Statistics 583: Final Exam

Wellner; 6/5/2000

Instructions: This exam is to be done with closed notes and closed books.

1. (30 points). **Define** the following terms, providing an appropriate context for your definition:
   A. A continuous functional $T(F)$ with respect to a metric d on distributions.
   B. A Fréchet differentiable functional $T(F)$ with respect to a metric d.
   C. The cumulative hazard function corresponding to a distribution function F.
   D. A maximal invariant with respect to a group G.
   E. An invariant test with respect to a group G.
   F. A metric d between distribution functions which is "compatible" with respect to the empirical measure.

2. (27 points). **State** any *three* of the following four theorems, providing an appropriate context in each case:
   A. The Neyman-Pearson lemma.
   B. A Theorem about UMP $G-$invariant tests in the presence of monotone likelihood ratio of the $G-$maximal invariant.
   C. A theorem about asymptotically normality of the natural empirical estimator $T_n = T(\mathbb{F}_n)$ of a Fréchet - differentiable functional $T(F)$.
   D. Some version of *Hoeffding's formula* for the distribution of the ranks under alternatives.

3. (40 points). A simple test for asymmetry of a distribution function F is based on the difference of the mean and median: $T(F) = \int x dF(x) - F^{-1}(1/2)$. Note that $T(F) = 0$ if F is symmetric about some point, while $T(F)$ is positive for F skewed to the right, and negative for F skewed to the left. Suppose that $X_1, \ldots, X_n$ are i.i.d. with distribution function F, and let $\mathbb{F}_n$ be the empirical distribution of the sample, $\mathbb{F}_n(x) = n^{-1} \sum_{i=1}^n 1_{(-\infty,x]}(X_i)$.
   A. Show that T is invariant under location shifts: If F_θ is defined by $F_\theta(x) = F(x-\theta)$, then $T(F_\theta) = T(F)$.
   B. Under appropriate assumptions on F (make these explicit), find the influence function of $T(F)$. [Hint: we already essentially did this in class, since we calculated the influence functions of $F^{-1}(1/2)$ and $\int x dF(x)$.]
   C. Under appropriate assumptions on F (make these explicit), state a limit theorem for $\sqrt{n}(T(\mathbb{F}_n) - T(F))$ and compute the asymptotic variance $V^2(F)$.
   D. Consider the bootstrap and jackknife estimators of $V_n(F) \equiv Var_F(\sqrt{n}(T(\mathbb{F}_n) - T(F)))$. Would either or both of these "work" for estimation of $V^2(F)$? Explain why or why not.

4. (40 points). Consider testing $H : F = G$ versus $K : (1 - G) = (1 - F)^\theta$ where $\theta > 1$ based on $X_1, \ldots, X_m$ i.i.d. F and $Y_1, \ldots, Y_n$ i.i.d. G. We want a test which is invariant with respect to monotone transformations of the data.
A. Draw a picture to show the relationship of F and G under the alternative hypothesis K. What is the relationship of the corresponding cumulative hazard functions under K? Draw a picture showing this.
B. Give an appropriate version of Hoeffding's formula for the distribution of the order Y-ranks $Q_1 < \ldots < Q_n$ under the alternative K.
C. Use the formula in B (or any other way) to compute the complete distribution of $\underline{Q}$ when $m = n = 2$ (so $N = m + n = 4$) and $\theta = 2$. For testing H versus $\theta = 2$, what observed value of $\underline{Q}$ would lead you to reject H at size $\alpha = 1/6$?
D. What statistic would you use for testing H versus K? In the case $m = n = 2$ in part C, what are its possible values? What observed value of $\underline{Q}$ leads you to reject H at size $\alpha = 1/6$?

5. (30 points). Suppose that X_{ij} for $j = 1, \ldots, n_i$, $i = 1, \ldots, I$ are independent, normally distributed random variables with common variance σ^2, and suppose that $EX_{ij} = \theta_i$. Thus $X_{i1}, \ldots, X_{i,n_i}$ is a sample of size n_i from the $N(\theta_i, \sigma^2)$ distribution. Consider testing $H : \theta_1 = \ldots = \theta_I$ versus $K : \theta_i \neq \theta_j$ for some $i \neq j$.
A. Explain briefly the canonical form of this testing problem, and the groups which leave the testing problem invariant.
B. Find the UMP-invariant test of H versus K, and specify its distribution under the null hypothesis H.
C. What is the distribution of the test statistic you found in B under the general hypothesis K?

6. (40 points). Suppose that $X_1, \ldots, X_m$ are i.i.d. exponential(λ), and that $Y_1, \ldots, Y_n$ are i.i.d. exponential(μ); thus the density of X_1 is $\lambda \exp(-\lambda x) 1_{[0,\infty)}(x)$. Consider testing $H : \lambda \leq \mu$ versus $K : \lambda > \mu$.
A. Show that this testing problem is invariant with respect to the group of scale changes, G given by $g_c(\underline{x}, \underline{y}) = (c\underline{x}, c\underline{y})$ where $c > 0$.
B. Find the UMP G-invariant test of H versus K. [Hint: you may use the fact that the family of distributions $\{\delta^{-1} F_{r,s} : \delta > 0\}$ has monotone decreasing likelihood ratio.]
C. Specify exactly how you would carry out the test derived in B.
D. Find a UMP - unbiased test of H versus K. Is this the same test as in part B or a different test?