Statistics 582, Problem Set 1

Wellner; 1/4/2006

Reading: Chapter 4, Sections 4-6; Lehmann and Casella, TPE, section 6.6, pages 469 - 483; Ferguson, ACLST, Chapter 16-17, pages 107-118.
Due: Wednesday, January 11, 2006.

1. Suppose that $(X, Y), (X_1, Y_1), \ldots, (X_n, Y_n)$ are i.i.d. with bivariate normal distribution $N_2(\mu, \Sigma)$ where $\mu \in R^2$ and

$$\Sigma = \begin{pmatrix} \sigma^2 & \sigma\tau\rho \\ \sigma\tau\rho & \tau^2 \end{pmatrix}$$

where $\sigma^2 > 0$, $\tau^2 > 0$, and $\rho \in (-1, 1)$.
A. If we assume that $\mu_1 = \mu_2 \equiv \theta$ and Σ is known, what is the MLE of θ?
B. If we assume that μ is known and $\sigma^2 = \tau^2 \equiv \theta$, what is the MLE of θ?
C. What is the asymptotic distribution of the estimator you found in B?
D. Under the same assumption as in B, what is the MLE of ρ?
E. What is the asymptotic distribution of the estimator you found in D?

2. Problem 1, page 117, Ferguson, ACILST. What happens if $\Theta = [1, \infty)$ or $(0, \infty)$?

3. Suppose that $X, X_1, \ldots, X_n$ are i.i.d. Weibull(α_0, β_0) (if X has the Weibull(θ) distribution where $\theta = (\alpha, \beta)$, then $1 - F_\theta(x) = P_\theta(X > x) = \exp(-(x/\alpha)^\beta)$ for $x \geq 0$). Recall that the MLE $\widehat{\alpha}$ of α is given by

$$\widehat{\alpha} = \left\{ \frac{1}{n} \sum_{i=1}^n X_i^{\widehat{\beta}} \right\}^{1/\widehat{\beta}}$$

where $\widehat{\beta}$ is the MLE of β. As a simpler alternative to maximum likelihood, I propose to use the alternative estimator $\overline{\beta}$ of β obtained from the slope of an ordinary least squares fit of a Weibull Q-Q plot, and then estimate α by

$$\overline{\alpha}_n = \left\{ \frac{1}{n} \sum_{i=1}^n X_i^{\overline{\beta}} \right\}^{1/\overline{\beta}} .$$

A. Suppose that $\overline{\beta}_n \rightarrow_p \beta_0$ is known. Show that $\overline{\alpha}_n \rightarrow_p \alpha_0$. [Hint: use a uniform strong law of large numbers.]
B. Show that $\overline{\alpha}_n$ is a "pseudo-MLE" in the sense that $\overline{\alpha}_n$ maximizes $l_n(\alpha, \overline{\beta}_n)$.

4. (Profile likelihood) [For nice plots to accompany this exercise, see pages 41 - 43 of Cox, D. R. and Oakes, D. (1984); *Analysis of Survival Data*, Chapman and Hall.] As in problem 1.3, consider the Weibull family of example 3.2.5: $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ with $\Theta \subset \mathbb{R}^{+2}$ given by the (Lebesgue) densities

$$p_\theta(x) = \frac{\beta}{\alpha} (\frac{x}{\alpha})^{\beta - 1} \exp(-(\frac{x}{\alpha})^\beta) 1_{[0,\infty)}(x)$$

where $\theta \equiv (\alpha, \beta) \in (0, \infty) \times (0, \infty) \subset \mathbb{R}^2$.
A. For a sample of n observations from p_θ, we know that, for each fixed value of β the value of α which maximizes the likelihood as a function of α is

$$\widehat{\alpha}(\beta) = \{\frac{1}{n}\sum_{i=1}^{n} X_i^{\beta}\}^{1/\beta}.$$

Use this to compute the *profile likelihood* $l_{\text{profile}}(\beta) = l_{\text{profile}}(\beta|\underline{X})$ defined by

$$l_{\text{profile}}(\beta) = l(\widehat{\alpha}(\beta), \beta) = l(\widehat{\alpha}(\beta), \beta|\underline{X}).$$

B. Use what we know from problem 10.3 to show that the profile likelihood is strictly concave and hence has a unique maximum. Show that maximizing the profile likelihood as a function of β yields the maximum likelihood estimate: i.e. that $(\widehat{\alpha}, \widehat{\beta}) = (\widehat{\alpha}(\widehat{\beta}_{\text{profile}}), \widehat{\beta}_{\text{profile}})$.

C. What is the relationship of the score function for β from the profile likelihood, $\dot{l}_{\beta,\text{profile}}$ to the (efficient) score for β from the full likelihood? Prove or disprove my claim: the profile score for β (based on n observations) is asymptotically equivalent to the sum of efficient scores for β over the sample in the sense that their difference divided by $\sqrt{n}$ converges to 0 in probability.
D. What is the relationship of the observed information from the profile likelihood $-\ddot{l}_{\beta\beta,\text{profile}}$ to information quantities from the full likelihood?

5. **Optional bonus problem.** On pages 116-117 of ACILST, Ferguson (see also Ferguson, T. S. (1982). An inconsistent maximum likelihood estimate. *J. Amer. Statist. Assoc.* **77**, 831–834) shows that $\hat{\theta}_n \to_{a.s.} 1$ no matter what θ_0 is true if $\delta(\theta) \to 0$ "fast enough".
(a) Show that $\hat{\theta}_n \to_{a.s.} 1$ continues to hold if

$$\delta(\theta) = (1-\theta)\exp(-(1-\theta)^{-c} + 1)$$

with $c > 2$. (Ferguson shows that $c = 4$ works.)
(b) Show that when $c = 2$, Ferguson's argument yields

$$\sup_{0 \le \theta \le 1} n^{-1} \log L_n(\theta) \ge \frac{n-1}{n}\log(M_n/2) + \frac{1}{n}\log\frac{1-M_n}{\delta(M_n)} \to_d D$$

where

$$P(D \le y) = \exp\left(-\frac{1}{2(y - \log 2)}\right), \quad y \ge \log(2).$$

That is, $D \stackrel{d}{=} \log 2 + 1/(2E)$ where E is an Exponential(1) random variable.