BIOS 760 Final Exam (Fall, 2016)

Consider n i.i.d random vectors $(Y_i, X_i), i = 1, ..., n$, and they satisfy

$$Y_i = \beta X_i + \epsilon_i, \quad X_i \sim N(0, 1),$$

where ϵ_i is independent of X_i and follows distribution $N(0, 1)$. Only β is an unknown parameter. Due to non-responses, some Y_i's are not observed and we use R_i to indicate whether Y_i is observable ($R_i = 1$: Y_i is observed; otherwise, $R_i = 0$). Therefore, the observed data can be expressed as

$$(R_i Y_i, R_i, X_i), \quad i = 1, ..., n.$$

1. We assume $P(R_i = 1|Y_i, X_i) = p$ for a known constant $p \in (0, 1)$.
   (a) (1 point) What is the missing mechanism?
   (b) (2 point) Write down the joint likelihood of the observed data.
   (c) (2 point) Treat $X_1, ..., X_n$ as fixed values. Identify a complete and sufficient statistic for β.
   (d) (2 point) Treat $X_1, ..., X_n$ as fixed values. Derive the UMVUE for β.
   (e) (3 point) Treat $X_1, ..., X_n$ as fixed values. Calculate the Cramér-Rao lower bound for β based on these n observations. Does the UMVUE achieve this lower bound?
2. [1] The missingness, R_i, depends on (Y_i, X_i, W_i) where W_i is some other observed random variable and may be correlated with (Y_i, X_i). Assume that $P(R_i = 1|Y_i, X_i, W_i) = p(W_i)$ for a known function $p(w)$ satisfying $0 < m < p(W_i)$ for a constant m.
   (a) (1 point) What is the missing mechanism?

[1]

(**Continued on the other side**)

(b) (2 point) Show that

$$\sum_{i=1}^{n} \frac{R_i}{p(W_i)} X_i(Y_i - \beta X_i) = 0$$

is an estimating equation for β.

(c) (2 point) What is the asymptotic distribution of the estimator that solves the above equation? You don't have to simplify the final variance expression.

3. Now assume that subject i does not respond if and only if $Y_i > c$ for some known value c.

   (a) (1 point) What is the missing mechanism?

   (b) (2 point) Write down the joint likelihood of the observed data.

   (c) (2 point) We use the EM algorithm to calculate the MLE for β. Give the details of the EM algorithm. You can leave integrations in the expressions.