

Homework 3

Due Thur, Jun 4 11 AM (please send a pdf)

For all questions, please show your work or include a copy of the output, whichever is relevant. Please write your answers in report form including relevant statistics, as if you were describing results in a published study. Be sure to describe each finding in terms of the research problem. Most responses to a particular question should be approximately one paragraph in length. **All answers should be in your own words.**

For all questions, please show your work or include a copy of the output, whichever is relevant. Please write your answers in report form, as if you were describing results in a published study. Most responses should not be longer than one paragraph. The data for each problem can be downloaded from the website: <http://web.pdx.edu/~newsomj/data.htm>. **All answers should be in your own words.**

The data (`schoolhb4.sav`) for the questions below come from a study of race and ethnicity and health behaviors among students in grades six through ten.¹ There are several variables in the dataset that will be used in this assignment: `hispanic`, a binary indicator for Hispanic self-identity (0='Anglo/non-Hispanic', 1='Hispanic'); `hwhelp` student perceptions of how willing parents are to help with homework (1='strongly disagree' to 5='strongly agree'); `friends`, a sum of the number of close male and female friends the student has; `mj`, whether the student has ever used marijuana (0='no', 1='yes'); `brkfast`, frequency with which the student eats breakfast on weekends (1='never', 2='usually only one day', 3='usually both days'); `racecat`, indicating student self-identified race, simplified to three categories for this assignment (1 = "White," 2 = "Black or African American," and 3="Other"). For all of the R analyses, I suggest you read in data with the package `haven`, with the code `d = read_sav("c:/myfile/path/schoolhb4.sav")`. (Some packages like the `foreign` package create problems with automatic recoding of values. Otherwise, you will need to recode `hispanic`, `friends`, `mj`, `brkfast`, and `pa` so that they begin with 0.)² The `haven` package method also keeps your variables as numeric. Otherwise, you will need to convert your predictors to numeric.

1. Obtain a cross-tabs (contingency frequency table) and chi-square test of Hispanic ethnicity (`hispanic`) and marijuana use (`mj`) in either SPSS or R and compute the odds ratio by hand (please show your work). Then in both SPSS and R, conduct a simple logistic regression with `hispanic` as predictor of `mj`. Report and interpret your results, including the overall fit, the pseudo- R^2 , all of the odds ratios, significance tests, and confidence intervals. (There is no need to report the results of the initial chi-square analysis).

2. In both SPSS and R, conduct a multiple logistic regression with `hispanic`, `hwhelp`, and `friends` as predictors of `mj`. Report and interpret your results. Be sure to report the likelihood ratio chi-square, pseudo- R^2 , and the regression coefficients, coefficient significance, odds ratios, and confidence intervals for all of the predictors.

3. In SPSS or R, test an ordinal logistic model of `brkfast` using `hispanic`, `hwhelp`, and `friends` as predictors. Report and interpret your results. Be sure to report the likelihood ratio chi-square, pseudo- R^2 , and the regression coefficients, coefficient significance, odds ratios, and confidence intervals for all of the predictors.

¹ A random subsample of the full sample was used for this data set from a WHO study of US school children. United States Department of Health and Human Services. Health Resources and Services Administration. Maternal and Child Health Bureau. Health Behavior in School-Aged Children, 2001-2002 [United States]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2008-07-24. <https://doi.org/10.3886/ICPSR04372.v2>

² I've found that the `dplyr` package is the most convenient way to recode values, for example, where `d` is data frame and `x` is variable:

```
library(dplyr)
d$x = recode(d$x, `3` = 2, `2` = 1, `1` = 0)
```

Note: for recoding when using `haven` to read data, change variable to numeric first, e.g., `d$x <- as.numeric(d$x)`

4. In SPSS or R, test an ordinal probit model of `brkfast` using `hispanic`, `hwhelp`, and `friends` as predictors. Report and interpret your results. Be sure to report the likelihood ratio chi-square, pseudo- R^2 , the regression coefficients, standardized coefficients, coefficient significance, and confidence intervals. Briefly describe any similarities or differences in the conclusions compared with the ordinal logistic model in the previous question.

5. For this problem use your own data or the school health behavior data set. If you use your own data, you should compare at least two groups on at least three continuous dependent variables (which should be at least modestly positively correlated). Use SPSS or R to conduct a MANOVA to compare the groups on the multiple dependent variables. Be sure to provide a brief paragraph description of the study and variables in your answer. If using the school health behavior data set, use SPSS or R to conduct a MANOVA to compare the three race categories (`racecat`) on the dependent variables number of friends (`friends`), perceived availability to help with homework (`hwhelp`), perceived ease of talking to the parents (`partalk`). Report the multivariate results, including the multivariate test that Olsen recommends, significance, variance accounted for, and the univariate tests. Be sure to include the univariate means of each and indicate the direction of the differences to describe your results.

6. Read **one** of the following articles (copies available from the class website (<http://web.pdx.edu/~newsomj/mvclass/>) and write **two paragraphs** summarizing the article. First, describe the study design (e.g., randomized experiment, non-equivalent control group design, cross-sectional survey; for a quick refresher, see <https://www.appinio.com/en/blog/market-research/quasi-experimental-design>) and purpose of the study **in your own words**. Be sure to include who/what was studied (e.g., who were the participants?) and the number of cases. Then, **choose one regression analysis** used in the article that you have learned about in the course so far (i.e., simple or multiple regression), and, **in your own words**, describe the hypothesis that is being tested, the results obtained, and what the findings mean. Be sure to **include the relevant statistical values** and whether the results were significant. Write your paragraphs as if you were describing results in a published article or reporting someone else's results as in a review article.

Bell, M. F., Bayliss, D. M., Glauert, R., & Ohan, J. L. (2018). Using linked data to investigate developmental vulnerabilities in children of convicted parents. *Developmental psychology*, 54(7), 1219.

Newhouse, N. K., & Cerniak, J. (2016). Student Success Factors in Graduate Psychology Professional Programs. *Online Learning*, 20(1), 70-91.

Nuzulia, S., & Why, F. Y. P. (2020). When the Dark Shines: The role of dark personality traits in leadership role occupancy and hiring decisions in a collectivistic culture. *Social Psychological and Personality Science*, 1948550619893956.

Rosen, K. D., Curtis, M. E., & Potter, J. S. (2020). Pain, psychological flexibility, and continued substance use in a predominantly hispanic adult sample receiving methadone treatment for opioid use disorder. *Drug and alcohol dependence*, 206, 1076-81.