

Homework 2

PubH 7406: Biostatistical Inference II – Jared D. Huling

Due: Tues, Feb 21, 2023 at the beginning of class

Homework 2 - GLMs/Logistic Regression - 40 points

Instructions

Turn in the homework in the form of a PDF. It is fine to use existing functions to answer questions. However, note that simply providing output from statistical software is not sufficient and will not receive full points. Any output/results must be interpreted in the context of the problem and accompanying explanations of the results are necessary. Use clearly-defined and explained statistical notation to accompany results. For example, any statistical tests should be accompanied by a formal statement of the null and alternative hypotheses with additional context explaining what the hypotheses mean in the context of the problem. Please follow the instructions on homeworks in the syllabus in order to receive full credit.

If you have any questions, please ask in the course Q&A on Canvas so that others can see any responses.

The Questions

1. **(4 points)** Problem 4.2a, pg 154 of CDA.
2. **(8 points)** Problem 4.16 of CDA: For binary data, define a GLM using log link (define any notation you use). Show that the effects refer to the *relative risk*. Why do you think this link is not often used? (Hint: what happens if the linear predictor takes a positive value?)
3. **(9 points)** Problem 5.2 (all parts) of CDA, pg 196-197
4. **(5 points)** Problem 5.14 of CDA, pg 200-201
5. **(6 points)** Problem 5.18 of CDA, pg 202
6. **(8 points)** Problem 5.34 of CDA, pg 205

Extra problems (not graded): CDA 4.7, 4.11, 5.4, 5.5, 5.19, 5.29