

Fall 2016 Course Syllabus PSY 510/610 Categorical Data Analysis

Instructor

Jason Newsom, Ph.D., Office: 317F Cramer Hall Center (3rd floor), Phone: 503-725-5136, Email: newsomj@pdx.edu. Office hours: Mondays 1:30-3:00 PM, Wednesdays 11-noon, 1-2 PM, and by appointment. Website: <http://web.pdx.edu/~newsomj>

Meeting Times and Location

Class: Monday, 8:45 AM – 12:30 PM, Cramer Hall 307. Lab sessions will be held the last hour (from 11:30 – 12:30) of some class periods in NH 450.

Text

Azen, R., & Walker, C. M. (2011). *Categorical data analysis for the behavioral and social sciences*. New York: Routledge. ISBN: 1848728360

Recommended, Optional Texts

Agresti, A. (2013). *Categorical data analysis*. New York: Wiley. ISBN: 0470463635

Wickens, T.D. (1989). *Multiway contingency tables analysis for the social sciences*. Hillsdale, NJ: Erlbaum.

Bilder, C.R. (2015). *Analysis of categorical data with R*. Boca Raton, FL: CRC Press.

Overview

This course is intended to introduce students to categorical data analyses. The general goal is to provide a thorough background in the conceptual aspects, statistical underpinnings, and application of this method rather than a tutorial on a specific software package. By the end of the course, students should be able analyze use categorical data analysis methods to analyze real data using current statistical software, write about, critique applications of, and read methodological articles about categorical data analysis. Prerequisites: Students should have at least one graduate statistics course covering chi-square, ANOVA, and regression analysis, such as PSY 521/621 and PSY 522/622

Homeworks

There will be three homework assignments consisting primarily of data analysis and reporting of categorical analyses using R, SPSS, or SAS. I will supply homework questions and data files (usually 10-12 questions) and you will be asked to analyze your own data for at least one question on each assignment. I can assist you in finding a data set if needed. You will be able to choose among the software programs for many questions, but some analyses may not be available or convenient in all software programs. Illustrations are available in the text or will be provided in class for all types of analyses on the homework assignment. **Homework due dates are: Mon 10/24, Mon 11/14, Mon 12/5.** Late assignments are not accepted without substantial penalty except for cases of illness or family emergencies. Please contact me ahead of time if you are going to miss the deadline for any reason. You are welcome to work with others when running the analyses or consult them on interpretation, but **your assignment must be written in your own words.**

Grades

Grades are based on an average of the three homework assignments with total percentages assigned the following grades: $\geq 90 = A$, $85-89.9 = B+$, $80-84.9 = B$, $75-79.9 = C+$, $70-74.9 = C$.

Other Resources

There are several useful electronic links on the class website.

Disabilities

I am happy to make any necessary arrangements with students who have a disability and are in need of academic accommodations. If you have not done so already, please contact the Disability Resource Center, 116 Smith Memorial Student Union, <http://www.pdx.edu/drc/>, Email: drc@pdx.edu, for assistance and any testing arrangements. I would appreciate it if you would check with me as soon as possible to discuss any needed accommodations and to make sure that I have received a faculty notification letter. If any aspects of instruction or course design result in barriers to your inclusion or learning, please let me know.

Fall 2016 Course Readings PSY 510/610 Categorical Data Analysis

All supplemental readings available online at the class website: <http://web.pdx.edu/~newsomj>

10/3 Descriptive and Univariate Statistics

Levels of scale/measurement, review of probability, descriptive statistics, distributions for binary and categorical variables, test of single proportion, univariate chi-square, estimation basics

Azen & Walker, Chapters 1-3. "Introduction and Overview," "Probability Distributions," and "Proportions, Estimation, and Goodness-of-Fit"

Fife-Schaw, C. (2006). Levels of measurement. In G. M. Breakwell, S. Hammond, C. Fife-Schaw & J. A. Smith (Eds.), *Research methods in psychology* (3rd ed) (pp. 50-63). London: Sage.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological methods*, 7(1), 19-40.

Briefly review: "Graphs, Charts, and Tables—Describing your data," <http://catalogue.pearsoned.co.uk/samplechapter/0130934917.pdf>

Optional: DeCoster, J., Iselin, A. M. R., & Gallucci, M. (2009). A conceptual and empirical examination of justifications for dichotomization. *Psychological methods*, 14(4), 349-366.

10/10 Two Categorical Variables

Contingency of 2 x 2 using Pearson and likelihood ratio chi-square comparing two proportions, I x J tables, measures of association, tetrachoric correlations, interrater agreement statistics

Azen & Walker, Chapter 4 "Association between Two Categorical Variables"

Delucchi, K. L. (1993). On the use and misuse of chi-square. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences* (pp. 295-319). Hillsdale, NJ: Lawrence Erlbaum.

Banerjee, M., Capozzoli, M., McSweeney, L., & Sinha, D. (1999). Beyond kappa: A review of interrater agreement measures. *Canadian Journal of statistics*, 27(1), 3-23.

