

## P8131 Spring 2022 Homework #2

Due on February 20 at 11:59pm

1. The table below gives the data collected from a bioassay study in which  $X$  variable (treated as continuous variable) is the concentration level. At each of five different dose levels (0-4), 30 animals are tested and the number of dying are recorded.

| Dose ( $X$ )    | 0 | 1 | 2  | 3  | 4  |
|-----------------|---|---|----|----|----|
| Number of dying | 2 | 8 | 15 | 23 | 27 |

Fit the model  $g(P(\text{dying})) = \alpha + \beta X$ , with logit, probit, and complementary log-log links.

- (a) Fill out the table and give comments.

| Model     | Estimate of $\beta$ | CI for $\beta$ | Deviance | $\hat{p}(\text{dying} x = 0.01)$ |
|-----------|---------------------|----------------|----------|----------------------------------|
| logit     |                     |                |          |                                  |
| probit    |                     |                |          |                                  |
| c-log-log |                     |                |          |                                  |

- (b) Suppose that the dose level is in natural logarithm scale, estimate LD50 with 90% confidence interval based on the three models.
2. The table below contains the enrollment data of some MPH program in a year
    - Amount: one-time two-year scholarship
    - Offer: the number of offers made with the corresponding scholarship
    - Enrolls: the number of offer accepted

| Amount (in thousand dollars) | Offers | Enrolls |
|------------------------------|--------|---------|
| 10                           | 4      | 0       |
| 15                           | 6      | 2       |
| 20                           | 10     | 4       |
| 25                           | 12     | 2       |
| 30                           | 39     | 12      |
| 35                           | 36     | 14      |
| 40                           | 22     | 10      |
| 45                           | 14     | 7       |
| 50                           | 10     | 5       |
| 55                           | 12     | 5       |
| 60                           | 8      | 3       |
| 65                           | 9      | 5       |
| 70                           | 3      | 2       |
| 75                           | 1      | 0       |
| 80                           | 5      | 4       |
| 85                           | 2      | 2       |
| 90                           | 1      | 1       |

Please analyze the data using a logistic regression and answer the following questions:

- How does the model fit the data?
- How do you interpret the relationship between the scholarship amount and enrollment rate? What is 95% CI?
- How much scholarship should we provide to get 40% yield rate (the percentage of admitted students who enroll?) What is the 95% CI?