Statistics 582, Problem Set 1

Wellner; 1/5/2011

Reading: Chapter 6, Sections 4-6 and 10; Lehmann and Casella, TPE, pages 457-461 and 515-519. Ferguson, ACLST, Chapter 16-17, pages 107-118.
Due: Wednesday, January 12, 2011.

1. Lehmann and Casella, problem 3.15, page 502.

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

3. Consider the model introduced in Ferguson, ACILST, problem 17.2, page 117:

$$p(x|\theta) = 2\left(\frac{x}{\theta} 1_{[0,\theta]}(x) + \frac{1-x}{1-\theta} 1_{(\theta,1]}(x)\right), \quad \theta \in [0,1].$$

Show that Theorem 4.3, page 28, of the Chapter 4 notes (or Theorem 17, Ferguson, ACILST, page 114) applies to the MLE of θ in this model.

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.] Consider the Weibull family of example 3.2.5 (581 Course Notes) $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ 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 Statistics 581 problem 9.2 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 $(\widehat{\alpha}, \widehat{\beta}) = (\widehat{\alpha}(\widehat{\beta}_{\text{profile}}), \widehat{\beta}_{\text{profile}})$.

5. 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 $\overline{\beta}_n$ of β obtained from the slope of an ordinary least squares fit of a Weibull Q-Q plot, and then estimate α by

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

(a) Suppose that $\overline{\beta}_n \to_p \beta_0$ is known. Show that $\overline{\alpha}_n \to_p \alpha_0$. [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 $l_n(\alpha, \overline{\beta}_n)$.

6. **Optional bonus problem 1.** (a) Suppose that $X_1, \ldots, X_n$ are i.i.d. with distribution P on R. Consider generalizing the result discussed in class on 1/3/2011 for $r = 1$: if $V_n(r)$ is defined for $1 \le r \le 2$ by

$$V_n(r) \equiv \frac{1}{n} \sum_{i=1}^{n} |X_i - \overline{X}_n|^r.$$

If $E|X|^r < \infty$, show that

$$V_n(r) \to_{a.s.} v(r)$$

where $v(r) \equiv E|X_1 - \mu|^r$.
(b) Now suppose we generalize the problem considered in (a) by considering $X_1, \ldots, X_n$ i.i.d. P on R^d. Let $\|\cdot\|$ be the usual Euclidean metric in R^d, and consider

$$V_n(r) \equiv \frac{1}{n} \sum_{i=1}^{n} \|X_i - \overline{X}_n\|^r$$

for $1 \le r \le 2$ where $\overline{X}_n$ is the (multivariate) sample mean of the X_i's. Can the same method be used to show that $V_n(r) \to_{a.s.} v(r)$ where $v(r) \equiv E\|X_1 - \mu\|^r$ (assuming that $E\|X_1\|^r < \infty$)?

7. **Optional bonus problem 2.** This is a continuation of problem 5 above.
(a) 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.
(b) 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?

8. **Optional bonus problem 3.** 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.
(c) What hypothesis in Wald's consistency theorem is violated in this example?