|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sl. No.** | **Question** | **M** | **BT** | **CO** |
| 1 | Ordinal, ratio | 1 | 2 | 1 |
| 2 | (Measures, Consistency) | 2 | 2 | 2 |
| 3 | ( Accuracy, Consistency) | 2 | 2 | 1 |
| 4 | (Research) | 1 | 2 | 2 |
| 5 | (Instruments) | 1 | 1 | 1 |
| 6 | (Graphs, Summary statistics) | 1 | 1 | 1 |
| 7 | (Probability) | 1 | 1 | 1 |
| 8 | (Parameters) | 1 | 1 | 1 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Marks Distribution** | **Particulars** | | **CO1** | **CO2** | **CO3** | **CO4** | **L1** | **L2** | **L3** | **L4** | **L5** | **L6** |
| **Quiz** | **Max Marks** | 6 | 4 | - | - | 4 | 6 |  | - | - | - |

Scheme and Solution

1(a)

**1. Nominal Scale:**

* Definition: The nominal scale classifies data into distinct categories without any order or ranking. Each category is unique and mutually exclusive.
* Example: Types of fruits (apple, banana, orange). Each fruit type is a separate category without a specific order.
* Data Analysis: Nominal scales are used for categorizing and counting occurrences. Statistical analysis is limited to frequency counts and mode.

**2. Ordinal Scale:**

* Definition: The ordinal scale classifies data into categories with a meaningful order or ranking but does not specify the magnitude of differences between categories.
* Example: Customer satisfaction ratings (poor, fair, good, excellent). The categories have a clear order but do not quantify the difference between them.
* Data Analysis: Ordinal scales allow for ranking and median calculations. Analysis can include determining the mode, median, and non-parametric tests.

**3. Interval Scale:**

* Definition: The interval scale includes ordered categories with equal intervals between values but lacks a true zero point. It allows for the measurement of the size of the differences between values.
* Example: Temperature in Celsius (10°C, 20°C, 30°C). The differences between temperature readings are meaningful, but there is no absolute zero.
* Data Analysis: Interval scales allow for the use of mean, standard deviation, and various parametric statistical tests. However, ratios are not meaningful (e.g., 20°C is not “twice as hot” as 10°C).

**4. Ratio Scale:**

* Definition: The ratio scale possesses all the features of the interval scale, with the addition of a true zero point. This allows for the measurement of absolute quantities and the calculation of meaningful ratios.
* Example: Weight (10 kg, 20 kg, 30 kg). Zero weight means the absence of weight, and ratios like “30 kg is twice as heavy as 15 kg” are meaningful.
* Data Analysis: Ratio scales support a full range of statistical analyses, including mean, median, mode, and all parametric tests. Ratios and proportions are meaningful.

1(b)

Differences Between Probability and Non-Probability Sampling Methods:

**1. Probability Sampling:**

* Definition: In probability sampling, every member of the population has a known and non-zero chance of being selected. This method uses random selection to ensure that the sample is representative of the population.
* Types:
* Simple Random Sampling: Every member of the population has an equal chance of being selected.
* Stratified Sampling: The population is divided into subgroups (strata) based on specific characteristics, and random samples are taken from each stratum.
* Cluster Sampling: The population is divided into clusters (usually geographically), and a random sample of clusters is selected. All members within the chosen clusters are then surveyed.
* Systematic Sampling: Every nth member of the population is selected after a random starting point.

**2. Non-Probability Sampling:**

* Definition: In non-probability sampling, not all members of the population have a known or equal chance of being selected. Selection is based on non-random criteria, which may introduce bias.
* Types:
* Convenience Sampling: Sampling is done based on ease of access to the participants. For example, surveying people who are readily available.
* Judgmental (Purposive) Sampling: The researcher selects participants based on their judgment and specific criteria relevant to the study.
* Snowball Sampling: Existing study subjects recruit future subjects from their acquaintances. This method is often used for hard-to-reach populations.
* Quota Sampling: The researcher ensures that certain characteristics of the sample match the characteristics of the population, but selection within each group is non-random.

