

## Statistics 581, Problem Set 9

Wellner; 11/19/2014

**Reading:** Chapter 4, Sections 1-2;

Ferguson, ACLST, Chapters 18- 20, pages 119-125, 133-139; Chapter 22, pages 144-150; Lehmann and Casella, Chapter 6, especially section 6.5, pages 461-468.

**Due:** Wednesday, November 26, 2014.

1. Consider the Weibull family of example 3.2.5 and problem set #6, problem 1:  $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$  with  $\Theta \subset R^{+2}$  given by the (Lebesgue) densities

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

where  $\theta \equiv (\alpha, \beta) \in (0, \infty) \times (0, \infty) \subset R^2$ . Suppose that  $X, X_1, \dots, X_n$  are i.i.d. with density function  $p_\theta$ .

(a) If  $X \sim P_\theta \in \mathcal{P}$ , show that the distributions of  $\log X$  form a location and scale family from a Gumbel (extreme value) density on  $R$ . (This amounts to a rephrasing of the statement of problem 1 in problem set 6.)

(b) Use the result of (a) to construct method of moments estimators or quantile based estimators  $\bar{\theta}_n$  of  $\theta = (\alpha, \beta)$ .

(c) Show that the method of moments or quantile estimators  $\bar{\theta}_n$  of  $\theta$  are asymptotically normal, and find the asymptotic distribution; i.e. show that

$$\sqrt{n}(\bar{\theta}_n - \theta) \rightarrow_d N_2(0, \Sigma) \quad \text{for some } \Sigma.$$

[We will use these estimators as “starting points” approximate (or one-step) maximum likelihood estimators in the next problem.]

2. (Problem #1, continued).

(a) Does a maximum likelihood estimate of  $\hat{\theta} = (\hat{\alpha}, \hat{\beta})$  exist? Is it unique? (See Lehmann and Casella, Example 6.1, page 468.)

(b) Compute an approximate (one - step) maximum likelihood estimate  $\check{\theta}$  of  $\theta$  using the method of moment (or quantile) estimators  $\bar{\theta}_n$  as the preliminary estimators based on the following data (with  $n = 12$ ):

1, 1, 2, 3, 12, 21, 46, 54, 65, 109, 317, 413.

[These are failure times in seconds for “breakdown” of an insulating fluid between two electrodes subject to a voltage of 40 kV. – from Nelson, *Applied Life Data Analysis*, page 252, modified slightly.]

(c) Compute the maximum likelihood estimator  $\hat{\theta}_n$ , and compare it with the one step estimator computed in (b).

3. (a) Ferguson, ACILST, problem 17.2, page 117: I would suggest modifying Ferguson's definition of the density to:

$$p_\theta(x) \equiv f(x|\theta) = 2 \left\{ \frac{x}{\theta} 1_{[0,\theta]}(x) + \frac{1-x}{1-\theta} 1_{(\theta,1]}(x) \right\}.$$

- (b) Do our hypotheses A0-A2 hold in this example?  
 (c) Compute  $K(P_{\theta_0}, P_\theta)$  where  $P_\theta$  has density as given in this problem.  
 (d) Do our hypotheses A3 and A4 hold in this example? Why or why not?  
 (e) Does there exist an estimator  $\bar{\theta}_n$  of  $\theta$  which is  $n^{1/2}$ -consistent?
4. (a) Lehmann and Casella, problem 6.3.1, page 501.  
 (b) Lehmann and Casella, problem 6.3.2, page 501.  
 (c) Lehmann and Casella, problem 6.3.4, page 501.

5. **Optional bonus problem 1:** Suppose that  $X_1, \dots, X_n$  are i.i.d. log-normal( $\mu, \sigma^2$ ) with density

$$p_\theta(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right) 1_{(0,\infty)}(x).$$

Here  $\theta = (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}^+ \equiv (-\infty, \infty) \times (0, \infty)$ .

