

---

---

# Types of Statistical Tests

DR. MIKE MARRAPODI

# Tests

- *t* tests
- ANOVA
- Correlation
- Regression
- Multivariate Techniques
- Non-parametric

## *t* tests

- One sample *t* test
- Independent *t* test
- Paired sample *t* test

# One sample $t$ test

- Measures: Mean of a single variable
- When to use: Comparing a known mean against a hypothetical value
- Assumptions: Variable should be normally distributed
- Interpretation: If the  $p$  value is less than .05, the results are significant
- What to use if assumptions are not met:
  - Normality violated, use the Mann-Whitney or Wilcoxon Rank Sum

# Independent $t$ tests

- Measures:
  - Dependent variable (continuous)
  - Independent variable (binary)
- When to use: Compare the means of 2 independent groups
- Assumptions:
  - Dependent variable should be normally distributed
  - Homogeneity of variance (Levene's Test)
- Interpretation: If the  $p$  value is less than .05, the results are significant
- What to use if assumptions are not met:
  - Normality violated, use the Mann-Whitney or Wilcoxon Rank Sum
  - Homogeneity violated, use the second row of results on the  $t$  test table

# Paired samples $t$ tests

- Measures:
  - Dependent variable (continuous)
  - Independent variable (2 points in time or 2 conditions with same group)
- When to use: Compare the means of a single group at 2 points in time (pre test/post test)
- Assumptions:
  - Paired differences should be normally distributed (check with histogram)
- Interpretation: If the  $p$  value is less than .05, the results are significant
- What to use if assumptions are not met: Wilcoxon Signed Rank Test

# One-way ANOVA (Analysis of Variance)

- Measures:
  - Dependent (continuous)
  - Independent (categorical, at least 3 categories)
- When to use: When assessing means between 3 or more groups
- Assumptions:
  - Normal distribution of residuals (check with histogram)
  - Homogeneity of Variance (Levene's Test)
- Interpretation: Null hypothesis states all means are equal; if rejected, conduct a post hoc test to see where actual differences occur
- What to use if assumptions are not met:
  - Normality violated, use Kruskal-Wallis test
  - Homogeneity violated, use Welch test and Games-Howell post hoc test

# One-way ANOVA (with repeated measures)

- Measures:
  - Dependent (continuous)
  - Independent (categorical, with levels as within subject factor)
- When to use: When assessing means between 3 or more groups with dependent variable repeated
- Assumptions:
  - Normal distribution of residuals (check with histogram)
  - Sphericity (Mauchly's Test)
- Interpretation: If the main ANOVA is significant, there is a difference between at least two time points (check where difference occur with Bonferroni post hoc test).
- What to use if assumptions are not met:
  - Normality violated, use Friedman test
  - Sphericity violated, use Greenhouse-Geisser correction



# Two-way ANOVA

- Measures:
  - Dependent (continuous)
  - Independent (categorical, with 2+ levels within each)
- When to use: There are three sets of hypothesis with a two-way ANOVA.  $H_0$  for each set is as follows:
  - The population means of the first factor are equal – equivalent to a one-way ANOVA for the row factor.
  - The population means of the second factor are equal – equivalent to a one-way ANOVA for the column factor.
  - There is no interaction between the two factors – equivalent to performing a test for independence with contingency tables (a chi-squared test for independence).

## Two-way ANOVA (continued)

- Assumptions:
  - Normal distribution of residuals (check with histogram)
  - Homogeneity of variance (Levene's Test)
- Interpretation: When interpreting the results, you need to return to the hypotheses and address each one in turn. If the interaction is significant, the main effects cannot be interpreted from the ANOVA table. Use the means plot to explain the effects or carry out separate ANOVA by group.
- What to use if assumptions are not met:
  - Normality violated, use Friedman test
  - Homogeneity violated, compare p-values with smaller significance level, e.g, .01

# Pearson's correlation coefficient

- Measures:
  - Dependent (continuous)
  - Independent (continuous)
- When to use: When assessing the correlation between 2 or more variables
- Assumptions:
  - Continuous data for each variable (check data)
  - Linearly related variables (create a scatter plot for each variable)
  - Normally distributed variables (create a histogram for each variable)

## Pearson's correlation coefficient (continued)

- Interpretation

<u>Correlation coefficient value</u>	<u>Association</u>
-0.3 to +0.3	Weak
-0.5 to -0.3 or 0.3 to 0.5	Moderate
-0.9 to -0.5 or 0.5 to 0.9	Strong
-1.0 to -0.9 or 0.9 to 1.0	Very strong

- What to use if assumptions are not met:

- If ordinal data, use Spearman's rho or Kendall tau
- Linearity violated, transform the data
- Normality violated, use rank correlation: Spearman's or Kendall tau

# Regression

- Measures:
  - Dependent (continuous)
  - Independent (continuous)
- When to use: When predicting the dependent variable given the independent variable
- Assumptions:
  - Observations are independent (conduct a Durbin Watson test)
  - Residuals should be normally distributed (create a histogram)
  - Linear relationship between independent and dependent variable (create a scatterplot)
  - Homoscedasticity (create a scatterplot)
  - No observations have a large overall influence (look at Cook's and Leverage distances)

# Regression (continued)

- Interpretation
  - ANOVA table: Use this value to make a decision about the null
  - Coefficients table: The 'B' column in the coefficients table provides the values of the slope and intercept terms for the regression line. For multiple regression, (where there are several predictor variables), the coefficients table shows the significance of each variable individually after controlling for the other variables in the model.
  - Model summary: The  $R^2$  value shows the proportion of the variation in the dependent variable which is explained by the model. The level for a 'good model' varies but above 70% is generally considered to be good for prediction.
- What to use if assumptions are not met:
  - Independent observations, check with a statistician
  - Normally distributed residuals violated, transform the dependent variable
  - Linearity is violated, transform either the independent or dependent variable
  - Homoscedasticity is violated, transform the dependent variable
  - Leverage, remove observation with very high leverage

