

Descriptive/Exploratory Statistics:

1. Mean
 - a. Use: Measure of central tendency, the type of data you can expect from the group
 - b. Pros: Useful when data is evenly distributed
 - c. Cons: If you have large outliers, skewed data
 - d. How: Calculate the average
2. Median
 - a. Use: Finding the middle number as a measure of tendency
 - b. Pros: Good for skewed data
 - c. Cons:
 - d. How: Find the middle number
3. Mode
 - a. Use: Central tendency for most commonly occurring data
 - b. Pros: When there is a large sample, nonnumeric data
 - c. Not well defined, does not use all the data points
 - d. How: Find the most common occurrences
4. Range
 - a. Use: Most basic measure of spread and variability. Determines how close numbers are together or in relation to the mean
 - b. Pros: Simple
 - c. Cons: Only uses two values, tells you very little if anything. Highly susceptible to outliers
 - d. How: The distance between the highest and lowest values
5. Variance
 - a. Use: A measure of dispersion
 - b. This gives an arbitrary number, so you should find standard deviation to standardize it
6. Standard deviation
 - a. Use: Finding the variance in standard units of how far away units are on average from the mean
 - b. Pros:
 - c. Cons:
 - d. How: Take the square root of the variance
7. Correlation and Pearson's r
 - a. Use: Determine the relationship between 2 variables
 - b. Pros: Determining if two things relate
 - c. Cons:
 - d. r is the correlation value, Range: -1 to +1
8. Covariance
 - a. Use: Covariance determines how two variables differ, basically an unstandardized correlation
 - b. Pros:
 - c. Cons: Is dependent on the units, you cannot easily compare covariances

- d. How: $\text{Correlation} * \text{sd of each variable}$
- 9. Binomial Distribution
- 10. Gaussian (standard) Distribution
- 11. Z-scores and Z-Tests
- 12. P Values
- 13. T Tests/t-test/(independent sample t test)
 - a. Use: test the difference between two independent means for hypothesis testing
 - b. Pros:
 - c. Cons:
 - d. How: Needs independent means, continuous dependent variable. You need a calculated t value and critical t value. If calculated is greater than critical, reject null.
- 14. Computing beta (type II error probability)
- 15. F Test (multiple uses)
 - a. Tests the equality of two or more population variances.
- 16. Omega Squared and Partial Omega Squared
 - a. Omega squared is a measure of the strength of association in analysis of variance for fixed effects treatments

Inferential Statistics:

- 1. ANOVA (F test)
 - a. Analysis of variance
 - b. ANCOVA
 - c. MANOVA
- 2. Dunnett's Test
- 3. Dun Sidak Test
- 4. Bonferroni Test
- 5. Holm's Test
- 6. Fisher Hayter Test
- 7. Kruskal Wallace Test
- 8. Welch Test
- 9. Brown Forsythe Test
- 10. Levenes Test
- 11. McNemar Test
 - a.
- 12. Cochrans Test
- 13. Cronbachs Test
- 14. Tukey Test
- 15. Tukey Kramer Test
- 16. Games Howel Test
- 17. Scheffes Test
- 18. Newman Keuls Test
- 19. Duncan Test

- 20. Peritz Test
- 21. Barlett's Test
 - a.
- 22. Levene's Test
- 23. Hartley's Test of Fmax
- 24. Regression
 - a. Regression is
 - i. This is used over trend analysis when the independent variable is continuous with many distinct values.
 - b. b is the raw regression coefficient. This is not standardized and functions the same as β
 - c. β is the standardized regression coefficient, the number of units y increase when x increases by 1.
 - d. Linear regression
 - e. Multiple regression
 - f. Hierarchical regression
- 25. Hypothesis Tests
- 26. Confidence Interval
- 27. Power and Power Analysis
 - a. Power analysis is used to
 - b. Overall power, probability of rejecting false complete null hypothesis
 - c. P subset power focuses on detecting the heterogeneity of means
 - d. Any pair power is the probability of detecting at least one true difference among all pairs of means
 - e. All pairs power is the probability of detecting all true differences among all pairs of means

Key Terms

- 28. Bayesian
- 29. Frequentist
- 30. Experiment and Quasi-Experiment
 - a. Random sampling must be present in an experiment
 - b. Replication must be viable
 - c. Blocking
- 31. Predictive Statistics
- 32. Outlier
- 33. Covariate
 - a. A predictive variable, such that independent variables may be predictive of a dependent variable. A covariate is the independent variable.
- 34. Spread/Dispersion
 - a. Statistics that show how far numbers may be away from the mean or median. How much variation there is in the data. How close or far are data points. Does not tell where the data points are.

