BIOS760: SOLUTION TO 2013 FALL SEMESTER FINAL EXAM

1. (a) $T_n = \sum_{i=1}^n X_i$ is the complete and sufficient statistic for λ and $T_n \sim$Poisson$(n\lambda)$. Let $E[g(T_n)] = e^\lambda$. This gives

$$\sum_{k=0}^{\infty} g(k)(n\lambda)^k/k! = e^{(n+1)\lambda} = \sum_{k=0}^{\infty}(n+1)^k\lambda^k/k!.$$

Thus, $g(k) = (1+1/n)^k$ so the UMVUE is $(1+1/n)^{T_n}$.

(b) The information for λ is $E[T_n/\lambda^2] = n/\lambda$. Thus, the information bound for e^λ is $e^{2\lambda}\lambda/n$. The variance of UMVUE is

$$E[(1+1/n)^{2T_n}]-e^{2\lambda} = \sum_{k=0}^{\infty}(1+1/n)^{2k}(n\lambda)^k e^{-n\lambda}/k! - e^{2\lambda} = e^{2\lambda+\lambda/n} - e^{2\lambda} = e^{2\lambda}(e^{\lambda/n}-1)$$

so it does not attain this bound.

(c) The MLE for λ is T_n/n so the MLE for e^λ is $e^{T_n/n}$. Thus,

$$\sqrt{n}(e^{T_n/n} - e^\lambda) \to_d N(0, e^{2\lambda}\lambda).$$

(d) The variance of the MLE is

$$E[e^{2T_n/n}] - E[e^{T_n/n}]^2 = \sum_{k=0}^{\infty} e^{2k/n}(n\lambda)^k e^{-n\lambda}/k! - [\sum_{k=0}^{\infty} e^{k/n}(n\lambda)^k e^{-n\lambda}/k!]^2$$

$$= e^{n\lambda e^{2/n} - n\lambda} - [e^{n\lambda e^{1/n} - n\lambda}]^2.$$

The relative efficiency is given by

$$\frac{e^{2\lambda}(e^{\lambda/n}-1)}{e^{n\lambda e^{2/n}-n\lambda} - [e^{n\lambda e^{1/n}-n\lambda}]^2}.$$

2. (a) It is

$$\prod_{i=1}^n \left[(2\pi\sigma^2)^{-1/2} e^{-(Y_i-\beta X_i)^2/(2\sigma^2)} e^{-X_i^2/2} P(R=1|Y_i, X_i)\right]^{R_i}$$

$$\times \left[\int_x (2\pi\sigma^2)^{-1/2} e^{-(Y_i-\beta x)^2/(2\sigma^2)} e^{-x^2/2} P(R=0|Y_i, x)dx\right]^{1-R_i}.$$

(b) The M-step is to maximize

$$\sum_{i=1}^n R_i \left[-\frac{1}{2}\log\sigma^2 - \frac{1}{2\sigma^2}(Y_i - \beta X_i)^2\right]$$

$$+\sum_{i=1}^n (1-R_i)E\left[-\frac{1}{2}\log\sigma^2 - \frac{1}{2\sigma^2}(Y_i-\beta X_i)^2 \Big| Y_i\right].$$

Therefore,

$$\beta = \frac{\sum_{i=1}^n R_i Y_i X_i + \sum_{i=1}^n (1-R_i) Y_i E[X_i|Y_i]}{\sum_{i=1}^n R_i X_i^2 + (1-R_i)E[X_i^2|Y_i]}$$

and

$$\sigma^2 = n^{-1} \sum_{i=1}^{n} \left\{ R_i (Y_i - \beta X_i)^2 + (1 - R_i) E[(Y_i - \beta X_i)^2 | Y_i] \right\}.$$

The E-step computes $E[X_i|Y_i]$ and $E[X_i^2|Y_i]$ for $R_i = 0$. Since (X_i, Y_i) is bivariate normal with mean zero and covariance matrix

$$\begin{pmatrix} 1 & \beta \\ \beta & \beta^2 + \sigma^2 \end{pmatrix}.$$

Thus,

$$E[X_i|Y_i] = \beta/(\beta^2 + \sigma^2) Y_i, \quad E[X_i^2|Y_i] = 1 - \beta^2/(\beta^2 + \sigma^2) + E[X_i|Y_i]^2.$$

(c) In this case, the M-step is to maximize

$$\sum_{i=1}^{n} R_i \left[-\frac{1}{2} \log \sigma^2 - \frac{1}{2\sigma^2} (Y_i - \beta X_i)^2 \right]$$

$$+ \sum_{i=1}^{n} E \left[-\frac{(1 - R_i)}{2} \log \sigma^2 - \frac{1}{2\sigma^2} (1 - R_i)(Y_i - \beta X_i)^2 \Big| Y_i \right].$$

Thus,

$$\beta = \frac{\sum_{i=1}^n R_i Y_i X_i + \sum_{i=1}^n Y_i E[(1 - R_i) X_i | Y_i]}{\sum_{i=1}^n R_i X_i^2 + E[(1 - R_i) X_i^2 | Y_i]}$$

and

$$\sigma^2 = n^{-1} \sum_{i=1}^{n} \left\{ R_i (Y_i - \beta X_i)^2 + E[(1 - R_i)(Y_i - \beta X_i)^2 | Y_i] \right\}.$$

The E-step computes

$$E[(1 - R_i)|Y_i], \quad E[(1 - R_i) X_i | Y_i], \quad E[(1 - R_i) X_i^2 | Y_i]$$

and notice $R_i = I(X_i > 0)$.