# Multivariate Techniques

Parametric Test	Use
Principal Components Analysis	Data reduction
Factor Analysis (Exploratory/Confirmatory)	Data reduction
Correspondence Analysis	Data reduction
Cluster Analysis	Identify groups of similar subjects
MANOVA	Compare groups
MANCOVA	Compare groups
Linear Discriminant Analysis	Predict group membership

# Non-parametric Tests

Non-parametric Test	Parametric Equivalent
Mann-Whitney (Wilcoxon)	Independent samples <i>t</i> test
Wilcoxon Signed Rank	Paired samples <i>t</i> test
Kruskal-Wallis	One-way ANOVA
Friedman	Repeated measures ANOVA



# Examples from Walden

## Locating

- Library – Dissertations & Theses @ Walden University
- Advanced search
- Terms: Quantitative in Abstract (select from drop down menu)
- Publication date: Last 2 years

Questions?

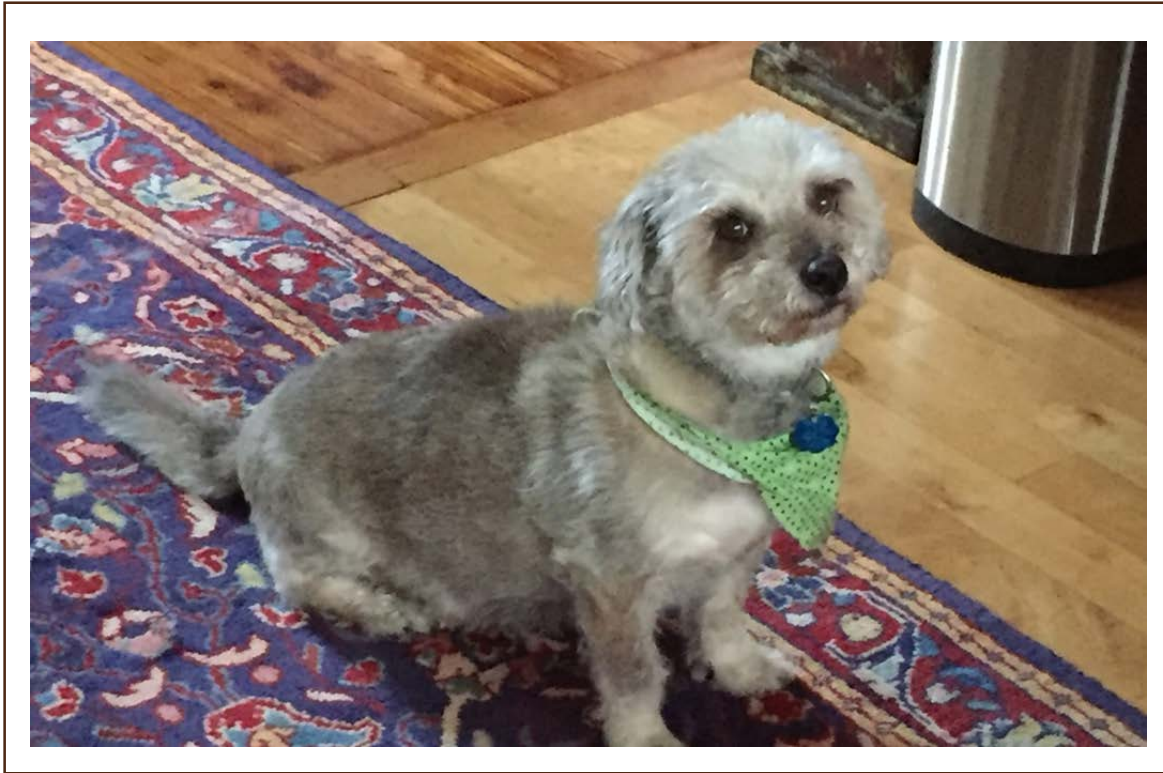
# References

- Antonius, R. (2003). *Interpreting quantitative data with SPSS*. Thousand Oaks, CA: SAGE.
- Balnaves, M., & Caputi, P. (2001). *Introduction to quantitative research methods*. Thousand Oaks, CA: SAGE.
- Cohen, L. (1992). Power primer. *Psychological Bulletin*, 112(1), 155-159.
- Cohen, B. H., & Lea, R. B. (2003) *Essentials of statistics for the social and behavioral sciences*. Hoboken, NJ: John Wiley & Sons, Inc.
- Frankfort-Nachmias, C., & Leon-Guerrero, A. (2018). *Social statistics for a diverse society* (8<sup>th</sup> ed.). Thousand Oaks, CA: SAGE.

# References

- Kaplan, D. (2004). *The SAGE handbook of quantitative methodology for the social sciences*. Thousand Oaks, CA: SAGE.
- Osborne, J. W. (2008). *Best practices in quantitative methods*. Thousand Oaks, CA: SAGE.
- Treiman, D. J. (2009). *Quantitative data analysis: Doing social research to test ideas*. Thousand Oaks, CA: SAGE.
- Vogt, W. P. (2011). *SAGE quantitative research methods*. Thousand Oaks, CA: SAGE.

Thank you!



Casanova says:

“See you next time!”

[michael.marrapodi@mail.waldenu.edu](mailto:michael.marrapodi@mail.waldenu.edu)