Research Methodology Statistics Comprehensive Exam Study Guide

References

Outline

• Basic definition
  – Measurement Scale
  – Population vs. sample
  – Discrete vs continuous variables
  – Qualitative vs quantitative variables
  – Sampling
  – Randomization
  – Sampling error

• Descriptive statistics
  – General
  – Central Tendency
    * Mean
    * Median
    * Mode
  – Variability
    * Range
    * Interquartile Range / Semi-Interquartile Range
    * Variance
    * Standard Deviation
  – Distribution
    * Frequency Distribution
    * Percentile Ranks/Percentiles
– Standard Scores
  * T-scores
  * z-scores
  * Linear Transformation
– Graphing data
  * Bar Graph
  * Polygon
  * Histogram
  * Boxplot
  * Stem-and-leaf
  * Ogive

• Theoretical distribution
  – General
  – Normal Distribution
  – t
  – F
  – Chi-square
  – Binomial
  – Relationships among theoretical distributions

• Correlation
  – General
  – Covariance
  – Scatterplot
  – Point-biserial correlation
  – Pearson product-moment correlation
  – Spearman Rank correlation
  – Phi correlation

• Simple Linear Regression
  – Linear regression model
  – Ordinary least square
  – R-square
  – Standard error of estimate

• Probability
  – General
  – Independent events
– Mutual exclusive
– Sample space
– Sampling with/without replacement

• Sampling distribution
  – General
  – Central Limit Theorem
  – Logic of sampling distribution
  – Standard error

• Hypothesis testing
  – Concepts
    * Logic of hypothesis testing
    * One-tailed two-tailed test
    * p-value
    * Type I and Type II error
    * Power
    * Sample Size
  – z-test
  – One-sample t-test
  – Paired-sample t-test
  – Independent-sample t-test
  – Test of Correlation Coefficient

• Estimation (confidence intervals)
  – Point estimate
  – Confidence Interval
  – Relationship between hypothesis testing and confidence interval

• Nonparametric (chi-square)
  – One-way Chi-square
  – Two-way Chi-square

• One-way between subjects ANOVA
  – Statistical model
  – Main effect
  – Assumption
  – Comparisons
    * Planned
post-hoc adjustments
orthogonality
  – effect size (eta-square & omega-square)
  – power

• two-way between subjects ANOVA
  – statistical model
  – main effect
  – interaction effect
  – simple effects
    * simple main effect
    * simple comparisons
    * marginal comparisons
    * interaction contrast
    * planned
    * post-hoc adjustments
  – assumption
  – effect size
  – power

• one-way within subjects ANOVA
  – statistical model
  – main effect
  – assumption
    * sphericity
    * compound symmetry
    * greenhouse-geiser
    * huynh-feldt
  – comparison
    * planned
    * post-hoc adjustments
    * orthogonality
  – effect size (eta-square & omega-square)
  – power

• two-way within subjects ANOVA
  – statistical model
  – main effect
  – interaction effect
– Simple effects
  * Simple main effect
  * Simple comparisons
  * Marginal comparisons
  * Interaction contrast
  * Planned
  * Post-hoc adjustments
– Assumption
  * Sphericity
  * Compound Symmetry
  * Greenhouse-Geiser
  * Huynh-Feldt
– Effect Size (eta-square & omega-square)
– Power

• Two-way mixed ANOVA
  – Statistical model
  – Main effect
  – Interaction effect
  – Simple effects
    * Simple main effect
    * Simple comparisons
    * Marginal comparisons
    * Interaction contrast
    * Planned
    * Post-hoc adjustments
  – Assumption
    * Sphericity
    * Compound Symmetry
    * Greenhouse-Geiser
    * Huynh-Feldt
  – Effect Size
  – Power

• One-Way Between-Subjects ANCOVA
  – Statistical model
  – Main effect
  – Comparisons
    * Planned
    * Post-hoc adjustments
  – Adjusted means
  – Assumption
    * Homogeneity of regression
**Terminology and Notation - ANOVA**

- $A$ and $B$ will denote IVs
- $Y$ will denote a DV
- $SS$ - sum of squares
- $MS$ - mean squares

**Between-subjects ANOVA**

Subjects appear only in one group/cell

- One-way between-subjects ANOVA: there is only 1 IV ($A$)
  
  - model

  \[ Y_{ij} = \mu + \alpha_j + \varepsilon_{ij} \]  

  where
  
  * $\mu$ - grand mean
  * $Y_{ij}$ is the dependent variable for subject $i$ in $j^{th}$ group
  * $\alpha_j = \mu_j - \mu$ where $\mu_j$ is the mean of a DV for group $j$
  * $\varepsilon_{ij} \sim N(0, \sigma^2)$

  - sum of squares

  \[ SS_{\text{total}} = SS_A + SS_{S/A} \]  

  * $SS_A$ - sum of squares of $A$ (aka sum of squares between groups)
  * $SS_{S/A}$ - sum of square error (aka sum of squares within groups)

- Two-way between-subjects ANOVA: there are 2 IVs ($A$ and $B$)
  
  - model

  \[ Y_{ijk} = \mu + \alpha_j + \beta_k + \alpha\beta_{jk} + \varepsilon_{ijk} \]  

