

HOMework #11

Read:

Chapters 12 in 'Regression Analysis by Example'.

R Assignment:

Solve the following questions using R. Hand in your R code and output file for Questions 1-3, as well as answers to all questions.

1. A psychologist conducted a study to examine the nature of the relationship, if any, between a person's emotional stability (X) and their ability to perform a certain task (Y). Emotional stability was measured by a written test for which the higher the score, the greater the emotional stability. Ability to perform the task (Y=1 if able, Y=0 if unable) was evaluated by the psychologist. The results for 27 subjects can be found at:

www.stat.columbia.edu/~martin/W2024/Data/ability.txt

- (a) Fit a logistic regression model to the data.
 - (b) What are the estimated parameters of the regression model (i.e. compute the values of b_0 and b_1)?
 - (c) Perform a significance tests to determine whether emotional stability has a significant effect on the probability that a person can perform the task.
 - (d) Calculate $\exp(b_1)$ and interpret this number.
 - (e) Provide a 95% confidence interval for $\exp(b_1)$.
 - (f) Write an equation for estimating the probability that a person will be able to perform the task as a function of their emotional stability score.
 - (g) What is the estimated probability that a subject with an emotional stability test score of 510 will be able to perform the task? Provide a 95% confidence interval for the estimate.
2. An equal-rights group claims that females are less likely to be hired by a particular firm than males with the same background and experience. Data was collected on years of education, years of experience and gender for 28 former applicants. In addition, it was recorded whether or not the applicant was hired. The data can be found on the web page:

www.stat.columbia.edu/~martin/W2024/Data/discrimination.txt

- (a) Fit a multiple logistic regression model to the data to determine whether the variables education, experience and gender have a significant impact on the probability of being hired.
- (b) What are the estimated parameters of the regression model? Give an interpretation of their values.
- (c) Is there sufficient evidence to conclude that gender has a significant impact on the probability somebody is hired?

- (d) Assuming gender and years of education are held constant, how do the odds of being hired change for each additional year of experience?
- (e) Write an equation for estimating the probability that a person will be hired.
- (f) What is the estimated probability that a *female* with 3 years of education and 1 year of prior work experience will be hired? Provide a 95% confidence interval for the estimate.
- (g) What is the estimated probability that a *male* with 3 years of education and 1 year of prior work experience will be hired? Provide a 95% confidence interval for the estimate.

3. A substance used in medical research is shipped by airfreight to users in cartons containing thousands of ampules. The data below, involving 10 shipments, were collected on the number of times the carton was transferred from one aircraft to another over the shipment route (X) and the number of ampules found to be broken upon arrival (Y).

X	1	0	2	0	3	1	0	1	2	0
Y	16	9	17	12	22	13	8	15	19	11

- (a) Fit a Poisson regression model to determine how the mean number of broken ampules changes as a function of the number of transfers.
- (b) State the estimated regression coefficients and the estimated response function.
- (c) How does the mean number of broken ampules change for each additional transfer? Provide a 95% confidence interval for the estimate.