Intro to Statistics with R: Analysis of Variance (ANOVA)

ANOVA

Analysis of variance (ANOVA)

- Compare groups means across 3+ groups

- Separate, independent groups of subjects

- More than 2 group means from same subjects (i.e. repeated measures)

Working memory training

- Four independent groups (8, 12, 17, 19 sessions)

- Measured IQ before and after training

- Dependent variable is IQ gain

- Null hypothesis: All groups are equal (i.e. all groups have equal IQ gain)

- Alternative hypothesis: More training leads to larger IQ gain

Intro to Statistics with R: Analysis of Variance (ANOVA)

Exploration of the F-test

Working memory ANOVA

- <u>Null</u> hypothesis: all groups are equal

- ANOVA provides a significance test

- Test statistic is the F-test (or F-ratio)

$$F = \frac{\text{Variance between groups}}{\text{Variance within groups}} = \frac{\text{Systematic variance}}{\text{Unsystematic variance}}$$

- Large F-ratio indicates significant effect

Getting to a p-value

- Similar to t-test and family of t-distributions

- Family of F-distributions depends on...

  - Number of subjects in sample

  - Number of groups being compared

Intro to Statistics with R: Analysis of Variance (ANOVA)

F-Ratio

F-ratio

$$F = \frac{\text{Variance between groups}}{\text{Variance within groups}}$$

$$F = \frac{\text{Systematic variance}}{\text{Unsystematic variance}}$$

$$F = \frac{\text{"Good" variance}}{\text{"Bad" variance}}$$

$$F = \frac{MS_{Between}}{MS_{Within}}$$

Intro to Statistics with R: Analysis of Variance (ANOVA)

More ANOVA

ANOVA (continued)

- Calculate a total sum of squares

- Calculate a total degrees of freedom

- Sum of between groups value and within groups value

Summary Table

Source	SS	df	MS	F
A	$n \sum (Y_j - Y_T)^2$	$a - 1$	SS_A/df_A	$MS_A/MS_{S/A}$
S/A	$\sum (Y_{ij} - Y_j)^2$	$a(n-1)$	$SS_{S/A}/df_{S/A}$	-
Total	$\sum (Y_{ij} - Y_T)^2$	$N - 1$	-	-

Effect size

- F-test biased by sample size

- Supplement with effect size

- $R^2 = \eta^2$ (eta-squared) $\longrightarrow$ **Percentage of variance in the dependent variable explained by the independent variable**

- $\eta^2 = SS_A / SS_{Total}$

Assumptions

- Dependent variable is continuous (i.e. interval or ratio variable)

- Dependent variable is normally distributed

- Homogeneity of variance (check with Levene's test)