Unit 4

1. Levels Of Scale Measurement (Nominal Ordinal Interval Ratio)

1. Nominal Scale

- **Definition:** Nominal scales are the most basic level of measurement.
- **Purpose:** Used for identification or classification purposes only; no quantities are represented.
- **Representation:** Can use letters or numbers without implying different quantities.
- **Common Instances:** Sports uniform numbers, school bus numbers, and other nominal numbering systems.
- **Example1:** The number 7 on a horse's colors in betting, serves as a label, not indicating the horse's speed or size.
- **Examples2:** Barq's Root Beer experiment, labeling drinks A, B, or C without indicating the sweetener type.

2. Ordinal Scale

- **Definition:**
  - Allows arrangement in order based on the quantity of a concept.
  - Often referred to as a rank order scale.
- **Purpose:**
  - Ranks items based on how much of a concept they possess.
  - Commonly used in preference ranking and classifying without precise measurement.
- **Representation:**
  - Assigns ordered numbers to represent the order of items.
  - The assigned numbers indicate relative ranking, not precise quantities.
- **Common Instances:** Class rankings in high school. Ranking preferences or performance.
- **Examples:**
  - Ranking students based on class performance.
  - Ordering preferences from most to least preferred.
  - Assigning positions like "win," "place," and "show" in a horse race

3. Interval Scale:

1. **Definition:**
   - Interval scales have nominal and ordinal properties but also measure differences in quantities.
   - They capture information about the degree of difference between observations.
2. **Purpose:**
   - Provide a precise measure of the difference between values in a scale.
   - Offer more information than ordinal measures, specifying not just order but also the extent of difference.
3. **Representation:**
   - Represented by a numeric scale where equal intervals indicate equal differences.
4. **Common Instances:**
   - Used in scenarios like grading systems where the gap between grades is consistent.
   - Applied in temperature scales like Fahrenheit.
5. **Examples:**
   - A professor grading term papers from 1.0 to 20.0, indicating both order and the extent of difference.
   - Temperature scale examples, e.g., June 6 at 80°F and December 7 at 40°F, showing a 40°F difference but not implying twice the coldness.

4. Ratio Scale

1. **Definition:**
   - Ratio scales encompass all properties of interval scales and additionally represent absolute quantities.
   - They provide iconic measurement, where zero indicates an absence of the measured concept.
2. **Purpose:**
   - Offer the highest level of measurement, allowing for precise quantitative comparisons.
   - Enable the representation of absolute values, distinguishing them from relative measurements.
3. **Representation:**
   - Possess both ordinal and interval properties, but crucially include a true zero point.
   - Zero on a ratio scale signifies a meaningful absence of the measured concept.
4. **Common Instances:**
   - Applied in scenarios where true zero points are meaningful, such as measuring quantities like money, age, weight, etc.
   - Utilized in economic valuation, where items like watches are priced and compared.
5. **Examples:**
   - Online auction prices of watches, demonstrating ordinal, interval, and ratio conclusions based on their values.
   - Timing information in a horse race, where knowing the ratio data allows determination of interval and ordinal information.

2. Ranking Rating Sorting

1. Ranking:

- **Definition**: Involves arranging a small number of items based on overall preference or a specific characteristic.
- **Purpose:** Determines the order of preference or importance among a set of items.
- **Example:** Respondents might rank order stores, brands, feelings, or objects based on overall preference.

2. Rating:

- **Definition:** Requires respondents to estimate the magnitude or extent of a characteristic using attitudinal or cognitive scales.
- **Purpose:** Quantifies the degree or extent to which a particular characteristic exists.
- **Example:** Involves marking responses on attitudinal or cognitive scales to indicate one's position regarding a characteristic.

3. Sorting:

- **Definition:** Involves classifying concepts by placing them into groups, typically presented on cards.
- **Purpose:** Classifies concepts into groups, revealing patterns and preferences.
- **Example:** Respondents classify concepts printed on cards into groups based on similarities or other criteria.

