random sampling to choose members from groups.
- Systematic Sampling means that you choose every “nth” participant from a complete list. For example, you could choose every 10th person listed.
- Cluster Random Sampling is a way to randomly select participants from a list that is too large for simple random sampling. For example, if you wanted to choose 1000 participants from the entire population of the U.S., it is likely impossible to get a complete list of everyone. Instead, the researcher randomly selects areas (i.e. cities or counties) and randomly selects from within those boundaries.
- Multi-Stage Random sampling uses a combination of techniques.
- Non-probability sampling involves non-random selection based on convenience or other criteria, allowing you to easily collect data.
- Convenience sampling is a non-probability sampling technique where samples are selected from the population only because they are conveniently available to the researcher. Researchers choose these samples just because they are easy to recruit, and the researcher did not consider selecting a sample that represents the entire population.

Consecutive sampling:

This non-probability sampling method is very similar to convenience sampling, with a slight variation. Here, the researcher picks a single person or a group as a sample, conducts research over a period, analyzes the results, and then moves on to another subject or group if needed. Consecutive sampling technique gives the researcher a chance to work with many topics and fine-tune his/her research by collecting results that have vital insights.

Quota sampling:

Hypothetically consider, a researcher wants to study the career goals of male and female employees in an organization. There are 500 employees in the organization, also known as the population. To understand better about a population, the researcher will need only a sample, not

the entire population. Further, the researcher is interested in particular strata within the population. Here is where quota sampling helps in dividing the population into strata or groups.

Judgmental or Purposive sampling:

In the judgmental sampling method, researchers select the samples based purely on the researcher's knowledge and credibility. In other words, researchers choose only those people who they deem fit to participate in the research study. Judgmental or purposive sampling is not a scientific method of sampling, and the downside to this sampling technique is that the preconceived notions of a researcher can influence the results. Thus, this research technique involves a high amount of ambiguity.

Snowball sampling:

Snowball sampling helps researchers find a sample when they are difficult to locate. Researchers use this technique when the sample size is small and not easily available. This sampling system works like the referral program. Once the researchers find suitable subjects, he asks them for assistance to seek similar subjects to form a considerably good size sample.

Q.3 What is Parameters and Frame?

In order to draw a random sample from a finite population, the first thing that we need is the complete list of all the elements in our population. This list is technically called the FRAME.

In statistics, a parameter is any measured quantity of a statistical population that summarises or describes an aspect of the population, such as a mean or a standard deviation.

Q.4 Define Quantitative variable & Qualitative variable.

A variable is called a **quantitative variable** when a characteristic can be expressed numerically such as age, weight, income or number of children. On the other hand, if the characteristic is non-numerical such as education, sex, eye-colour, quality, intelligence, poverty, satisfaction, etc. the variable is referred to as a **qualitative variable**. A qualitative characteristic is also called an **attribute**. An individual or an object with such a characteristic can be counted or enumerated after having been assigned to one of the several mutually exclusive classes or categories.

Q.5 What is Window Variable editor?

The data editor window is the default window when you run SPSS. The data worksheet works just like a spreadsheet, where a column represents a variable and a row represents a case or an observation.

Q.6 When Skewness was discovered?

Between the end of the nineteenth century and the beginning of the twentieth century, Pearson, Karl studied large sets of data which sometimes deviated significantly from normality and exhibited considerable skewness.

Q.7 What is Dispersion?

In statistics, dispersion is the extent to which a distribution is stretched or squeezed. Common examples of measures of statistical dispersion are the variance, standard deviation, and interquartile range. For instance, when the variance of data in a set is large, the data is widely scattered.

Q.8 What is Relative measure of dispersion?

Relative measures of dispersion are measures of the variance of a range of values regardless of its unit of measure. This means that the spread of two ranges of values with different measures can be compared directly with relative measures of dispersion.

Q.9 What is Regression analysis?

Regression analysis is a powerful statistical method that allows you to examine the relationship between two or more variables of interest. While there are many types of regression analysis, at their core they all examine the influence of one or more independent variables on a dependent variable.

Q.10 What is Data scale?

Scales of measurement in research and statistics are the different ways in which variables are defined and grouped into different categories. Sometimes called the level of measurement, it describes the nature of the values assigned to the variables in a data set. There are four types of data scale:

- i. Nominal
- ii. Ordinal
- iii. Interval
- iv. Ratio

Q.11 What is the difference between quota and cluster sampling?

Quota Sampling:-

In this type of sampling, the selection of the sampling unit from the population is no longer dictated by chance. A sampling frame is not used at all, and the choice of the actual sample units to be interviewed is left to the discretion of the interviewer. However, the interviewer is restricted by quota controls.

Cluster Sampling:-

Cluster sampling is a probability sampling method in which you divide a population into clusters, such as districts or schools, and then randomly select some of these clusters as your sample. The clusters should ideally each be mini-representations of the population as a whole.

Q.12 Explain Range rule of thumb.

The range rule of thumb is a handy method of estimating the range from the standard deviation. It tells us that the range is generally about four times the standard deviation. So if your standard deviation is 2, you might guess that your range is about eight.

Q.13 Write general and specific rule for dependent and independent events.

Independent Events:

Two events are independent if the occurrence of one does not change the probability of the other occurring. An example would be rolling a 2 on a dice and flipping a head on a coin. Rolling the 2 does not affect the probability of flipping the head.