  * grand mean: $\mu$
  * main effect of $A$: $\alpha_j = \mu_j - \mu$
  * main effect of $B$: $\beta_k = \mu_k - \mu$
  * interaction effect $A \times B$: $\alpha\beta_{jk} = \mu_{jk} - \mu - \alpha_j - \beta_k = \mu_{jk} - \mu_j - \mu_k + \mu$
  * individual difference: $\varepsilon_{ijk} = y_{ijk} - \mu_{jk}$

  - sum of squares

  \[ SS_{\text{total}} = SS_A + SS_B + SS_{A \times B} + SS_{S/AB} \]  

  * sum of squares between-groups breaks up into $SS_A$, $SS_B$, and $SS_{A \times B}$
  * $SS_{S/AB}$ - sum of squares error (aka sum of squares within groups)
– main effect of \(A\), main effect of \(B\), and interaction of \(AB\)

– simple effects
  * simple main effects
  * simple comparisons
  * marginal comparisons
  * interaction contrasts (tetrad differences)

**Within-subjects ANOVA**

aka repeated-measures ANOVA. Subjects appear in all the cells.

- one-way within-subjects ANOVA: there is only 1 IV
  
  – model
  \[
  Y_{ij} = \mu + \alpha_j + S_i + \alpha S_{ij} + \varepsilon_{ij}
  \]  
  (5)

  * grand mean: \(\mu\) - mean of the DV
  * iv effect: \(\alpha_j = \mu_j - \mu\)
  * individual difference: \(S_i = \mu_i - \mu\)
  * individual by iv: \(\alpha S_{ij} = y_{ij} - \mu - \alpha_j - S_i\). idiosyncratic response of the subject in a particular condition. differences in skill, ability, or predilection make some subjects perform better in one condition, others in another.
  * variability of the individual observation: \(\varepsilon_{ij}\)

  – sum of squares
  \[
  SS_{\text{total}} = SS_A + SS_S + SS_{A\times S}
  \]  
  (6)

  * sum of squares between-subjects, \(SS_S\)
  * sum of squares within-subjects breaks up into \(SS_A\) and \(SS_{A\times S}\)

- two-way within-subjects ANOVA: there are 2 IVs
  
  – model
  \[
  Y_{ijk} = \mu + \alpha_j + \beta_k + \alpha\beta_{jk} + S_i + \alpha S_{ij} + \beta S_{ik} + \alpha\beta S_{ijk} + \varepsilon_{ijk}
  \]  
  (7)

  * grand mean: \(\mu\)
  * main effect of \(A\): \(\alpha_j = \mu_j - \mu\)
  * main effect of \(B\): \(\beta_k = \mu_k - \mu\)
  * interaction effect of \(AB\): \(\alpha\beta_{jk} = \mu_{jk} - \alpha_j - \beta_k - \mu\)
  * \(S_i = \mu_i - \mu\)
  * \(\alpha S_{ij} = \mu_j - \alpha_j - S_i - \mu\)
  * \(\beta S_{ik} = \mu_k - \beta_k - S_i - \mu\)
  * \(\alpha\beta S_{ijk} = y_{ijk} - \alpha\beta_{jk} - \alpha S_{ij} - \beta S_{ik} - \mu - \alpha_j - \beta_k - S_i\)

  – sum of squares
  \[
  SS_{\text{total}} = SS_A + SS_B + SS_{A\times B} + SS_S + SS_{A\times S} + SS_{B\times S} + SS_{A\times B\times S}
  \]  
  (8)

  * sum of squares between-subjects, \(SS_S\)
  * sum of squares within-subjects breaks up into \(SS_A\), \(SS_B\), \(SS_{A\times B}\), \(SS_{A\times S}\), \(SS_{B\times S}\), \(SS_{A\times B\times S}\)
Two-way mixed ANOVA

Subjects appear in some of the cells (more than 1, less than all). There are 2 IVs (1 within-subjects, 1 between-subjects)

- model
  \[ Y_{ijk} = \mu + \alpha_j + \beta_k + \alpha\beta_{jk} + S_{ij} + \beta S_{ijk} + \epsilon_{ijk} \] (9)
  - grand mean: \( \mu \)
  - main effect of A: \( \alpha_j = \mu_j - \mu \)
  - main effect of B: \( \beta_k = \mu_k - \mu \)
  - interaction effect of AB: \( \alpha\beta_{jk} = \mu_{jk} - \alpha_j - \beta_k - \mu \)
  - between-subjects error: \( S_{ij} = \mu_{ij} - \mu \)
  - within-subjects error: \( \beta S_{ijk} = y_{ijk} - \alpha\beta_{jk} - S_{ij} - \mu - \alpha_i - \beta_k \)

- sum of squares
  \[ SS_{\text{total}} = SS_A + SS_B + SS_{A\times B} + SS_{S/A} + SS_{B\times S/A} \] (10)
  - sum of squares between-groups breaks up into \( SS_A \) and \( SS_{S/A} \)
  - sum of squares within-groups breaks up into \( SS_B, SS_{A\times B} \) and \( SS_{B\times S/A} \)