BIOS760: 2011 FALL SEMESTER FINAL EXAM

1. Assume that $X_1, ..., X_n$ are i.i.d from the uniform distribution in $[-\theta, \theta]$, where θ is a positive and unknown parameter. Let $|X|_{(n)}$ denote $\max(|X_1|, ..., |X_n|)$ and $g(\theta)$ be a continuously differentiable function of θ satisfying $g'(\theta) > 0$.

   (a) (5 points) Show that $|X|_{(n)}$ is a complete and sufficient statistic for θ.

   (b) (5 points) Show that
   $$n\left(1 - \frac{|X|_{(n)}}{\theta}\right)$$
   converges in distribution to a non-degenerate distribution. Give the explicit expression of the limiting distribution.

   (c) (5 points) What is the maximum likelihood estimator for $g(\theta)$? Derive the asymptotic distribution of the maximum likelihood estimator. Furthermore, construct an asymptotic confidence interval with coverage $(1-\alpha)$ for $g(\theta)$ based on the maximum likelihood estimator.

   (d) (5 points) Find the UMVUE for $g(\theta)$.

   (e) (5 points) Derive the asymptotic distribution of the UMVUE. Furthermore, construct an asymptotic confidence interval with coverage $(1-\alpha)$ for $g(\theta)$ based on the UMVUE.

2. Let X and ξ be two independent random variables and moreover, ξ follows a gamma distribution with density $\Gamma(\lambda)^{-1}\lambda^{\lambda}\xi^{\lambda-1}\exp\{-\lambda\xi\}$ where λ is a known constant larger than 1. We assume that conditional on (X, ξ), a response variable Y follows the distribution
   $$P(Y > t|X, \xi) = \exp\{-te^{\beta X}\xi\},$$
   where β is an unknown parameter. In a real study, ξ is not observed so we only obtain n i.i.d data $(X_i, Y_i), i = 1, ..., n$.

   (a) (5 points) Write down the conditional density of Y given X then give the observed likelihood function.

   (b) (2 points) Under what condition(s) is β identifiable?

   (c) (5 points) Show that $E[Y|X] = ce^{-\beta X}$ for some constant c and identify the constant c. Thus, we can construct an estimating equation for β as
   $$\sum_{i=1}^{n} X_i \left\{Y_i - ce^{-\beta X_i}\right\} = 0.$$
   Give the asymptotic distribution of the estimator which solves this equation. You do not need justify this distribution and may leave mathematical expressions in the result.

   (d) (5 points) Write the score equation for the maximum likelihood estimator and derive the Newton-Raphson iteration for solving this equation.

   (e) (8 points) Another way of calculating the maximum likelihood estimator is via the EM algorithm. In this case, we treat $\xi_1, ..., \xi_n$ as missing data. Give the details in the E-step and M-step.