35. Variable

- a. Quantitative
 - i. Discrete
 - 1. Interval
 - 2. Ratio
 - ii. Continuous
 - 1. Interval
 - 2. Ratio
- b. Qualitative
 - i. Nominal
 - ii. Ordinal
- c. Latent
- d. Confounding
- e. Control
- f. Nuisance
- g. Lurking

36. Nomological Network

37. Post Hoc, a posteriori, unplanned test

- a. Tests that are used for data snooping after a significant omnibus test

38. A Priori

39. Confirmatory Analysis

40. Pilot Studies

41. Hypothesis Testing

- a. 1 tailed
- b. 2 tailed

42. Standard Deviation

43. Confidence Intervals

44. Degree of freedom

45. Alpha level

46. Interactions

47. Normality

- a. Testing for normality includes making a histogram and seeing if it looks like a normal curve. You can test for outliers, there should be maybe 1 on each side at most. Quantile plots are useful.

48. Nonnormality and the effects of it

49. Type I Error

50. Type II Error

51. Additivity

52. Omnibus Test

53. False Discovery Rate

54. Standardized residuals

- a. measure of strength of the difference between observed and expected values. How significant your cells are to the chi square value

55. Error term, standard error, error variance
- Simply describes how data differs from actual population data
 - Standard error describes the accuracy of a mean when compared to the true mean of a population.
 - Error variance is variance in the statistical scores that is caused by uncontrollable independent variables.
56. Combinations and Permutations
- Rule of subtraction
 - Formal Rule of Multiplication
 - Intuitive Rule of Multiplication
 - Rule of Multiplication
 - Rule of Addition
57. Standardized meaning
-
58. Mediation
- Baron and Kenny is good for $x_1 \rightarrow \text{Mediator}_1 \rightarrow y_1$ only
 - Otherwise use SEM (Multilevel modeling, structural equations modeling)
 - Do not use the terminology “mediated effect,” use “indirect effect.”
59. Multimodal data
- Data that has repeating values, such as 2 modes
60. Main effect
61. Nonmain effect
62. Treatments, levels, interventions
63. Sum of Squares
- A SoS may reflect the effects of the independent variable, individual characteristics of an experimental unit, chance fluctuations in the subjects performance, other nuisance variables that may not have been controlled.
 - Sum of squares between groups (SSBG)
 - Between defined as the difference among the group means by calculating variation of each. This is typically larger than within subjects variance
 - Sum of squares within groups (SSWG), within group variance
 - Within defined as
 - Sum of squares total
64. Homogeneity
- Equality of variance = homogeneity of variance
 - A common assumption (t test, anova, etc)
 - Homogenous data would have equal variance, whereas heterogenous data could have equal means by differing variance
65. Sphericity
- Think about this similarly to homogeneity of variance. This is easier to meet than homogeneity
66. Independent means
- Each unit is only measured once

67. Mean Squares

a.

68. Control

- a. Hold nuisance variables constant
- b. Random assignment for experiments
 - i. Distributes error
 - ii. Permits unbiased error
 - iii. Statistical independence of error
- c. Include nuisance in design

69. Bias

- a. Sample bias
- b. Self-selection bias
- c. Nonresponse bias
- d. Social desirability bias

70. Trimming

a.

71. Validity

- a. Ecological validity
- b. Construct validity
- c. External validity
 - i. Cooke Chapter
- d. Internal validity
- e. Criterion related validity
- f. Content validity
- g. Discriminant validity
 - i. Multitrait multimethod matrix
- h. Convergent validity
 - i. Multitrait multimethod matrix

72. Threats to Validity and Inference

- a. Conclusional validity
 - i. Low power
 - ii. Violated assumptions
 - iii. Fishing for significance
 - iv. Reliability of variables and measures
 - v. Randomness
 - vi. Heterogeneity of respondents
- b. Internal validity
 - i. History
 - ii. Maturation
 - iii. Repeated testing
 - iv. Instruments
 - v. Regression
 - vi. Selection

- vii. Diffusion
 - viii. Imitation
- c. External
 - i. Interactions of testing
 - ii. Setting
 - iii. Selection
 - iv. Treatment
 - v. Reactions
- d. Experimenter expectancy
- e. Demand
- f. Subject predisposition
- g. Placebo effect