ISyE 6420
Fall 2019

Homework 6

Due November 17, 2019, 11:55pm. HW6 is not time limited except the due date. Late submissions will not be accepted.

Use of all available electronic and printed resources is allowed except direct communication that violates Georgia Tech Academic Integrity Rules.

1. Cancer of Tongue. Sickle-Santanello et al (1988)[1] provide data on 80 males diagnosed with cancer of the tongue. Data are provided in the file `tongue.csv|dat|xlsx`. The variables in the dataset are as follows:

- Tumor DNA profile (1 - aneuploid tumor, 2 - diploid tumor);
- Time to death or on-study time (in weeks); and
- Censoring indicator (0=censored, 1=observed)

Fit the regression with tumor profile as covariate. What is the 95% Credible Set for the slope β_1?

2. Airfreight Breakage with Missing Data. A substance used in biological and medical research is shipped by air freight to users in cartons of 2,000 ampules. The data below, involving 15 shipments, were collected on the number of times a 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	2	1	0	2	NA	3	1	0	1	2	3	0	1	NA	NA
Y	NA	16	9	17	12	22	13	8	NA	19	17	11	10	20	2

(a) Using OpenBUGS/WinBUGS, fit Y by Poisson regression, with X as a covariate. Report the deviance of your fit.

(b) According to your model, how many packages on average are expected will be broken if the number of shipment routes is $X = 4$? What is 95% CS for your estimate.

(c) For a particular shipment sent from Shenzhen you learned that it would involve $X = 4$ shipping routes. Predict the number of broken packages. What is here different from (b)?

(d) What are estimates for unobserved X_5, X_{14}, X_{15}, Y_1 and Y_9?

Hint: Note that for missing X, you need to specify the distribution. It could be any non-negative valued distribution, but since X is discrete, a good choice is Poisson(2), as in

```
for(i in 1:n){
x[i] ~ dpois(2)}
```

[1] Sickle-Santanello, B. J., Farrar, W. B., DeCenzo, J. F., Keyhani-Rofagha, S., Klein, J., Pearl, D., Laufman, H., and O'Toole R. V. (1988). Technical and statistical improvements for flow cytometric DNA analysis of paraffin-embedded tissue. *Cytometry*, **9**, 594–599.