PSY 653 Module 05: Analyses Involving Categorical Dependent Variables

Try it Yourself Activity

Description of the dataset (courtesy of Dr. Kim Henry):

A research team sought to examine factors associated with 21st birthday drinking among female students at a large University. Female students who were nearing age 21 and self-classified as regular drinkers were eligible for the study. In total, 200 students were recruited and agreed to take part in the study. Students were instructed to report to the lab two weeks prior to their 21st birthday. During this lab session, students completed a brief survey that measured alcohol use during the past month (using the Timeline Follow Back Method) and their weight was recorded. One week prior to their 21st birthday, participants were sent a link for an online survey to measure positive alcohol expectancies for drinking on their 21st birthday. Within three days prior to their 21st birthday, students reported to the lab and were given a diary-based data collection form to record several items on their 21st birthday. Students were instructed to record the food that they consumed during the day, the de- gree to which they were in a partying mood just prior to the celebration, and the quantity and type of drinks that they con- sumed during the first two hours of the celebration. The students were also given a small breathalyzer machine to measure BAC 2 hours after consumption of their first drink.

The dataset called bac_module5.csv contains the following variables:

- **weight:** weight in kilograms
- **weight_low:** participants with body weight lower than 1 standard deviation below the mean were coded as, and participants with body weight above 1 standard deviation below the mean were coded as 0
- **alcexp:** positive alcohol expectancy for drinking on the impending 21st birthday, a multi-item scale that ranges from 1-7, where a higher score indicates more positive expectations about the role alcohol will play
- **typ_drks:** the number of standard alcohol drinks consumed in the past 30 days
- pmood: a rating on a scale from 1-9 on the respondent's mood to party on the 21st birthday, where 1 means never been less in the mood to party, and 9 means never been more in the mood to party
- **absorb:** a score calculated from the food diaries to determine how full the participant was when they began drinking, the score ranges from 1 to 8, where 1 means a completely full stomach, and 8 means a completely empty stomach
- **alc_gm:** a score calculated from the drinking diary to estimate the grams of alcohol consumed on the 21st birthday

- **bac:** the participant's blood alcohol content, measured as grams of alcohol per deciliter of blood on the 21st birthday
- **bac_over:** participants were coded as 1 if their measured blood alcohol content was > 0.08 and coded as 0 if their measured BAC was < 0.08.

1. Download the "bac_module5.csv" dataset from the module 05 lab module on canvas
2. Create a new R notebook from your project file and name it "logistic_regression_notebook"
3. Create a first level header: "Load Libraries"
    a. In a new R chunk load in the psych & tidyverse packages
4. Create an R-chunk with the first level header: "Import data"
    a. Read in the datafile "bac_module5.csv". save it to an object named "bac"
5. Create a new first level header: Get variable descriptives
    a. Use any method to get the dataset descriptives
6. Create a first level header: "Logistic regression practice"
    a. Use logistic regression to regress bac_over on alcexp, pmood, weight_low, and typ_drks
        i. Describe what the values of each of the regression coefficients tells you. Are these interpretable?
    b. Create a second level header: "Get exponentiated coefficients and confidence intervals"
        i. Exponentiate your model coefficients to get Odds ratios
            1. Hint: exp(coefficients(model))
        ii. Exponentiate your 95% confidence intervals to get odds ratio confidence intervals
            1. Hint: exp(confint(model))
        iii. In the white space below, interpret each of your odds ratios.
    c. Create a second level header: "Use anova()" to compare deviance of each model"
        i. Use the anova() function to compare the deviance of each model.
            1. Hint: anova(model, test = "Chisq")
        ii. Describe what this test is telling you.
    d. Create a second level header: "Calculate Mcfadden's R^2"
        i. Calculate the Mcfadden's R^2 for each of your models
            1. Hint: Mcfadden's R^2 = 1 – (deviance model / deviance null)