University of Illinois
Spring 1996

Econ 476 Final Exam

Department of Economics
Roger Koenker

Please answer all 5 of these questions.

1. Let $X_1, \ldots, X_n$ be a random sample from the density $f(x, \theta_0)$ and $\hat{\theta}_n$ denote the maximum likelihood estimator of θ_o, i.e., $\hat{\theta}_n$ maximizes

$$l_n(\theta) = \sum_{i=1}^{n} \log f(X_i, \theta)$$

over $\theta \in \Theta \subseteq \mathbb{R}^p$. Presuming sufficient regularity that $\hat{\theta}_n$ is efficient in the sense that

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \rightsquigarrow \mathcal{N}(0, I(\theta_0)^{-1})$$

where $I(\theta_0)$ denotes the Fisher Information of f. Sketch the argument for fact that

$$2(l_n(\hat{\theta}) - l_n(\theta_0)) \rightsquigarrow \chi^2_p$$

2. Suppose $X_1, \ldots, X_{10}$ constitute a random sample from the density

$$f(x) = \frac{\theta^3}{2} e^{-\theta x} x^2 \qquad x > 0.$$

Knowing only that $\bar{X} = 2$, find an estimate of θ, and test the hypothesis that $\theta = 1$. Briefly, try to explain any special virtues of the procedures you use.

3. Consider the following 5 estimators of the location parameter θ

$$\begin{aligned}
\hat{\theta}_1 &= n^{-1} \sum X_i \\
\hat{\theta}_2 &= F_n^{-1}(1/2) \\
\hat{\theta}_3 &= \text{argmin} \{-\sum \log(1 + (X_i - \theta)^2)\} \\
\hat{\theta}_4 &= \hat{\theta}_2 - \left[\sum_{i=1}^{n} \frac{2(X_i - \hat{\theta}_2)^2 - 2}{(1 + (X_i - \hat{\theta}_2)^2)^2}\right]^{-1} \sum_{i=1}^{n} \frac{2(X_i - \hat{\theta}_2)}{(1 + (X_i - \hat{\theta}_2)^2)} \\
\hat{\theta}_5 &= \hat{\theta}_2 + \frac{4}{n} \sum_{i=1}^{n} \frac{X_i - \hat{\theta}_2}{(1 + (X_i - \hat{\theta}_2)^2)}
\end{aligned}$$

based on a random sample from the Cauchy density

$$f(x) = \frac{1}{\pi(1 + (x - \theta)^2)}.$$

(a) Explain briefly all the estimators concentrating on $\hat{\theta}_4, \hat{\theta}_5$. Where did these expressions come from? What is the difference between them? Would it matter if we replaced $\hat{\theta}_2$ by $\hat{\theta}_1$ in these expressions?

(b) Compare and contrast the asymptotic performance of the 5 estimators.

Hint:

$$\pi^{-1} \int_{-\infty}^{\infty} \frac{2(x-\theta)^2 - 2}{(1+(x-\theta)^2)^3} dx = -\frac{1}{2}$$

4. Consider the model,

$$\log y_i = x_i'\beta + u_i$$

with $\{u_i\}$ iid from a common distribution function F.

(a) Explain what the quantile regression estimator,

$$\hat{\beta}(\tau) = \text{argmin}_{\beta \in \mathbf{R}^p} \sum_{i=1}^{n} \rho_\tau(\log y_i - x_i'\beta)$$

is estimating, and enough about its asymptotic behavior to justify this claim.

(b) Now consider

$$\check{\beta}(\tau) = \text{argmin}_{\beta \in \mathbf{R}^p} \sum_{i=1}^{n} \rho_\tau(y_i - \exp(x_i'\beta))$$

compare and contrast the two estimators: do they estimate the same quantity? in what ways might one be preferable to another?

5. In regular parametric models with random sampling, the variance of an estimator $\hat{\theta}$ of the parameter, θ, from the common density $f(z, \theta)$ satisfies,

$$V(\hat{\theta}) \geq V(\partial \log f(Z, \theta)/\partial\theta)/n$$

a.) Explain briefly the significance of this result when $\theta \in \mathbf{R}^1$,

b.) Explain how to interpret the result if $\theta \in \mathbf{R}^p$,

c.) Explain the circumstances in which equality is attained in this expression,

d.) Give an example of an irregular case in which the inequality is violated, or becomes meaningless.