

Midterm: Conceptual Definitions

Below are the key concepts we have covered in class. This list is not exhaustive and is intended to complement, not replace, lecture material.

Lecture 01: Introduction

Statistics: A set of tools and techniques that are used for describing, organizing, and interpreting information or data

Descriptive Statistics: Used to organize or describe the characteristics of a dataset

Inferential Statistics: Used to make inferences (reach a conclusion) about the data

Lecture 02: Central Tendency and Variability

Sample: A subset of a population

Mean: The sum of a set of scores divided by the total number of scores in the set

Median: The point at which half (50%) of the values are above and half (50%) of the values are below

Mode: The value in a distribution of data that occurs most frequently

Scales of measurement: Describes the nature of the information contained in a given set of data

Measures of variability: Describe how scores in a given dataset differ from one another (e.g., the spread or clustering of points)

Range: The difference between the lowest and highest values in a dataset

Sum of squared deviations: The squared differences of each observation from the overall mean of a dataset

Standard Deviation: The standard (or typical) amount that scores deviate from the mean

Variance: The averaged squared deviation from the mean

Lecture 04: Correlation, Validity, Reliability

Correlational designs: Examine the extent to which two variables are associated

Independent variable: The variable that is hypothesized to have an effect on some outcome of interest

Dependent variable: The outcome of interest that the independent variable might have an effect on

Reliability: The overall consistency of a measure

Validity: Refers to how accurately a method measures what it is intended to measure

True score: The true value that we are trying to measure

Observed Score: The value that we actually measure

Test-Retest Reliability: Used to determine whether a test (or scale) is reliable over time

Parallel Forms Reliability: Used to examine the equivalence or similarity between two forms of the same test (or scale). Measured as the correlation r_{xy} between scores for the same individuals on Form A and Form B

Internal consistency: Used to determine whether the items on a test (or scale) are consistent with each other. Used to determine whether the items on a test (or scale) are consistent with each other, typically measured with Cronbach's α

Test-Retest Reliability: Measured as the correlation r_{xy} between scores on a measure at Time 1 and the same measure at Time 2

Interrater Reliability: The degree of agreement among independent observers who rate, code, or agree on their judgments of an outcome of interest

Content validity: The extent to which the items on a test are fairly representative of the entire domain the test seeks to measure

Criterion validity: Measures how well one measure predicts an outcome for another measure

Concurrent Criterion Validity: The degree to which scores on the measure are associated with scores on another measure taken at the same time

Predictive Validity: The degree to which test scores accurately predict scores on a criterion measure

Construct Validity: The degree to which a test, scale, or assessment measures the construct it claims to measure

Lecture 05: Using Hypotheses to test questions

Hypothesis: A specific, clear, and testable proposition or predictive statement about the possible outcome of a scientific research study

Hypothesis Test: A statistical method that uses sample data to evaluate a hypothesis about a population

Null Hypothesis: There is no change, no difference, or no relationship between variables, H_0

Research Hypothesis: There is change, a difference, or a relationship between variables. the research hypothesis, H_1 , is sometimes called the alternative hypothesis, H_a

Alpha: The probability value that is used to define which sample outcomes are considered very unlikely if H_0 is true

Critical Region: The region (of the sampling distribution) that contains the sample outcomes that are considered *very unlikely* if H_0 is true

p-value: The probability of getting the observed or more extreme data, *assuming* the null hypothesis is true

Test statistic: A numerical summary of the degree to which a sample is unlike the samples predicted by the null hypothesis, H_0

Lecture 06: Probability and the Normal Curve

Probability: For a situation in which several different outcomes are possible, the probability for any specific outcome is defined as a *fraction* or a *proportion* of all the possible outcomes.

z-score: Describes the position of a raw score in terms of its distance from the mean, when measured in *standard deviation units*. the *z-score* is positive if the value lies above the mean, and negative if it lies below the mean

Lecture 07: Introduction to Statistical Significance

Statistical Significance: A result is said to be *significant*, or *statistically significant*, if it is very unlikely to occur when the null hypothesis is true. That is, the result is sufficient to reject the null hypothesis. Thus, a treatment has a significant effect if the decision from the hypothesis test is to reject H_0

Type I error: Occurs when there is *no* effect present but the researcher rejects the null hypothesis

Type II error: Occurs when a real effect is present, but the researcher fails to reject the null hypothesis

Specificity: The probability of failing to reject H_0 when H_0 is true (also called the True Negative Rate)

Statistical power: The probability that the test will correctly reject a false null hypothesis. That is, power is the probability that the test will identify an effect if one really exists