2(a)

**1. Surveys:**

* Definition: Surveys involve the collection of data through structured questionnaires or forms that participants complete on their own or through electronic means.
* Advantages:
  + Efficiency: Can reach a large number of participants quickly and cost-effectively.
  + Standardization: Provides uniform questions and responses, which facilitates easier data analysis and comparison.
  + Anonymity: Allows for responses to be given anonymously, which can reduce social desirability bias.
* Disadvantages:
  + Limited Depth: May not capture detailed or nuanced information due to the predefined structure of the questions.
  + Response Quality: Dependent on the participant’s ability to interpret and understand questions accurately. Misunderstandings can lead to inaccurate responses.
  + Engagement: Can suffer from low response rates or incomplete surveys, impacting the quality and reliability of the data.

2. Interviews:

* Definition: Interviews involve direct, personal interactions between the interviewer and the participant, where questions are asked verbally, and responses are recorded.
* Advantages:
* Depth of Information: Allows for in-depth exploration of responses and follow-up questions to clarify or expand on answers.
* Flexibility: The interviewer can adjust questions based on responses, probe deeper into certain topics, and explore new avenues of inquiry.
* Immediate Feedback: Enables real-time clarification of questions and understanding of responses, which can improve data accuracy.
* Disadvantages:
* Time-Consuming: Typically requires more time and resources to conduct and analyze, especially if done in person.
* Interviewer Bias: The interviewer’s presence and behavior can influence responses, potentially introducing bias.
* Cost: More expensive due to the need for trained interviewers and potentially longer data collection periods.

2(b)

Measurement error refers to the difference between the actual value of a variable and the value obtained through measurement. It represents the inaccuracies or discrepancies that can arise when collecting or recording data. Measurement error can affect the validity and reliability of research findings, making it crucial to identify and minimize these errors.

Types of Measurement Errors

1. Systematic Error (Bias):

- Definition: Systematic errors are consistent, repeatable errors that occur due to flaws in the measurement system or procedure. They shift all measurements in one direction (either higher or lower) from the true value.

- Examples:

- Instrument Bias: A scale that consistently reads 2 pounds heavier than the actual weight.

- Observer Bias: A researcher interpreting results in a way that favors their hypothesis.

2. Random Error:

- Definition: Random errors are unpredictable and vary from one measurement to another. They arise from unknown or uncontrollable factors and tend to average out over multiple measurements.

- Examples:

- Environmental Factors: Variations in temperature or humidity affecting measurements.

- Human Factors: Minor inconsistencies in measurement technique or judgment.

3. Measurement Error due to Instrumentation:

- Definition: Errors arising from the tools or instruments used for measurement.

- Examples:

- Calibration Issues: An instrument not calibrated properly, leading to incorrect readings.

- Wear and Tear: Aging equipment that no longer provides accurate measurements.

4. Measurement Error due to Sampling:

- Definition: Errors occurring due to the way samples are selected or due to sample variability.

- Examples:

- Sampling Bias: If a sample is not representative of the population (e.g., surveying only urban residents for a national health study).

- Sampling Variability: Differences in measurements obtained from different samples drawn from the same population.

5. Measurement Error due to Data Entry:

- Definition: Errors that occur during the recording or entering of data into databases or spreadsheets.

- Examples:

- Transcription Errors: Incorrectly typing numbers or values into a dataset.

- Data Coding Errors: Misclassifying data into incorrect categories.

6. Measurement Error due to Response Bias:

- Definition: Errors that occur when respondents provide inaccurate answers due to various biases or influences.

- Examples:

- Social Desirability Bias: Respondents providing answers they think are more socially acceptable.

- Recall Bias: Respondents having difficulty remembering or accurately reporting past events.

Reducing Measurement Error

1. Calibration and Maintenance: Regularly calibrate and maintain instruments to ensure they function correctly.

2. Standardization: Use standardized procedures and protocols to minimize variability.

3. Training: Train personnel thoroughly to ensure consistency in measurement techniques.

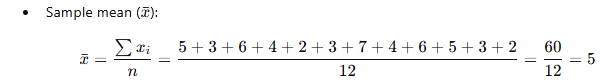
4. Pilot Testing: Conduct pilot tests to identify and address potential sources of error before the main study.

