STAT 610 **Quiz 2** **Newton: 3–13–07.**

Record all answers in the blue books. Show your work.

1. From a random sample of n pairs (X_i, Y_i) we compute the sample correlation coefficient $\hat{\rho}_n$, which estimates the population correlation ρ. It is known that $\sqrt{n}(\hat{\rho}_n - \rho) \longrightarrow_d N\left(0, (1-\rho^2)^2\right)$ as $n \to \infty$. With $\phi = g(\rho) = (1/2)\log\left(\frac{1+\rho}{1-\rho}\right)$, and with $\hat{\phi}_n = g(\hat{\rho}_n)$, show that $\sqrt{n}(\hat{\phi}_n - \phi) \longrightarrow_d N(0,1)$. How can this result be used to obtain an approximate 95% confidence interval for ρ?

2. Independent random variables $X_1, X_2, \ldots, X_n$ have a Pareto distribution with probability density function, for $x \geq 1$,

$$f(x;\theta) = \frac{\theta}{x^{\theta+1}}$$

where $\theta > 0$ is an unknown parameter. A sample is realized as $x_1, x_2, \ldots, x_n$.

   (a) Write the likelihood, log likelihood, and score functions and find the maximum likelihood estimate $\hat{\theta}$.

   (b) Describe how to use the nonparametric bootstrap to estimate the standard error of $\hat{\theta}$.

3. Counts $X_1, \ldots, X_n, X_{n+1}, \ldots, X_{2n}$ are independent and Poisson distributed. For $i = 1, 2, \ldots, n$, the Poisson mean is $a\theta$ and for $i = n+1, \ldots, 2n$ the Poisson mean is $b\theta$, where $a, b > 0$ are distinct known constants. The counts $\{X_i\}$ are not observed directly. Instead we observe $Y_i = 1[X_i \geq 1]$ for all i.

   (a) Give a formula for the log likelihood $l_c(\theta)$ which would have been obtained had we been able to observe $\{X_i\}$.

   (b) For an initial guess θ_1 of θ, evaluate

$$J(\theta, \theta_1) = E_{\theta_1}\left[l_c(\theta) | Y_i = y_i \text{ for all } i\right],$$

   where randomness is in the latent X_i's.

   (c) Derive the value θ_2 which maximizes $J(\theta, \theta_1)$ in θ.

   (d) In words, briefly explain the EM algorithm in the present context.