Stat 581 Homework 9: Due December 8, 2004

1. Suppose that $X_1, ..., X_n$ are i.i.d. with Weibull density $c\theta^{-c}x^{c-1}\exp(-(x/\theta)^c)I_{(0,\infty)}(x)$.

(a) Show that $(X/\theta)^c$ is a standard exponential r.v.

(b) Find the Information, $I(\theta)$, when c is known.

(c) When (c, θ) are both unknown, find the Rao (Score) test of the hypothesis $c = 1$. (You may here **assume** the information matrix given in JAW notes P. 3.13).

(d) Why would you use the Rao (Score) test in this example?

2. In a sample from a trinomial distribution, the cell counts are (n_1, n_2, n_3), $n_1 + n_2 + n_3 = n$, and the three cell probabilities are $(\theta, \theta^2 + \theta^4, 1 - \theta - \theta^2 - \theta^4)$, where $0 < \theta < \theta_0$, where θ_0 is the smallest root in (0,1) of $1 = \theta + \theta^2 + \theta^4$.

(a) Show that n_1/n is a consistent asymptotically Normal estimator of θ.

(b) Find the information for θ, as a (messy) function of θ.

(c) Show that n_1/n is not an asymptotically efficient estimator of θ.

(d) Find the one-step estimator of θ based of the initial estimator n_1/n. Is this estimator asymptotically efficient?

3. Suppose X_i, $i = 1, \ldots, n$, are i.i.d. from the mixture distribution

$$\theta\mathcal{P}(\lambda) \;+\; (1-\theta)\mathcal{P}(\mu) \;\; 0 \le \theta \le 1, \lambda \;>\; 0\;, \quad \mu \;>\; 0\;.$$

where $\mathcal{P}(\lambda)$ denotes the Poisson distribution mean λ.

(a) Let $Z_i = e_1 = (1, 0)'$ if X_i is from the first component $P(\lambda)$ and let $Z_i = e_2 = (0, 1)'$ if X_i is from the second component $P(\mu)$. Let $S_j = \{i : \; Z_i \; = \; e_j\}$, $j \; = \; 1, 2$. Show that the natural sufficient statistics for (θ, λ, μ) for the family of distributions of the "complete data" $\{(X_i, \; Z_i);\; i = 1, \ldots, n\}$ are $(\sum_{S_1} X_i, \; \sum_{S_2} \; X_i, \; \sum_{S_1} \; 1)$

(b) Hence construct the EM algorithm equations for obtaining the MLE of (θ, λ, μ) from a sample $x_1, ..., x_n$.

(c) Investigate briefly the performance of your algorithm when
(i) $n \; = \; 7$, $\mathbf{x} = (0,\ 1,\ 2,\ 6,\ 7,\ 9,\ 10)$
(ii) $n \; = \; 7$, $\mathbf{x} = (3,\ 3,\ 5,\ 5,\ 6,\ 6,\ 7)$.

4. A simple model for a quantitative trait (e.g. height) on a set of k related individuals gives rise to a multivariate Normal distribution for the observed trait values: $\mathbf{y}$ is distributed as $N_k(\mathbf{0}, \sigma^2\mathbf{G} + \tau^2\mathbf{I})$ where $\mathbf{G}$ is a known positive definite matrix, and $\mathbf{I}$ is the identity matrix. An EM-type formulation (which actually has a "real" interpretation for the latent variables $\mathbf{z}$) is $\mathbf{y} \; = \; \mathbf{z} \; + \; \mathbf{e}$ where $\mathbf{z}$ is $N_k(\mathbf{0}, \sigma^2\mathbf{G})$, and $\mathbf{e}$ is $N_k(\mathbf{0}, \tau^2\mathbf{I})$.

