Statistics 582, Midterm Exam

Wellner; 2/13/98

1. (24 points) **Define** *any three* of the following terms. In each case, provide an appropriate context for your definition.
(a) The *Rao or score statistic* for testing a simple null hypothesis $H : \theta = \theta_0$ in a regular parametric model.
(b) The *Rao or score statistic* for testing a composite null hypothesis $H : \theta_1 = \theta_{10}$ in a regular parametric model. with $\theta = (\theta_1, \theta_2)$, $\theta_1 \in R^m$, $\theta_2 \in R^{k-m}$.
(c) A *one-step "approximate MLE"* (starting from an $n^{1/4}$-consistent estimator).
(d) The *Wald statistic* for testing a simple null hypothesis $H : \theta = \theta_0$ in a regular parametric model.
(e) The *likelihood equations* for estimation of θ in a regular parametric model.
(f) The *Kullback-Leibler information* $K(P, Q)$ between two probability distributions P and Q on a measurable space $(\mathcal{X}, \mathcal{A})$.

2. (24 points) **State** *any three* of the following results:
(a) A theorem about the large sample distribution of the MLE in a regular parametric model with parameter set $\Theta \subset R^k$.
(b) A theorem about the behavior of the likelihood ratio statistic for testing a simple null hypothesis under a fixed alternative $\theta \neq \theta_0$.
(c) Any theorem / result about nonparametric nonparametric maximum likelihood estimation.
(d) A uniform strong law of large numbers (or Glivenko - Cantelli theorem).
(e) Wald's theorem on strong consistency of maximum likelihood estimators.
(f) A theorem about the limiting distribution of the likelihood ratio statistic for testing a composite null hypothesis $H: \ \theta_1 = \theta_{10}$ versus $K: \ \ \theta_1 \neq \theta_{10}$ under local alternatives $\theta_{1n} = \theta_{10} + n^{-1/2} t_1$.

Do **either** problem 3 **or** problem 4.

3. (30 points) Suppose that $X_1, \ldots, X_n$ are i.i.d. with mixture density (mass function)
$$p(x; \lambda, \mu, \theta) = \theta \frac{\lambda^x}{x!} e^{-\lambda} + (1-\theta) \frac{\mu^x}{x!} e^{-\mu}, x = 0, 1, \ldots,$$
where $0 < \theta < 1$, $0 < \lambda \neq \mu < \infty$; in other words, p is the mixture of two Poisson distributions with parameters λ and μ respectively.
A. Describe an EM - algorithm for estimation of (λ, μ, θ).
B. What is the natural corresponding nonparametric model for the data which

were modeled with the parametric mixture distribution in A? What is the natural nonparametric maximum likelihood estimator here?

4. (30 points) Suppose that (X_i, Y_i), $i = 1, \ldots, n$ are independent pairs of random variables with

$$X_i \sim \text{exponential}(\beta_i/\alpha)\,, \qquad Y_i \sim \text{exponential}(1/\beta_i\alpha)$$

independent. Here $\alpha > 0$ and $\beta_i > 0$ for $i = 1, \ldots, n$ are all unknown. Thus the joint density of (X_i, Y_i) is

$$f_{X_i,Y_i}(x_i, y_i) = \alpha^{-2} \exp(-\beta_i x/\alpha) \exp(-y_i/\alpha\beta_i) 1_{[0,\infty)}(x_i) 1_{[0,\infty)}(y_i)\,.$$

A. Find the maximum likelihood estimator $\widehat{\alpha}$ of α.
B. Do our theorems about consistency and asymptotic normality of maximum likelihood estimators apply to $\widehat{\alpha}$? Why or why not? To what (famous) model is the above model analogous?

5. (45 points) Suppose that $X_i = (Y_i, Z_i)$, $i = 1, \ldots, n$ are i.i.d. with $(X|Z = z) \sim Poisson(\alpha_0 \exp(\beta_0 z))$ where $\theta \equiv (\alpha, \beta) \in R^+ \times R$; suppose that the distribution H of Z is known and that Z is not degenerate at a single point. You may assume that Z is bounded: $|Z| \le c$ with probability 1 for some $c < \infty$.
A. Find the scores for (α, β) based on one observation $X = (Y, Z)$.
B. What is the information matrix for θ? Hint: compute conditionally on Z, and leave the information matrix in terms of expectations of functions of Z.
C. Find the score statistic for testing $H : \beta = 0$ (and $\alpha =$ anything) versus $K : \beta \neq 0$. What is its asymptotic distribution under H?
D. If β is fixed, show that the likelihood is maximized as a function of α by

$$\hat{\alpha}(\beta) = \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n \exp(\beta Z_i)}\,.$$

Use this to compute the profile log-likelihood function

$$l_n^{prof}(\beta) = l_n(\hat{\alpha}(\beta), \beta)\,.$$

E. Suppose that an ad hoc consistent estimator $\overline{\beta}_n$ of β_0 is available: $\overline{\beta}_n \to_{a.s.} \beta_0$. Prove that

$$\hat{\alpha}(\overline{\beta}) \to_{a.s.} \alpha_0\,.$$