- (a) Find the MLE  $\hat{\theta} = (\hat{\mu}, \hat{\sigma}^2)$  of  $\theta = (\mu, \sigma^2)$ .  
 (b) Show that  $\log X \stackrel{d}{=} \mu + \sigma Z \sim N(\mu, \sigma^2)$  where  $Z \sim N(0, 1)$ .  
 (c) Suppose that  $\nu(P_\theta) = q(\theta) = E_\theta(X)$ . Express  $q(\theta)$  explicitly as a function of  $\theta$ .  
 (d) Suggest a natural nonparametric estimator  $\bar{\nu}_n$  of  $E_\theta(X)$ .  
 (e) Find the asymptotic variance of  $\sqrt{n}(\bar{\nu}_n - \nu(P_\theta))$  for the estimator  $\bar{\nu}_n$  you proposed in (d).  
 (f) What is the MLE  $\hat{\nu}_n$  of  $\nu = \nu(P_\theta)$  assuming that the log-normal model is true? What do our results in chapter 3 say about the asymptotic distribution of  $\sqrt{n}(\hat{\nu}_n - \nu(P_\theta))$  (assuming that the model holds)?  
 (g) Compare the variances you found in (e) and (f). Which estimator do you prefer if the log-normal model holds?
6. **Optional bonus problem 2:** Lehmann and Casella, TPE, problem 6.3.22, page 503, reworded as follows. (In other words, prove (vi) of theorem 1.2, pages 5-6, chapter 4 notes). Suppose that  $X_1, \dots, X_n$  are i.i.d. with density  $p_\theta$ ,  $\theta \in \Theta \subset \mathbb{R}^k$ , satisfying the hypotheses of theorem 4.1, page 463 (the Cramér conditions given in (A) - (D) on pages 462-463). Show that the following Local Asymptotic Normality (LAN) result holds for the (local) log-likelihood ratios: with

$$L_n(\theta) \equiv \log\left(\prod_{i=1}^n p_\theta(X_i)\right) = \sum_{i=1}^n \log p_\theta(X_i),$$

for a fixed  $\theta_0 \in \Theta$ ,

$$\begin{aligned} L_n(\theta_0 + n^{-1/2}\underline{t}) - L_n(\theta_0) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \underline{t}^T \dot{\ell}_\theta(X_i) - \frac{1}{2} \underline{t}^T I(\theta_0) \underline{t} + o_p(1) \\ &\rightarrow_d N(0, \underline{t}^T I(\theta_0) \underline{t}) - \frac{1}{2} \underline{t}^T I(\theta_0) \underline{t} \\ &\stackrel{d}{=} N(-\sigma^2/2, \sigma^2) \end{aligned}$$

under  $P_{\theta_0}$  where  $\sigma^2 \equiv \underline{t}^T I(\theta_0) \underline{t}$ . (The convergence in the last display actually holds under the considerably weaker hypothesis of Hellinger differentiability of  $p_\theta$  at  $\theta_0$ , as stated in Corollary 3 of section 3.3, page 28, of the Chapter 3 notes.)

7. **Optional bonus problem 3:** Suppose that  $(Y|Z) \sim \text{Poisson}(\lambda e^{\gamma Z})$ , and  $Z \sim \text{Bernoulli}(\eta)$ , and  $\theta = (\lambda, \gamma, \eta)$ . Let  $X = (Y, Z)$ , and suppose that we observe  $X_1, \dots, X_n$  i.i.d. as  $X$ .
- Find the score equations for estimation of  $\theta$ .
  - Give conditions on the data  $X_1, \dots, X_n = (Y_1, Z_1), \dots, (Y_n, Z_n)$  guaranteeing that the score equations have a unique solution which maximizes the likelihood. Call the resulting estimators  $\hat{\theta}_n = (\hat{\lambda}_n, \hat{\gamma}_n, \hat{\eta}_n)$ .
  - What does theorem 4.1.2 (Chapter 4, page 5), say about the asymptotic distribution of  $\sqrt{n}(\hat{\theta} - \theta_0)$  when the distribution of the data is given by  $P_{\theta_0}$ ?
  - Suppose that  $\theta_1 \neq \theta_0$  is the “true” value of the parameter  $\theta$ , and we consider the likelihood ratio  $L_n(\theta_1)/L_n(\theta_0)$  where  $L_n(\theta) \equiv \prod_{i=1}^n p_\theta(X_i)$ . Show that  $n^{-1} \log(L_n(\theta_1)/L_n(\theta_0)) \rightarrow_p$  some constant, and identify the constant explicitly in terms of  $\theta_1, \theta_0$ .