(a) Show that if $\mathbf{z}$ and $\mathbf{y}$ could be observed, then the MLEs of σ^2 and τ^2 would be $k^{-1}\mathbf{z}'\mathbf{G}^{-1}\mathbf{z}$ and $k^{-1}(\mathbf{y} - \mathbf{z})'(\mathbf{y} - \mathbf{z})$

(b) Let $\mathbf{V} = \mathrm{var}(\mathbf{y}) = \sigma^2\mathbf{G} + \tau^2\mathbf{I}$. Show

$$\begin{aligned} \mathrm{var}(\mathbf{z} \mid \mathbf{y}) &= (\sigma^{-2}\mathbf{G}^{-1} + \tau^{-2}\mathbf{I})^{-1} = \sigma^2\tau^2\mathbf{V}^{-1}\mathbf{G} \\ \mathbf{a} &= \mathrm{E}(\mathbf{z} \mid \mathbf{y}) = \sigma^2\mathbf{V}^{-1}\mathbf{G}\mathbf{y} \end{aligned}$$

(c) Hence show

$$\begin{aligned} \mathrm{E}(\mathbf{z}'\mathbf{G}^{-1}\mathbf{z} \mid \mathbf{y}) &= \mathbf{a}'\mathbf{G}^{-1}\mathbf{a} + \sigma^{\mathbf{2}}\tau^{\mathbf{2}}\mathrm{tr}(\mathbf{V}^{-1}) \\ \mathrm{E}((\mathbf{y}-\mathbf{z})'(\mathbf{y}-\mathbf{z}) \mid \mathbf{y}) &= (\mathbf{y}-\mathbf{a})'(\mathbf{y}-\mathbf{a}) + \sigma^{\mathbf{2}}\tau^{\mathbf{2}}\mathrm{tr}(\mathbf{V}^{-1}\mathbf{G}) \end{aligned}$$

and write down the EM equations to find the MLE of (σ^2, τ^2) given data $\mathbf{y}$.

5. Suppose that U, V, and W are independent Poisson random variables with means λ, μ, and ψ respectively. However, only $X \equiv U + W$ and $Y \equiv V + W$ are observable.

(a) Show that, if $\theta = (\lambda, \mu, \psi)$, then for non-negative integers x and y

$$p(x,y) \equiv P_\theta(X = x, Y = y) = \exp(-(\lambda+\mu+\psi)) \sum_{w=0}^{\min(x,y)} \frac{\lambda^{x-w}\mu^{y-w}\psi^w}{w!(x-w)!(y-w)!}$$

and hence that, for positive integers x and y,

$$\mathrm{E}_\theta(W|X = x, Y = y) = \psi \frac{p(x-1, y-1)}{p(x,y)}$$

(b) A sample of i.i.d pairs (X_i, Y_i), $i = 1, ..., n$ are taken from the above distribution. Assume λ, μ and ψ are all strictly positive. Treat (X_i, Y_i, W_i) as the *complete data* and use (a) to propose an EM algorithm for the estimation of $\theta = (\lambda, \mu, \psi)$.

(c) Suppose now that is is desired to test $\psi = 0$. Show that when $\psi = 0$, X and Y are independent Poisson r.vs with means λ and μ, and hence that the constrained MLE's are $\tilde{\lambda} = n^{-1}\sum_1^n X_i$ and $\tilde{\mu} = n^{-1}\sum_1^n Y_i$. Explain why you might prefer to use the Rao (Score) test, rather than either the likelihood ratio test or the Wald test. Explain also why the usual theory about the asymptotic distribution of the test statistic might not apply to testing this particular hypothesis.

(d) Show that for a single observation (X, Y) the score $\frac{\partial}{\partial\theta}\log p(X,Y)$ at $\theta \equiv (\lambda, \mu, \psi) = (\lambda, \mu, 0)$ is

$$((-1 + X/\lambda), (-1 + Y/\mu), (-1 + XY/\lambda\mu)),$$

and explain how you would use this to estimate the information matrix at $\psi = 0$.
DO NOT attempt to find the information matrix.
(Hint: Consider $p(x, y)$ when ψ is very small.)

(e) **Given** that when $\psi = 0$, $I_{\psi\psi\cdot(\lambda,\mu)} = (\lambda\mu)^{-1}$, determine the Rao (Score) statistic for testing $\psi = 0$ as explicitly as you can.