Chapter 5 "Seeing Frequency Data" in Young, F. W., Valero-Mora, P. M., & Friendly, M. (2006). *Visual statistics: seeing data with dynamic interactive graphics*. New York: John Wiley & Sons. (pp 148-164 only)

Optional: Chapter 6 "Categorical Outcomes" In Kline, Rex B. (2013). In *Beyond significance testing: Statistics reform in the behavioral sciences*. (2nd ed.). Washington, DC: American Psychological Association.

Optional: Pilhoefer, A., & Unwin, A. (2013). New approaches in visualization of categorical data: R package extracat. *Journal of Statistical Software*, 53(7), 1-25.

10/17 Three Categorical Variables, Matched Pairs and Repeated Measures

Mantel-Haenszel, Cochran-Mantel-Haenszel, McNemar's chi-square and related conventional tests, measures analysis for categorical variables

Azen & Walker Chapter 5 "Association Between Three Categorical Variables"

Newsom (2012) Chapter 5 "Basic Longitudinal Analysis Approaches for Continuous and Categorical Variables" (pp. 163-166 only). In J.T. Newsom, R.N. Jones, & S.M. Hofer (Eds.) (2012). *Longitudinal Data Analysis: A Practical Guide for Researchers in Aging, Health, and Social Science*. New York: Routledge.

Agresti, A. (2013). *Categorical data analysis*. New York: Wiley. (pp. 413-418 only).

10/24 HW 1 Due

10/24 Ordinal Analyses for Contingency Tables and Loglinear Models

Azen & Walker Chapter 7 "Log-Linear Models"

Wickens, T.D. (1989). Chapter 13 "Ordered Categories." *Multway contingency tables analysis for the social sciences*. Hillsdale, NJ: Erlbaum.

Nussbaum. Chapter 6 "Basic Nonparametric tests for Ordinal Data." In Nussbaum, E. M. (2014). *Categorical and nonparametric data analysis: choosing the best statistical technique*. New York: Routledge.

10/31 Regression Models for Binary and Ordinal Outcomes I: Binary Outcomes, The Basics

Logistic regression (continuous and binary predictors), interactions with logistic regression, mediation, alternative models: probit models and complementary log-log models

Azen & Walker Chapters 8 & 9 "Logistic Regression with Continuous Predictors" and "Logistic Regression with Categorical Predictors"

Hayes & Mathes (2009) Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations. *Behavior Research Methods*, 41 (3), 924-936

MacKinnon, 2008. Chapter 11 "Mediation in Categorical Data Analysis" New York: Routledge.

11/7 Regression Models for Binary and Ordinal Outcomes II: Diagnostics and Longitudinal Applications

Diagnostics with logistic regression, lagged regression, GEE, conditional logistic models, and multilevel regression for discrete variables; propensity scores

Menard (2010). Chapter 7 "Logistic Regression Diagnostics and Problems of Inference." In Menard, S. (2002). *Applied logistic regression analysis*, second edition. Sage.

Newsom (2012) Chapter 5 "Basic Longitudinal Analysis Approaches for Continuous and Categorical Variables". In Newsom, J.T., Jones, R.N., & Hofer, S.M. (Eds.) (2012). *Longitudinal Data Analysis: A Practical Guide for Researchers in Aging, Health, and Social Science*. New York: Routledge (pp. 168-170)

Hanley, J. A., Negassa, A., & Forrester, J. E. (2003). Statistical analysis of correlated data using generalized estimating equations: an orientation. *American journal of epidemiology*, 157(4), 364-375.

11/14 HW 2 Due

11/14 Regression Models for Binary and Ordinal Outcomes III: Generalized Linear Models, Ordinal Outcomes, Multicategory Outcomes

Generalized linear models, ordered/cumulative logit and probit, multinomial logistic, discrete choice, discrete survival analysis

Azen & Walker Chapter 6 "Generalized linear models"

Long 1997 Chapter 5 "Ordinal Outcomes. Ordered Logit and Ordered Probit Analysis" Long, S.J. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA: Sage.

Azen & Walker Chapter 10 "Logistic Regression for Multicategory Outcomes"

11/21 Sample Size, Estimation, and Practical Issues

Hagenaars, J. A. (2015). Methodological Issues in Categorical Data Analysis. *Methodology*, 11, 126-141.

Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European Sociological Review*, 26(1), 67-82.

Allison, P.D. (2008). Convergence Failures in Logistic Regression. *SAS Global Forum 2008*. Paper 360-2008

11/28 Psychometric Issues

Some basics of IRT, introduction to latent class modeling concepts

DeMars, C. (2010). Chapter 1 "Introduction" and Chapter 4 "Results," from DeMars, C. (2010). *Item response theory*. Oxford University Press.

Clogg, C.C. (1995). Chapter 5 "Latent Class Models" in G. Arminger, C.C. Clogg, M.E. Sobel (Eds.), *Handbook of Statistical Modeling for the Social and Behavioral Sciences* (pp. 311-359). New York: Plenum Press.

12/5 HW 3 due (finals week)