Unit 4

3. Coding And Reverse Coding

1. **Coding in Research:**

   - **Definition:** Coding in research refers to the process of assigning labels or categories to data to identify patterns and themes.
   - **Purpose:** It helps organize and analyze qualitative or quantitative data, facilitating the extraction of meaningful insights.

2. **Reverse Coding in Research:**

   - **Definition:** Reverse coding involves assigning values to items in a way that reverses the original scale, often done to counteract response biases or wording effects.
   - **Example:** If the original scale is from 1 to 5, reverse coding would involve subtracting respondents' scores from 6, where 1 becomes 5, 2 becomes 4, and so on[1].
   - **Application:** Commonly used in surveys and questionnaires where some questions are negatively framed.

Unit 4

4. Criteria for good measurement

1. Reliability

- **Definition:** Reliability is the consistency or repeatability of a measure.
- **Purpose:** Ensures dependable and repeatable measurements in research, minimizing errors.
- **Example:** In a survey, if a question can be interpreted differently by respondents, it becomes unreliable. Even clear questions may yield unreliable answers based on interpretation.

5 Types of Reliability:

1. **Test-Retest Reliability**
   - **Definition:** Administer the same test twice to a group over time.
   - **Purpose:** Assessing the stability of test scores over time.
   - **Example:** Assessing student learning stability through repeated psychology tests.
2. **Split-Half Reliability**
   - **Definition**: Randomly divide items measuring the same construct into two sets.
   - **Purpose:** Assessing internal consistency of a measure.
   - **Example:** Administer the split instrument, calculate scores for each half using the odd-even strategy.
3. **Inter-Rater Reliability**
   - **Definition:** It measures the consistency of judgments from multiple raters..
   - **Purpose:** Subjective assessments where interpretations may vary.
   - **Example:** Art portfolio evaluations by different judges.
     Ensuring consistent assessment of subjective constructs or skills.
4. **Parallel-Forms Reliability**
   - **Definition:** Administering different versions of an assessment tool to the same group of individuals.
   - **purpose:** Evaluating the consistency of results across alternate versions of the assessment.
   - **Example:** If two versions of a test are given to the same group, the correlation of scores helps assess the reliability of the measurement.
5. **Coefficient Alpha (α)**
   - **Definition:** Coefficient α is a widely used estimate for assessing the reliability of multiple-item scales.
   - It calculates the average of all potential split-half reliabilities for a scale, indicating internal consistency.
   - **Purpose:**
   - Assessing the reliability of multiple-item scales in research and measurement.
   - Evaluating whether different items in a scale converge, indicating consistency.
   - **Example:**
   - Researchers commonly use coefficient α as the primary indicator of a scale's quality, despite not addressing validity.
   - The range of coefficient alpha values is from 0 (no consistency) to 1 (complete consistency)

Unit 4

2. Validity

Definition:

- Validity refers to whether a measure accurately assesses what it intends to measure.
- An unreliable measure is also invalid, as it fails to consistently capture the intended phenomenon.

Purpose: - Indicates the soundness of research by ensuring the accuracy of measurements.

- Applies to both research design and methods, emphasizing the credibility of findings.

Example: If income is measured using years of education, it yields consistent but invalid results, as education doesn't directly represent income.

❖ **2 Main Types of Validity**

1. **Internal Validity**: Focuses on the accuracy of the measurement and test itself.
2. **External Validity**: Concerns the ability to generalize findings to the target population.

10 Types of Validity

1. **Face Validity**
   - **Definition:** Assess how well a test appears to measure what it claims.
   - **Purpose:** Engaging stakeholders and motivating participants.
   - **Example:** Ensure questions in an art appreciation test relate to different art components, avoiding unrelated topics like historical periods.

2. **Predictive Validity**
   - **Definition:** Evaluate if a new measure predicts outcomes as the old one did.
   - **Purpose:** Testing if a measure can predict theoretical relationships.
   - **Example:** Correlate math ability measure with engineers' salaries to validate its predictive power.

3. **Criterion-Related Validity**
   - **Definition:** Correlate parts of a compound measure with the variable it relates to.
   - **Purpose:** Predicting future or current performance based on current scores.
   - **Example:** Correlate a physics program's measure with a standardized ability test to assess cumulative student learning.

