Eric Slud Mar. 26, 2014

Topics & Sample Problems for S701 In-Class Test

The following are the topics that are covered and emphasized on the In-Class Test to be held 3/31/14.

• Modes of probabilistic convergence, rules for manipulating O_P, o_P, and consistency of estimators.

• CLT, including multivariate and Lindeberg/Liapunov versions

• Method of Moments and MLE's in Exponential Families

• Fisher Information and MLE Asymptotic Distribution Theory, including corollaries on Bayesian and One-Step Estimators

• Relative efficiency of Generalized Method of Moment Estimators

• Exact Moments and Asymptotic Distribution of (1-Sample) U-statistics

---

Here are some sample problems of the kind that might be asked on the test. There will be 3 or 4 problems of this general level of difficulty (with fewer parts each) on the exam.

(1) Two samples of data are observed, with a shared unknown parameter μ,

$$X_1, \ldots, X_{50} \sim \text{Poisson}(\lambda \cdot \mu) \quad , \qquad Y_1, \ldots, Y_{50} \sim \mathcal{N}(\mu, \sigma^2)$$

and nuisance parameters μ, σ^2 also unknown. Find the MLE of λ based on the combined $\mathbf{X}$ and $\mathbf{Y}$ samples, and find its asymptotic variance.

(2) Define **asymptotic variance** of a sequence of estimators $T_n = T_n(X_1, \ldots, X_n)$ for which there exist sequences a_n, b_n such that $b_n\,(T_n - a_n)$ converge in distribution to a nondegenate limiting distribution. Allthough the sequences a_n, b_n are not uniquely defined for T_n, show that the asymptotic variance is.

(3) Suppose that random pairs $(W_i,\, X_i)$, iid across $i = 1, \ldots, n$, are distributed such that

$$X_i \sim \mathcal{N}(0,1) \quad , \qquad W_i \sim \text{Expon}(\lambda\, e^{a\, X_i}) \quad \text{given} \quad X_i$$

(a) Find the large-sample distribution of the MLE $\hat{\lambda}$.

(b) Find the method of moments estimator of λ, and find its relative efficiency.

(c) Give a formula for an efficient estimator of λ. (*The likelihood equations for the MLE are explicit, but the MLE itself is not.*)

(4) Based on a sample $X_1, \ldots, X_n$ of scalar random variables with a continuous distribution function F, define the U-statistic U_n with kernel $h(x_1, x_2) \equiv I_{[x_1 \cdot x_2 < 0]}$. Tell what parameter ϑ (defined in terms of the unknown F) is unbiasedly estimated by U_n, and find the nondegenerate large-sample distribution of $U_n - \vartheta$ scaled by an appropriate multiplicative constant sequence.

(5) In a Bernoulli(p_0) sample $Y_1, \ldots, Y_{1000}$, with sample size $n = 1000$, it is observed that $\hat{p} = \bar{Y} = 0.43$. If you know that $p_0 = 0.40$, then give and justify a good approximation to the score statistic value $S_n(p_0)$ at the true parameter value, and explain what is the order of accuracy of the approximation.

(6) Suppose you have a likelihood function $L_n(\theta, \mathbf{X})$ based on a sample $\mathbf{X}$ of data of size n, where θ is a scalar unknown parameter. Suppose that by some special feature of the density of the X_i that the log-likelihood function has the known property that on an interval $J = (\theta_0 - \epsilon,\, \theta_0 + \epsilon)$ (where $\epsilon > 0$ does not depend on n, and n becomes as large as you like)

$$\sup_J \, \left|\frac{\partial^2}{\partial \theta^2} \log L(\theta) + I(\theta_0)\right| \; < \; \frac{1}{10} \, |I(\theta_0)|$$

where $I(\theta_0) > 0$, and always with iid data we know that

$$\text{argmax}_J \, \frac{1}{n} \, E(\log L_n(\theta)) \; = \; \theta_0$$

Then show that there is a unique RLE (Root of the Likelihood Equation $\partial L_n(\theta)/\partial\theta \; = \; 0$) estimator $\hat{\theta}_n$ within the interval J, and that this sequence of estimators is consistent.