

Assignment 09

Logistic Regression

This assignment is intended to give you experience working with logistic regression models to analyze dichotomous outcome data. *Do not include any R syntax or output unless it is specifically required in the question.* Please submit your responses to each of the questions below. Please submit your responses to each of the questions below in a printed document. All graphics should be resized so that they do not take up more room than necessary and should have an appropriate caption. All tables should also have an appropriate caption.

This assignment is worth 20 points. Each question is worth 1 point unless otherwise noted.

For this assignment, you will use the file, *same-sex-marriage.csv*. This file contains data collected from the 2008 American National Election Study conducted by the University of Michigan and Stanford University. These particular data consist of 1,746 American's responses. The variables in this file are:

- **support:** Does the respondent support gay marriage? (1=Yes; 0=No)
- **attendance:** How often does the respondent attend religious services? (0=Never; 1=Few times a year; 2=Once or twice a month; 3=Almost every week; 4=Every week)
- **denomination:** What is the respondent's religious denomination?
- **friends:** Does the respondent have family or friends that are LGBT? (1=Yes; 0=No)
- **age:** Respondent's age, in years
- **female:** Is the respondent female? (1=Yes; 0=No)

Schwartz (2010) broadly examined the demographic factors that play a role in shaping public opinion regarding same-sex marriage. Drawing on previous literature, Schwartz fitted a regression model to predict an individual's support for gay marriage by looking at several different explanatory variables. His findings religion is an important factor in the prediction of whether or not a person will support same-sex marriage. You will use these data to examine the effects of two aspects of religion (denomination, and frequency of attendance of religious services) on the support of same-sex marriage.

Schwartz, J. (2010). Investigating Differences in Public Support for Gay Rights Issues. *Journal of Homosexuality*, 57(6), 748-759.

Effect of Attendance

You will begin the analysis by examining the effect of attendance on support of same-sex marriage. Because the data for this predictor come from a Likert scale (ordinal in nature), we need to examine whether we can treat it as a continuous predictor in the model, or whether we should treat it as categorical.

1. Begin by computing the proportion of people that support same-sex marriage for each of the attendance categories. Create a line plot that shows the relationship between proportion of support and attendance.
2. Based on the plot you just created, describe the relationship between proportion of support and attendance.

Although a "one-unit" difference on the frequency scale does not have the same meaning at all levels of the scale, we can still treat the data as continuous, using a line (or polynomial) to fit the relationship. The only caveat is that we need to be careful with interpretations, writing or talking generally about the relationship rather than interpreting things as "a one-unit difference in x...".

3. Fit a logistic model to the data using attendance, use a linear and quadratic effect, to predict variation in support for same-sex marriage. Also, write the fitted equation.
4. Is there evidence that the quadratic term is necessary? Explain.
5. Use the fitted model to predict the (a) log-odds, (b) odds, and (c) probability of someone supporting same-sex marriage if that person attends religious services almost every week.

Effect of Denomination

6. Use the `CrossTables()` function in the **gmodels** package (or some other equivalent function) to summarize the cross-tabulations between support of same-sex marriage and religious denominations. Are there sample differences in the proportion of people who support same-sex marriage between the denominations? Explain.
7. Fit a logistic model to the data using denomination to predict variation in support for same-sex marriage. In this model, use Protestant as the reference group. Write the fitted equation.
8. Is there an effect of denomination? Explain.
9. Interpret the effect associated with the **Jewish** coefficient in terms of (a) log-odds, and (b) odds.

Model Taxonomy

10. Begin a table of model results. Add both previously fitted models to this table.
11. Fit the logistic model that includes effects for attendance (all that you adopted earlier) and denomination to predict variation in support of same-sex marriage. Add the results to the table of model results.
12. Are there still effects of denomination? What about attendance? Explain.

Controls

Now you will fit a series of logistic models that include the effects for attendance (all that you adopted earlier) and denomination to predict variation in support of same-sex marriage. In these models, you will also be controlling for other characteristics. You will do this by adding one control at a time, building up the model.

13. To help think about the order in which the controls should be added, create a correlation matrix of the outcome, and all of the control predictors. Based on these results, in which order will you add the controls? Explain.
14. Fit the three additional logistic models. Which control predictors seem important to adopt?
15. Although we fitted three additional models, perhaps not all of them need to be added to the table of model results. Add any models that you think should be added to help tell a compelling narrative about the effects of religious attendance and denomination on the support of same-sex marriage.
16. Write a few sentences that tell the data narrative about the effects of religious attendance and denomination on the support of same-sex marriage. Refer to the models in your table to help readers understand. (2pts.)

Plot of the Final Model

17. Create a display that visually represents your fitted model. Be sure to visually show the effects of all the focal predictors (attendance and denomination). Also show any pertinent controls you think are necessary to help readers understand.
18. Explain how this plot adds to the data narrative; in other words, how does it further help your reader understand the effects of religious attendance and denomination on the support of same-sex marriage? **(2pts.)**