Statistics 582, Midterm Exam

Wellner; 2/12/2010

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) A minimax decision rule d.
   (d) A Bayes decision rule d with respect to the prior distribution Λ.
   (e) An *integrable 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 and prove** *any two* of the following results:
   (a) A theorem concerning admissibility of Bayes rules (in the context of finite parameter, action, and sample spaces).
   (b) A formula expressing a survival function $1 - F(x) = P(X > x)$ of a non-negative random variable X in terms of the corresponding cumulative hazard function $\Lambda(x) = \int_{[0,x]} (1 - F(y-))^{-1} dF(y)$. Give the proof when F is continuous.
   (c) An inequality satisfied by the Kullback-Leibler information or divergence, $K(P, Q)$.
   (d) A relationship between the score for "incomplete" data Y and "complete" data X when $Y = T(X)$ for some measurable function T and when $\mathcal{P} = \{P_\theta : \theta \in \Theta \subset \mathbb{R}^d\}$ is a (regular) model for the complete data X.

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

3. (42 points) Suppose that $X_1, \ldots, X_n$ are i.i.d. with distribution function F on $\mathbb{R}^+$ and $Y_1, \ldots, Y_n$ are i.i.d. on $\mathbb{R}^+$ with distribution function G. We observe $(Z_i, \Delta_i) \equiv (X_i \wedge Y_i, 1_{[X_i \le Y_i]})$ for $i = 1, \ldots, n$, and our goal is to estimate the distribution function F (or, equivalently, the survival function $1 - F$.
(a) Describe the distribution of (Z_1, Δ_1) in terms of two sub-distribution functions expressed in terms of F and G, and give $P(Z_1 \ge z)$ in terms of F and G
(b) Relate the cumulative hazard function Λ of F to the two sub-distribution functions and survival function of Z_1 that you found in (a).
(c) Give the resulting natural nonparametric estimators of the three functions involved in (a) in terms of the observed data.
(d) Use the estimators in (c) to give the nonparametric MLE of the cumulative hazard function Λ.
(e) State a general formula expressing an arbitrary survival function $1 - F$ in terms of the corresponding hazard function Λ.
(f) Combine the results of (d) and (e) to give an explicit expression for the Kaplan-Meier (nonparametric maximum likelihood) estimator of $1 - F$.

4. (42 points) Suppose that $X \sim \text{Uniform}(0, 1)$.
(a) Find the hazard rate function $\lambda(t) = f(t)/(1 - F(t))$ where $f = f_X$ and $F(t) = F_X(t) = P(X \le t)$.
(b) Find the cumulative hazard function $\Lambda(t) = \int_{[0,t]} \lambda(s)ds$.
(c) State a general formula expressing a survival function $1 - F(t)$ to its corresponding cumulative hazard function Λ.
(d) Specialize the formula in (c) to the particular case in (a) and (b).
(e) Show that the formula in (c) continues to hold with F as in (a) replaced by $G = (1/2)F + (1/2)1_{[1/2,1]} = (1/2)F + (1/2)\delta_{1/2}$; compute the cumulative hazard function Λ_G corresponding to G, and show that the identity in (c) holds.

Do **one** of problems 5 - 7.

5. (36 points) Suppose that $X_1, \ldots, X_n$ are i.i.d. P_{θ_0} where $\mathcal{P} = \{P_\theta : \theta \in [1, \infty)\}$ and where P_θ is the Pareto(θ, α) distribution with density

$$p_\theta(x) = \frac{\alpha}{\theta}\left(\frac{\theta}{x}\right)^{\alpha+1} 1_{[\theta,\infty)}(x);$$

here α is assumed to be known.
(a) Find the MLE $\widehat{\theta}_n$ of θ.
(b) Give a proof of consistency of $\hat{\theta}_n$ using Wald's theorem and assuming that $\Theta = [1, M]$ for some $1 < M < \infty$.
(c) Give a direct proof that $\widehat{\theta}_n \to_p \theta_0$.

6. (36 points). Consider the (counter-)example given by Ferguson ACILST, pages 116-117 and Ferguson (1982):

$$p_\theta(x) = \frac{(1-\theta)}{\delta(\theta)} g_1\left(\frac{x-\theta}{\delta(\theta)}\right) + \theta g_0(x)$$

for $0 \le \theta \le 1$ where

$$\begin{aligned} g_1(x) &= (1-|x|)1_{[-1,1]}(x), \quad \text{the triangular density on } [-1,1], \\ g_0(x) &= 2^{-1}1_{[-1,1]}(x), \quad \text{the uniform density on } [-1,1], \\ \delta(\theta) &= (1-\theta)\exp(-(1-\theta)^{-c}+1) \end{aligned}$$

for some $c > 0$. Consider the class of functions $\mathcal{F} = \{\log p_\theta(x) - \log p_{\theta_0}(x) : \ \theta \in [0,1]\}$. Show that any envelope function F of the class $\mathcal{F}$ satisfies $EF(X) = \infty$ if $c \ge 1$.
Hint: Note that: (i) $p_{\theta_0}(x) \le p_{\theta_0}(\theta_0) =$ a constant; (ii) $\theta \mapsto p_\theta(x)$ is large when $\theta = \max\{0, x\}$. Use these two facts together with the form of δ to get a lower bound for

$$\sup_{\theta \in [0,1]} (\log p_\theta(x) - \log p_{\theta_0}(x))$$

which is not integrable under the stated condition.

7. (36 points) Suppose that $X_1, \ldots, X_n$ are i.i.d. with mixture density

$$p(x;\mu,\nu,\theta) = \frac{\theta}{2}\exp(-|x-\mu|) + \frac{1-\theta}{2}\exp(-|x-\nu|), \qquad x \in \mathbb{R},$$

where $0 < \theta < 1$, $\mu, \nu \in \mathbb{R}$, $\mu \ne \nu$. In other words, p is the mixture of two Laplace distributions with parameters μ and ν respectively.
(a) Describe an EM - algorithm for estimation of (μ, ν, θ).
(b) What hypothesis do you need to conclude that the EM algorithm in (a) converges to stationary points of the incomplete data log-likelihood? Does it hold in this case?