4. **Content Validity**
   - **Definition:** Check if the operationalization aligns with the relevant content domain.
   - **Purpose:** Ensuring the test represents the entire content of the behavior or construct.
   - **Example:** Test on Bangladesh Geography should cover diverse aspects, not limited to Australia.

5. **Convergent Validity**
   - **Definition:** Assess consistency between two measures of the same thing.
   - **Purpose:** Ensuring different measures of the same concept are consistent.
   - **Example:** Correlate scores on an arithmetic test with other math ability tests.

6. **Concurrent Validity**
   - **Definition:** Compare scores on a new test with an established one at the same time.
   - **Purpose:** Distinguishing between groups the measure should theoretically differentiate.
   - **Example:** Assessing a new measure of empowerment by comparing scores between migrant farm workers and farm owners.

7. **Construct Validity**
   - **Definition:** Use a panel of experts to ensure the measure reflects the intended construct.
   - **Purpose:** Confirming the measure measures the intended hypothetical construct.
   - **Example:** Validate a measure of anxiety by correlating it with the threat of an electric shock.

8. **Formative Validity**
   - **Definition:** Assess how well a measure provides information to improve a program.
   - **Purpose:** Improving courses or programs based on assessment results.
   - **Example:** Design a history rubric to identify areas where students lack knowledge, like the Civil Rights Movement

9. **Sampling Validity**
   - **Definition:** Ensure the measure covers a broad range of areas within the concept under study.
   - **Purpose:** Ensuring a representative sample of the content area.
   - **Example:** Designing an assessment in the theatre department that covers various aspects, not just acting.

10. **Discriminant Validity**
    - **Definition:** Examine the degree to which the operationalization differs from other measures.
    - **Purpose:** Showing that the measure is not similar to theoretically distinct measures.
    - **Example:** Demonstrating the discriminant validity of a Head Start program by showing its differences from other early childhood program.

3. Sensitivity

- **Definition:** Sensitivity in measurement refers to an instrument's capacity to accurately capture variations in a concept, crucial in assessing changes in attitudes or hypothetical constructs.
- **Purpose:** It enables precise measurement of subtle changes that a dichotomous response category may overlook. For instance, enhancing a scale with more nuanced response options, like "strongly agree" to "strongly disagree," increases sensitivity.
- **Example:** Improving the sensitivity of a scale can involve adding response points or expanding the scale with additional items. Composite measures, incorporating various items or questions, offer greater sensitivity compared to single-item scales

5. Purpose of sampling

- The basic purpose of sampling is to provide an estimate of the population parameter and to test the hypothesis.
  - *Advantages of sampling are –*

    i. Save time and money
    ii. Enable collection of comprehensive data
    iii. Enable more accurate measurement as it conducted by trained and experienced investigators
    iv. Sampling remains the only way when population contains infinitely many members
    v. In certain situation, sampling is the only way of data collection.
       For example, in testing the pathological status of blood, boiling status of rice, etc
    vi. It provides a valid estimation of sampling error

8C.3 Purpose of Sampling

The basic purpose of sampling is to provide an estimate of the population parameter and to test the hypothesis. Advantages of sampling are –

- ✓ Save time and money
- ✓ Enable collection of comprehensive data
- ✓ Enable more accurate measurement as it conducted by trained and experienced investigators
- ✓ Sampling remains the only way when population contains infinitely many members
- ✓ In certain situation, sampling is the only way of data collection. For example, in testing the pathological status of blood, boiling status of rice, etc
- ✓ It provides a valid estimation of sampling error

6. Probability vs non probability sampling

Probability Sampling	Non-Probability Sampling
Each unit of the population has an equal and positive chance of being selected.	Some elements of the population have no chance of selection or the probability of selection cannot be accurately determined.
Sample truly represents the overall population.	Sample may not be representative of the overall population.
Examples: Simple Random Sample, Systematic Random Sample, Stratified Random Sample, Cluster/Multistage Sample.	Examples: Accidental/Convenience Sampling, Quota Sampling, Judgment/Subjective/Purposive Sampling, Snowball Sampling.
Allows estimation of sampling errors.	Does not allow estimation of sampling errors.
More suitable for scientific research and generalizing findings to the population.	More suitable for exploratory research, pilot testing, or when it is difficult to undertake a stratified sample.