Dependent Events:

Two events are dependent when the outcome of the first event influences the outcome of the second event. The probability of two dependent events is the product of the probability of X and the probability of Y AFTER X occurs.

Q.14 What are Outliers?

An outlier is an observation that lies an abnormal distance from other values in a random sample from a population. A value that "lies outside" (is much smaller or larger than) most of the other values in a set of data. For example in the scores 25,29,3,32,85,33,27,28 both 3 and 85 are "outliers".

Q.15 What are Distribution types?

There are two types of probability distribution which are used for different purposes and various types of the data generation process. Normal or Cumulative Probability Distribution. Binomial or Discrete Probability Distribution.

Q.16 Define Correlation.

Correlation is a measure of the strength or the degree of relationship between two random variables. A numerical measure of the strength of the linear relationship between two random variables X and Y is known as Pearson's Product-Moment Coefficient of Correlation. Correlation can be linear, positive, perfect positive, negative, perfect negative and there can be no correlation at all.

Q.17 How many shapes of distribution are there?

Generally there are two basic types of distribution. Distribution may be symmetric or asymmetric but in advanced statistics, there are seven types of distribution. They are:

- i. Symmetric
- ii. Bell shaped
- iii. U shaped
- iv. Skewed
- v. Positively skewed
- vi. Negatively skewed
- vii. Uniform

Q.18 Name types of probability sampling.

- Simple random sampling
- Stratified Random Sampling
- Systematic Sampling
- Cluster Random Sampling
- Multi-Stage Random sampling

Q.19 What is Regression equation?

A regression equation is used in stats to find out what relationship, if any, exists between sets of data. For example, if you measure a child's height every year you might find that they grow about 3 inches a year. That trend (growing three inches a year) can be modeled with a regression equation.

$$Y_i = f(X_i, \beta) + e_i$$

Y_i = dependent variable

f = function

X_i = independent variable

B = unknown parameters

e_i = error terms

Q.20 What are categories of diagnostic test?

There are many different types of tests in statistics like t-test, Z-test, chi-square test, anova test, binomial test, one sample median test etc. Parametric tests are used if the data is normally distributed .

Q.21 Write field names where statistics is used.

Not only in the field of natural and physical science but statistics has made much impact on social sciences and is of immense use in these also. These include subjects like History, Political Science, Education, Psychology, Sociology and Logic etc. Even in mathematics, many theories have been drawn from statistics.

Q.22 What is Data and its resources?

Data is basically divided into two categories.

- i. Primary Data
- ii. Secondary Data

Primary data can further be divided into:

- a. Interview method
- b. Survey method
- c. Observation method
- d. Experimental method

Secondary data can further be classified as:

- a. Internal source
- b. External source

Q.23 What is primary and secondary data?

PRIMARY DATA

Primary data is a type of data that is collected by researchers directly from main sources through interviews, surveys, experiments, etc. Primary data are usually collected from the source—where the data originally originates from and are regarded as the best kind of data in research.

SECONDARY DATA

Secondary data is the data that has already been collected through primary sources and made readily available for researchers to use for their own research. It is a type of data that has already been collected in the past.

A researcher may have collected the data for a particular project, then made it available to be used by another researcher. The data may also have been collected for general use with no specific research purpose like in the case of the national census.

Q.24 What is Percentile?

In statistics, a k-th percentile is a score below which a given percentage k of scores in its frequency distribution falls or a score at or below which a given percentage falls. For example, the 50th percentile is the score below which or at or below which 50% of the scores in the distribution may be found.

Q.25 What are different types of graphs?

Types of Graphs and Charts

- Bar Chart/Graph.
- Pie Chart.
- Line Graph or Chart.
- Histogram Chart.
- Area Chart.
- Dot Graph or Plot.
- Scatter Plot.
- Bubble Chart.

Q.26 What is regression and correlation?

Regression is a statistical method used in finance, investing, and other disciplines that attempts to determine the strength and character of the relationship between one dependent variable (usually denoted by Y) and a series of other variables (known as independent variables).

Correlation is a statistical measure that expresses the extent to which two variables are linearly related (meaning they change together at a constant rate). It's a common tool for describing simple relationships without making a statement about cause and effect.

Q.27 Define Discrete and Continuous variable.

A discrete variable is a variable whose value is obtained by counting. Examples: number of students present. number of red marbles in a jar. number of heads when flipping three coins.

Continuous variables are numeric variables that have an infinite number of values between any two values. A continuous variable can be numeric or date/time. For example, the length of a part or the date and time a payment is received.

Q.28 What is pie chart?

A pie chart, sometimes called a circle chart, is a way of summarizing a set of nominal data or displaying the different values of a given variable (e.g. percentage distribution). This type of chart is a circle divided into a series of segments. Each segment represents a particular category.

Q.29 How to make box and whisker plot?

In a box plot, we draw a box from the first quartile to the third quartile. A vertical line goes through the box at the median. The whiskers go from each quartile to the minimum or maximum.

Q.30 Write two regression equation.

The equation has the form $Y = a + bX$, where Y is the dependent variable (that's the variable that goes on the Y axis), X is the independent variable (i.e. it is plotted on the X axis), b is the slope of the line and a is the y-intercept.