Mathematical Statistics (281ABC)
Qualifying Exam, May 19, 2011

Problem 1 Consider i.i.d. data X_1, ..., X_n from a parametric family of distributions $\mathcal{P}_\theta$, where θ is a unknown scalar parameter, i.e. of dimension one. Denote $f(x;\theta)$ either the probability function or the density function. Let $\hat{\theta}$ be the maximum likelihood estimator (MLE). Give, as complete as possible, a set of conditions under which

(a) the MLE $\hat{\theta}$ exists (*5 points*);

(b) $\hat{\theta}$ is consistent for the true value of θ (*5 points*);

(c) $\hat{\theta}$ is asymptotically normal. (*5 points*).

Problem 2 Now consider i.i.d. data (X_1, δ_1), ..., (X_n, δ_n) where the $X_i = \min(T_i, C_i)$ is possibly right-censored, and $\delta_i = I(T_i \leq C_i)$. Assume that C_i follows the random censorship assumption. Again denote $f(t;\theta)$ the density function for the distribution of T_i.

(a) Write down the likelihood function $L(\theta)$ for this data; (*5 points*)

(b) Give conditions under which the MLE $\hat{\theta}$ is consistent and asymptotically normal. (*5 points*)

(c) Choose a favorite parametric distribution of your own for $f(t;\theta)$, find $\hat{\theta}$; (*5 points*)

(d) Estimate the variance of $\hat{\theta}$. (*5 points*)

Problem 3 Consider the term 'asymptotic relative efficiency' (ARE); it appears a few times during our course.

(a) Give its definition in the context of point estimation; (*5 points*)

(b) Give its definition (i.e. Pitman ARE) in the context of hypothesis testing. (*5 points*)

(c) Now consider the score test based on the likelihood theory, show that it's fully efficient asymptotically; (*5 points*)

(d) Finally consider the problem of the weighted log-rank tests for the G^ρ family, where we derived the ARE. Under further assumptions that the censoring distributions are the same in the two comparison groups, the asymptotic efficacy for the G^ρ weighted log-rank test under the alternative which is given by the $G^{\rho'}$ distribution simplifies to

$$(\mu_\rho/\sigma_\rho)^2 = \frac{\{\int_0^\infty \pi(u)S(u)^{\rho+\rho'}d\Lambda(u)\}^2}{\int_0^\infty \pi(u)S(u)^{2\rho}d\Lambda(u)},$$

where $\pi(u) = P(T \geq u, C \geq u)$ is the probability of being 'at risk' at time u for any member of the two groups under the null hypothesis, and $S(\cdot)$ and $\Lambda(\cdot)$ are the survival function and the cumulative hazard function, respectively, again all under the null hypothesis.

Can you deduce the following from the above expression: [**Hint**: these parts do not require that you remember any of the formulas we talked about related to the G^ρ family or during the derivation of the ARE]

i. for given alternative distribution $G^{\rho'}$, the maximum efficacy of the G^ρ test is achieved when $\rho = \rho'$; (*5 points*)

ii. if we further assume that the survival function for the censoring distribution $S_c(u) = S(u)^\alpha$, then the asymptotic efficacy becomes $(2\rho + \alpha + 1)(2\rho' + \alpha + 1)/(\rho + \rho' + \alpha + 1)^2$; (*5 points*)

iii. use the above from part ii. to compute the ARE of Peto's Wilcoxon log-rank test (i.e. G^1) versus the unweighted log-rank test under the proportional hazards alternative assuming $\alpha = .5$ and $\alpha = 0$. (*5 points*)

Problem 4 Using likelihood ratio test, show that the Pearson's test for $r \times c$ contingency table has approximately chi-squared distribution with the correct number of degrees of freedom under the null hypothesis that the column and row variables are independent. (*15 points*)

Question V (Please use separate sheets of paper for this one.)

(a) Define what's meant by saying a sequence of random variables (X_n) is bounded in probability (or "tight", i.e. $X_n = O_p(1)$). Also define what's meant by another sequence $Y_n = o_p(X_n)$ as $n \to \infty$.

(b) Prove that if $X_n = O_p(1)$ and $Y_n = o_p(X_n)$ then $Y_n = o_p(1)$.

(c) State and prove the theorem known as the delta method.

(d) $X_1, \dots, X_n \sim$ iid $\text{Pois}(\lambda)$. Compute the ARE $(\hat{\lambda}_2^2, \hat{\lambda}_1^2)$, where $\hat{\lambda}_1$ is the UMVUE of λ, and $\hat{\lambda}_2 = -\log\left(\frac{1}{n}\sum_i 1[X_i = 0] + 1\right)$

(The target function of the squared estimators $\hat{\lambda}_i^2$ is of course λ^2. Also note the "+1" in $\hat{\lambda}_2$ simply prevents taking the log of 0, which would occur anyway with probability $\to 0$.)

(e) How would you tell, from much simpler considerations, that $\hat{\lambda}_2$ is $\sqrt{n}$-consistent?

(f)* Suppose $W_1, \dots, W_n \sim$ iid. $f_\theta \in \{f_\theta\}$, a location family.

Show that you can base a consistent estimator for θ on $X_{(n)}$ alone (the largest order statistic), provided there are sequences (a_n) and (b_n) with $b_n = o(1)$, such that $\left\{\frac{X_{(n)} - a_n}{b_n}\right\} = O_{f_0}(1)$.

Show that if $\text{var}\, X_{(n)} \to 0$ the condition $b_n = o(1)$ can be guaranteed.

(For interest only: this is a very fine tail condition on the shape f_0; the uniform on $[0,1]$ has it (easy), the normal $(0,1)$ has it (tricky), and the exponential (1) does not have it (easy).)