7. Relationship between sample size and random error

1. **Increasing Sample Size Reduces Error:** Larger sample sizes lead to a reduction in random sampling error, instilling confidence in the accuracy of collected data[1][4].
2. **Diminishing Returns in Error Reduction:** The reduction in sampling error diminishes at a decreasing rate as the sample size increases. Doubling the sample size may result in only a small reduction in error due to the law of diminishing returns[1].
3. **Inverse Proportion to Square Root:** Random sampling error is inversely proportional to the square root of the sample size. As the sample size increases, the error decreases, but at a slower rate[6].
4. **Crucial Role of Optimal Sample Size:** Determining the optimal sample size is crucial, considering factors like population heterogeneity, acceptable error range, and confidence level[4].
5. **Magnitude of Error and Decision Importance:** The magnitude of error, represented by the confidence interval, indicates the precision needed in estimates. Decision importance in terms of profitability influences the acceptable error range[2].

Overall, larger sample sizes generally result in lower random sampling error, but the rate of reduction in error diminishes as the sample size increases.

8. Normal Distribution

- **Definition:** The normal distribution is a common probability distribution in statistics, often represented by the bell-shaped normal curve.
- **Purpose:** It provides a mathematical and theoretical framework for understanding the distribution of data. The standardized normal distribution, with a mean of 0 and a standard deviation of 1, is especially valuable in inferential statistics.
- **Utility:** Valuable for finding probabilities of various portions under the curve.
- **Theoretical Nature:** The standardized normal distribution is purely theoretical but highly valuable in inferential statistics.
- **Transformation:** Statisticians transform data from observed normal distributions to the standardized normal curve for probability calculations.
- **Example:** An example of a normal distribution is seen in the distribution of IQ scores, where almost all values fall within ±3 standard deviations from the mean.

❖ ***Normal Distribution characteristics:***

i. Symmetrical about its mean, with equal tails on both sides.
ii. The mode, mean, and median coincide, forming a vertical symmetrical line.
iii. It's a continuous distribution with an infinite number of cases, and the area under the curve has a probability density of 1.0.
iv. Mean of 0 and standard deviation of 1.

EXHIBIT 17.8
Normal Distribution: Distribution of Intelligence Quotient (IQ) Scores

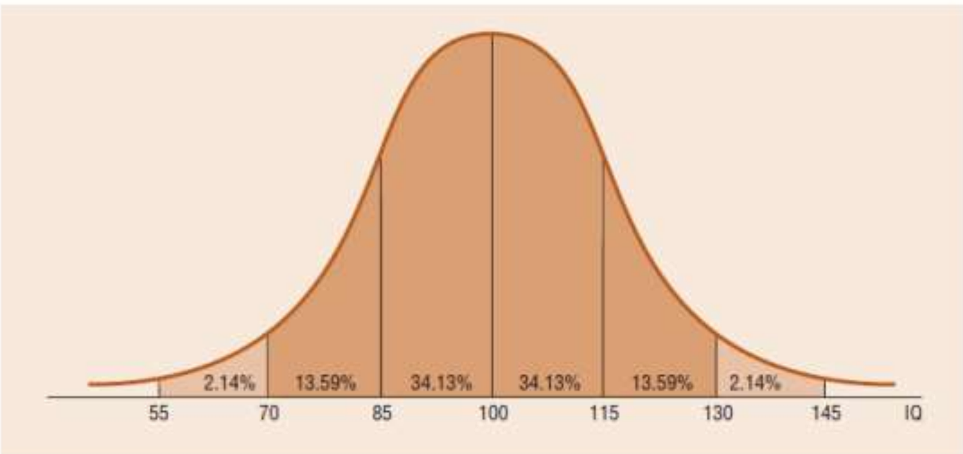

EXHIBIT 17.9
Standardized Normal Distribution