BIOS760: 2012 FALL SEMESTER FINAL EXAM

1. Let $X_1, \ldots, X_n$ be an i.i.d. sample of random variables with density
$$f_\theta(x) = e^{-(x-\theta)} 1\{x > \theta\},$$
where $1\{A\}$ is the indicator of A and the parameter θ is a real number; and let $X_{(1)}$ be the smallest value in the sample. Do the following:

   (a) (3 points) Show that $f_\theta(x)$ is a density and that $X_{(1)}$ is a sufficient statistic for θ.

   (b) (3 bonus points) Show that $X_{(1)}$ has density
   $$g_\theta(x) = ne^{-n(x-\theta)} 1\{x > \theta\},$$
   and verify that $X_{(1)}$ is also a complete statistic for θ.

   (c) (2 bonus points) Derive the UMVU estimator of θ.

2. Let $U_1, \ldots, U_n$ be an i.i.d. sample of random variables with density
$$h_p(u) = 2\left[pu + (1-p)(1-u)\right] 1\{0 \le u \le 1\},$$
where $p \in [0, 1]$. Do the following:

   (a) (2 points) Compute the score $\dot{\ell}_n(p)$ and information $I_n(p) = -\ddot{\ell}_n(p)$ for the sample.

   (b) (3 points) Verify that the Newton-Raphson update for finding the maximum likelihood estimator for p has the form
   $$p^{(k+1)} = p^{(k)} + \left[\sum_{i=1}^n \frac{(2U_i - 1)^2}{[p^{(k)}U_i + (1-p^{(k)})(1-U_i)]^2}\right]^{-1} \sum_{i=1}^n \frac{2U_i - 1}{p^{(k)}U_i + (1-p^{(k)})(1-U_i)}.$$

   (c) (3 points) Now suppose we use the EM algorithm to stabilize computation of the estimator, and so we add a latent Bernoulli random variable Δ with probability of success p to form the random pair (Δ, U) with joint density
   $$m_p(\delta, u) = 2(pu)^\delta \left[(1-p)(1-u)\right]^{1-\delta}.$$
   Verify that the marginal density of u based on $m_p(\delta, u)$ equals $h_p(u)$ given above, and show that the sample log-likelihood for the full model is
   $$\sum_{i=1}^n \log(2) + \Delta_i \left[\log(p) + \log(U_i)\right] + (1 - \Delta_i)\left[\log(1-p) + \log(1-U_i)\right].$$

   (d) (3 points) Denoting $p^{(k)}$ as the current iteration of the estimator for p, show that
   $$\mathrm{E}\left[\Delta_i | U_i, p^{(k)}\right] = \frac{p^{(k)}U_i}{p^{(k)}U_i + (1-p^{(k)})(1-U_i)}.$$

   (e) (4 points) If we denote the term on the right side of the above expression as $W_k(U_i)$, show that the EM update for estimating p is
   $$p^{(k+1)} = n^{-1} \sum_{i=1}^n W_k(U_i).$$

   (f) (2 points) What happens to $p^{(k+1)}$ in the above EM algorithm if $p^{(k)} = 0$? What happens to $p^{(k+1)}$ in the Newton-Raphson update if $p^{(k)} = 0$. What does this say about the EM algorithm in this situation?