5. Cross-Validation: Use multiple methods or instruments to validate measurements and reduce reliance on a single source of data.

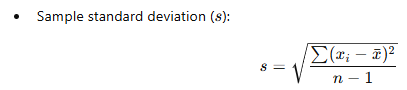
6. Randomization: Use random sampling and random assignment to minimize selection bias and ensure representative samples.

By understanding and addressing these different types of measurement errors, researchers can improve the accuracy and reliability of their data, leading to more valid and generalizable conclusions.

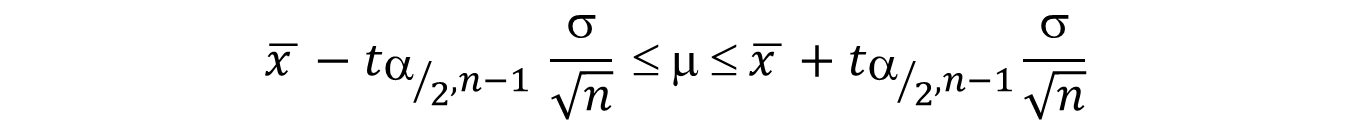
3)



= 4.166667



= 1.642245



5

3.95656 6.04344

The 95% confidence interval for the average number of books purchased per customer visit is approximately [3.95565,6.04344]

4

 **State the Hypotheses:**

* **Null Hypothesis (H0​)**: The mean cholesterol level of the patients is equal to the general population mean (μ=200 mg/dL).
* **Alternative Hypothesis (HA​)**: The mean cholesterol level of the patients is different from the general population mean (μ≠200 mg/dL).

 **Calculate the Sample Mean and Sample Standard Deviation:**

* **Sample Data**: 170, 185, 195, 210, 220, 190, 175, 205
* **Sample Size (n)**: 8
* **Sample Mean ()**:

= 193.75

**Sample Standard Deviation (s)**:

 = 17. 26888

Perform the t-Test:



**Calculate the t-Statistic**:

= -1.14313

**Degrees of Freedom (df)**:

df=n−1=8−1=7

**Compare the t-Statistic to the Critical t-Value**:

The calculated t-statistic is -0.82.

The critical t-values are approximately ± 2.364

Since -1.14313 is within the range −2.364-to +2.364, we fail to reject the null hypothesis.

### Summary:

There is insufficient evidence to conclude that the mean cholesterol level of the sample is significantly different from the general population mean of 200 mg/dL. The test indicates that the sample mean is not significantly different from the population mean at the 5% significance level.

5.

**State the Hypotheses:**

* **Null Hypothesis (H0)** The median compressive strength is at least 1500 psi (Median≥1500 psi).
* **Alternative Hypothesis (HA)**: The median compressive strength is less than 1500 psi (Median<1500 psi).

Rank the Data Relative to the Hypothesized Median (1500 psi):

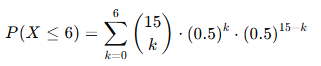
Data values below 1500 psi: 1470, 1490, 1400, 1450, 1480, 1430 (6 values)

Data values above 1500 psi: 1520, 1550, 1580, 1530, 1600, 1570, 1620, 1510, 1590 (9 values)

Perform the Sign Test:

Calculate the probability of obtaining 6 or fewer successes (values below 1500) out of 15 trials.

Using the binomial test (or a binomial distribution calculator):



From binomial distribution tables or calculators, we find that the p-value for k≤6 is approximately 0.295.

Compare the p-value to the Significance Level:

The significance level (α) is 0.05.

Since the p-value (0.295) is greater than 0.05, we fail to reject the null hypothesis.

here is insufficient evidence at the 5% significance level to conclude that the median compressive strength of the blocks is less than 1500 psi. Thus, the test indicates that the median compressive strength is likely to be at least 1500 psi, as required.