Statistics 582, Midterm Exam

Wellner; 2/9/2007

1. (24 points) **Define** *any three* of the following terms. In each case, provide an appropriate context for your definition.
   (a) The *risk function* $R(\theta, d)$, $\theta \in \Theta$, of a decision rule d.
   (b) The *Bayes risk* $\mathcal{R}(\Lambda, d)$ of a decision rule d with respect to a prior distribution Λ.
   (c) An *inadmissible* decision rule d.
   (d) An *envelope function* F of a class of functions $\mathcal{F} = \{f : \mathcal{X} \to \mathbb{R}\}$ in the context of a Uniform Strong Law of Large Numbers (or Glivenko-Cantelli theorem).

2. (24 points) **State** *any two* of the following results:
   (a) A uniform strong law of large numbers (or Glivenko - Cantelli theorem).
   (b) Wald's theorem on strong consistency of maximum likelihood estimators.
   (c) An identity expressing the survival function $1 - F(t)$ of a random variable X with distribution function $F(t) = P(X \le t)$ on $[0, \infty)$ in terms of the hazard function $\Lambda(t) \equiv \int_{[0,t]} (1 - F(s-))^{-1} dF(s)$.
   (d) Huber's $Z-$theorem for solutions $\hat{\theta}_n$ of $\Psi_n(\hat{\theta}_n) = 0$.

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

3. (40 points) Suppose that P and $Q_1, \ldots, Q_r$ are probability measures on a measurable space $(\mathcal{X}, \mathcal{A})$, and $\lambda_1, \ldots, \lambda_r$ satisfy $0 \le \lambda_j \le 1$ and $\sum_{j=1}^r \lambda_j = 1$. Let $K(P, Q)$ be the Kullback-Leibler information (or divergence) between P and Q.
   (a) Define $K(P, Q)$ in terms of P and dP/dQ.
   (b) Show that $K(P, \sum_{j=1}^r \lambda_j Q_j) \le \sum_{j=1}^r K(P, Q_j)$.
   (c) Show that $g(x) = x \log x$ is a convex function for $x > 0$.
   (d) Show that $K(\sum_{j=1}^r \lambda_j Q_j, P) \le \sum_{j=1}^r \lambda_j K(Q_j, P)$.
   Hint: note that $K(Q, P) = E_Q \log(dQ/dP) = E_P\{(dQ/dP) \log(dQ/dP)\}$.

4. (40 points) Suppose that $X_1, \ldots, X_n$ are i.i.d. with $E|X_1| < \infty$, and let $D_n \equiv n^{-1} \sum_{i=1}^n |X_i - \overline{X}_n|^{3/4}$. Use a uniform law of large numbers to show that $D_n \to_{a.s.} d \equiv E\{|X_1 - \mu|^{3/4}\}$ with $\mu = E(X_1)$.

Do **either** problem 5 **or** problem 6.

5. (40 points). Suppose that $X \equiv X_1, \ldots, X_n$ are i.i.d. F and $Y \equiv Y_1, \ldots, Y_n$ are i.i.d. G, and that we observe (Z_i, Δ_i), $i = 1, \ldots, n$, where $Z_i = X_i \wedge Y_i$, $\Delta_i = 1\{X_i \le Y_i\}$.
(a) Show how the cumulative hazard function $\Lambda = \Lambda_F$ corresponding to F can be expressed in terms of the sub-distribution functions

$$H^{uc}(t) = P(Z \le t, \Delta = 1), \qquad H^{c}(t) = P(Z \le t, \Delta = 0),$$

and the survival function $1 - H(t) = P(Z > t) = 1 - (H^{uc}(t) + H^{c}(t))$.
(b) Show how the result of (a) leads to a natural nonparametric maximum likelihood estimator of Λ, and hence of F via problem 2(c) above.

6. (40 points) Suppose that $X_1, \ldots, X_n$ are i.i.d. with mixture density

$$p(x; \lambda, \mu, \theta) = \left(\theta \frac{\lambda(\lambda x)^{r-1}}{\Gamma(r)} e^{-\lambda x} + (1-\theta) \frac{\mu(\mu x)^{s-1}}{\Gamma(s)} e^{-\mu x} \right) 1_{(0,\infty)}(x)$$

where $0 < \theta < 1$, $0 < \lambda \ne \mu < \infty$, and $r, s > 0$ are known. In other words, p is the mixture of two Gamma densities with parameters (r, λ) and (s, μ) respectively.
(a) Describe an EM - algorithm for estimation of (λ, μ, θ).
(b) What is a natural corresponding nonparametric model for the data which were modeled with the parametric mixture distribution in (a)? What is a natural nonparametric maximum likelihood estimator here?