Statistics 581, Problem Set 9

Wellner; 11/21/2018

Reading: Chapter 4, Sections 1-2;
Ferguson, ACLST, Chapters 18- 20, pages 119-125, 133-139; Chapter 22, pages 144-150;
vdV, Asymp. Statist., pages 41 - 75; Sections 5.1 - 5.7.
Due: Wednesday, November 28, 2018.

1. Suppose that $X_1, \ldots, X_n$ are i.i.d. Geometric(θ) random variables; that is, $P_\theta(X_1 = k) = \theta(1-\theta)^{k-1}$ for $k = 1, 2, \ldots$ where $\theta \in (0, 1)$.
   (i) Show that the Geometric distribution with parameter θ satisfies the conditions of Lemma 7.6 of van der Vaart (1998), page 95.
   (ii) Compute the information for θ.
   (iii) Suggest three different estimators of θ based on the data.
   (iv) Which of your estimators are asymptotically efficient in the sense of Hájek's convolution theorem?

2. Repeat problem 1 above when the X_i's are i.i.d. Cauchy(θ); i.e. each of the X_i's has the common density $p_\theta(x) = f(x - \theta)$ where $f(x) = 1/(\pi(1 + x^2))$ for $x \in \mathbb{R}$ and $\theta \in \mathbb{R}$.

3. Ferguson, problem 4, page 124: Let X and Y be independent random variables with densities p_θ and q_θ depending on θ. Assume that the Fisher informations $I_X(\theta)$ and $I_Y(\theta)$ for θ based on observing X or Y both exist. Show that the Fisher information for θ based on observing the pair (X, Y) is given by $I_{X,Y}(\theta) = I_X(\theta) + I_Y(\theta)$.

4. Consider the Laplace location family $p_\theta(x) = 2^{-1}\exp(-|x-\theta|)$ for $x \in \mathbb{R}$ and $\theta \in \mathbb{R}$.
   (a) Does the hypothesis (M5) of Theorem 3.2.22, page 11 of the Course Notes hold in this case? Does the hypothesis (M4) of Theorem 3.2.22. hold?
   (b) Show that the Laplace location family is differentiable in quadratic mean. What is the consequence of this for the behavior of the local log-likelihood ratios? What is the resulting information for the location parameter θ?
   (c) Apply the methods of section 3.5 to show that with $\theta_0 \in \mathbb{R}$ fixed and $\theta_n = \theta_0 + n^{-1/2}h$, then for any estimator T_n of θ we have
   $$\liminf_{n\to\infty} \inf_{T_n} \max\{E_{\theta_n} n|T_n - \theta_n|^2,\ E_{\theta_0} n|T_n - \theta_0|^2\} \geq cI(\theta_0)^{-1}$$
   for some choice of h and an absolute constant c.

5. Ferguson, problem 6, page 125: What was thought to be a certain species of moth is attracted to a capture tank at rate λ per day. On the first day, the number X of moths caught was recorded. It is assumed that X has a Poisson distribution with mean λ. Later it was pointed out that this species is, in fact, two different similar species, so a second day of capture was undertaken. This time, the numbers Y_1 and Y_2 of moths caught of these species separately were noted. It is assumed that these are Poisson random variables with means λ_1 and λ_2 where $\lambda_1 + \lambda_2 = \lambda$, and it is assumed that X, Y_1, and Y_2 are independent.
   (a) Using X, Y_1, and Y_2, find the maximum likelihood estimate of λ_1 and λ_2.
   (b) Assuming λ_1 and λ_2 large, what is the approximate variance of your estimator?

6. **Optional bonus problem 1:** Find the form of the score functions for a location-scale family $p_\theta(x) = f((x-\mu)/\sigma)$ with $\theta = (\mu, \sigma) \in \mathbb{R} \times \mathbb{R}^+$. Then apply Lemma 7.6 of van der Vaart (1998) to find a sufficient condition for differentiability in quadratic mean.

7. **Optional bonus problem 2:** Consider the shift (or location) family based on the Gamma$(\beta, 1)$ density. That is, $p_\theta(x) = f(x - \theta)$ for $\theta \in \mathbb{R}$ where

$$f(x) = \frac{1}{\Gamma(\beta)} x^{\beta-1} \exp(-x) 1_{(0,\infty)}(x)$$

for some $\beta > 0$. For what values of β is the family $\{p_\theta : \theta \in \mathbb{R}\}$ differentiable in quadratic mean (or Hellinger differentiable in the terminology of the course notes)? What happens if differentiability in quadratic mean "just barely fails"?

8. **Optional bonus problem 3:** Suppose that X has density p_θ and that the Fisher information for θ, $I_X(\theta)$ based on observing X exists. What is the Fisher information for θ based on observation of $Y = g(X)$ where g is a measurable map from the sample space $\mathcal{X}$ of X to the sample space $\mathcal{Y}$ for Y?