

Problem 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.

i) Fill out the table and give comments.

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

ii) Suppose that the dose level is in **natural logarithm** scale, estimate LD50 with 90% confidence interval based on the three models.

Problem 2

Amount	Offers	Enrolls
\$10,000	4	0
\$15,000	6	2
\$20,000	10	4
\$25,000	12	2
\$30,000	39	12
\$35,000	36	14
\$40,000	22	10
\$45,000	14	7
\$50,000	10	5
\$55,000	12	5
\$60,000	8	3
\$65,000	9	5
\$70,000	3	2
\$75,000	1	0
\$80,000	5	4
\$85,000	2	2
\$90,000	1	1

The table 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

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

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