

Homework 8 – Stat 230 – Fall 2022

Due date: Friday, November 4

Complete the following exercises and submit your assignment via gradescope (linked on the course webpage).

Note

When the book asks you to “Interpret the model in terms of the odds ratio,” it is asking you to interpret the slope of the logistic regression model.

Problems to start after class Oct 28

Q1

(Adapted from Statistical Sleuth problem 12.13)

A large number of volunteers participated in a randomized experiment to assess the effect of large doses of vitamin C on the incidence of colds. The subjects were given tablets to take daily, but neither the subjects nor the doctors who evaluated them were aware of the dose of vitamin C contained in the tablets. You can load the data from this study by running the following code:

```
data("ex2113", package = "Sleuth3")
```

The data set contains the `Dose` of vitamin C (in g), the `Number` of subjects, the number of subjects who did not report any illnesses during the study period (`WithoutIllness`), and the proportion of subjects who did not report any illnesses during the study period (`ProportionWithout`).

- (a) For each of the four dose groups, calculate the logit of the estimated proportion. Plot the logit versus the dose of vitamin C.

- (b) Fit the binomial logistic regression model using dose to describe the proportion of subjects who did not report illness. Report the estimated coefficients and their standard errors (print the table is fine).
- (c) Report the results of Wald's test for $H_0 : \beta_1 = 0$ vs. $H_a : \beta_1 \neq 0$. Be sure to include the test statistic and p-value, as well as a conclusion in context.
- (d) Report the results of drop-in-deviance test for $H_0 : \beta_1 = 0$ vs. $H_a : \beta_1 \neq 0$. Be sure to include the test statistic, d.f., and p-value, as well as a conclusion in context.
- (e) Report the results of the deviance goodness-of-fit test. Be sure to include the test statistic, d.f., and p-value, as well as a conclusion in context.
- (f) What can be concluded about the adequacy of the binomial logistic regression model? What evidence is there that the odds of a cold are associated with the dose of vitamin C?

Q2

An experiment in the 1980's looked at the levels of resistance to a certain toxin in the adult Budworm, a type of moth. Batches of 20 moths each of each sex were exposed to different doses of pyrethroid *trans*-cypermethrin and the number in each batch that were dead or knocked down were recorded (`numdead`).

You can load the data using the following code:

```
budworm <- read.csv("http://aloy.rbind.io/data/budworm.csv")
```

- (a) Fit a binomial logistic regression model using `sex` and `dose` as predictors of the proportion fo moths that were dead or knocked down. Report the fitted model equation and the results of the deviance goodness-of-fit test. Be sure to include the test statistic, d.f., and p-value, as well as a conclusion in context.
- (b) Create a plots of the Pearson residuals versus the predictors for the model from part (a). Is there any evidence of nonlinearity?
- (c) Fit a binomial logistic regression model using `sex` and `log2(dose)` as predictors of the proportion fo moths that were dead or knocked down. (Note: `log2()` is \log_2 .) Report the fitted model equation and the results of the deviance goodness-of-fit test. Be sure to include the test statistic, d.f., and p-value, as well as a conclusion in context.
- (d) Create a plots of the Pearson residuals versus the predictors for the model from part (c). Is there any evidence of nonlinearity?

- (e) Interpret the regression coefficients from the transformed model in terms of the odds ratio (i.e., on the odds scale). Remember that **dose** has a log base 2 transformation applied. While we haven't discussed this interpretation in class, it will be just like how we interpret a transformed coefficient in linear regression, so think about a doubling of dose.