STA 3000, Fall 2014 — Assignment #2

Due November 21, at start of lecture. Worth 8% of the course grade.

This assignment is to be done by each student individually. You may discuss it in general terms with other students, but the work you hand in should be your own. In particular, you should not leave any discussion of this assignment with any written notes or other recordings, nor receive any written or other material from anyone else by other means such as email.

$X_1, \ldots, X_n$ are non-negative integers that are i.i.d. from the Poisson distribution with mean λ, which has probability mass function $(\lambda^x/x!)\exp(-\lambda)$. We do not necessarily observe $X_1, \ldots, X_n$, however. Instead, we observe $Y_1, \ldots, Y_n$, in which (independently for each i) Y_i is X_i with probability p, and otherwise Y_i is -1. (You can think of this as a model with some "missing data".)

For parts (a) to (e) below, suppose that the parameters of the model are $p \in (0, 1)$ and $\lambda \in (0, \infty)$).

a) Show that this is an exponential family model, and identify natural parameters and natural sufficient statistics for this model.

b) Find the expectations of the natural sufficient statistics for your formulation of this model, by differentiating the normalizing factor of the distribution (*not* by some other method).

c) Find simple formulas for the maximum likelihood estimates for the model parameters λ and p.

d) Find a simple formula for the Fisher information, $I(\lambda, p)$, for this model.

e) Use the Fisher information and maximum likelihood estimates to obtain a formula (in terms of the data, not the unknown parameters) for the standard error of the maximum likelihood estimate for λ.

For parts (f) to (i) below, suppose that p is fixed at a known value, p_0, so the model has only one parameter, $\lambda \in (0, \infty)$.

f) Is this an exponential family model? What is the minimal sufficient statistic for this model?

g) Find a simple formula for the maximum likelihood estimate for λ in this model.

h) Find a non-constant ancillary statistic for this model. Is it independent of the minimal sufficient statistic?

i) If you apply the Conditionality Principle, how would you find a standard error for the maximum likelihood estimate for λ in this model?

And finally,

j) Compare the maximum likelihood estimates found in (c) and (g) above, and the standard errors for these maximum likelihood estimates found in (e) and (i). Discuss to what extent these results are of more general interest.