Eric Slud Fall 2007

Solutions to Stat 710 Problem Set 4

#7.1. Here the dominating measure μ on $\mathbf{N} = \{0, 1, 2, \ldots\}$ is counting measure. For each $x \geq 0$, the function

$$s(x, \theta) \equiv \sqrt{p(x, \theta)} = \exp((x \log \theta - \theta)/2)/\sqrt{x!}$$

is obviously continuously differentiable on $\theta \in (0, \infty)$. Then $\dot{p}_\theta / p_\theta = x/\theta - 1$, and (using the change of index $k = x - 1$ as needed)

$$I_\theta = \sum_{x \geq 0} (\frac{x}{\theta} - 1)^2 \, \frac{\theta^x}{x!} \, e^{-\theta} = 1 - 2\frac{\theta}{\theta} + \sum_{k \geq 0} \frac{k+1}{k! \theta^2} \theta^{k+1} e^{-\theta} \; = \; \frac{1}{\theta}$$

which is well-defined and continuous.

#7.5. Now X_i are *iid* $\mathcal{N}(\theta, 1)$ which implies

$$log \prod_{i=1}^n \frac{p_{\theta + h_n/\sqrt{n}}}{p_\theta}(X_i) \; = \; \frac{1}{2} \sum_{i=1}^n \Big\{ (X_i - \theta)^2 - (X_i - \theta - \frac{h_n}{\sqrt{n}})^2 \Big\} = h_n \sqrt{n} \bar{X} - \frac{h_n^2}{2} - \sqrt{n} \frac{\theta h_n}{2}$$

Comparing to the expression in Theorem 7.2, the term $o_{P_\theta}(1)$ is explicitly

$$(h_n - h)\sqrt{n}(\bar{X} - \theta) \; - \; \frac{1}{2}(h_n^2 - h^2) \; = \; (h_n - h)\Big(\sqrt{n}(\bar{X} - \theta) - \; \frac{h_n + h}{2} \Big)$$

#7.6. Now $f(x, \theta) = e^{-|x - \theta|}/2$, and we cannot apply Lemma 7.6 because the condition there does not hold for *every* x. But with $g_\theta(x)$ defined as the a.e. derivative of $\log f(x, \theta)$, that is, $g_\theta(x) = sgn(x - \theta)$, we find (using the change of variable $w = -v$ on the negative half-line)

$$\int \Big[\sqrt{\frac{1}{2} e^{-|x - \theta - h|}} \; - \; (1 + \frac{h}{2} sgn(x - \theta)) \sqrt{\frac{1}{2} e^{-|x - \theta|}} \Big]^2 dx \; = \qquad (*)$$

$$\int_0^\infty \Big(e^{(v - |v - h|)/2} - 1 - \frac{h}{2} \Big)^2 \frac{e^{-v}}{2} \, dv \; + \; \int_0^\infty \Big(e^{w - |w + h|)/2} - 1 + \frac{h}{2} \Big)^2 \frac{e^{-w}}{2} \, dw$$

The first of these integrals is split into two terms, an integral on $[0, h)$ and another on (h, ∞), the first equal to

$$\int_0^h (e^{v - h/2} - 1 - \frac{h}{2})^2 \, e^{-v} \, \frac{dv}{2} \; = \; O\Big(\int_0^h (v - \frac{h}{2} + \frac{1}{2}(v - h/2)^2 - h/2 \Big)^2 e^{-v} dv \Big) \; = \; O(h^3)$$

and the second equal to

$$\int_h^\infty \Big(e^{h/2} - 1 - h/2 \Big)^2 e^{-v} \, \frac{dv}{2} \; = \; e^{-h} \, (e^{h/2} - 1 - h/2)^2 \; = \; O(h^4)$$

The second integral on the right-hand side of (*) is

$$\int_0^\infty \left(e^{-h/2} - 1 + h/2\right)^2 e^{-w}\,\frac{dw}{2} \;=\; \frac{1}{2}\,(e^{-h/2} - 1 + h/2)^2 \;=\; O(h^4)$$

Thus we have shown that (*) is $o(h^2)$ as $h \to 0+$, completing the verification of quadratic mean differentiability.

#7.10. Here as elsewhere we try to define $g_\theta(x)$ as the a.e.(x) gradient of the log density. It is not hard to see that when x approaches ξ from above, the integrands (including terms from the gradient $g_\theta(x)$ (or from its square) increase as a power $(x-\xi)^{-k}$ with k no larger than 4, while the exponential term $\exp(-\frac{1}{2\sigma^2}(\log(x-\xi)-\mu)^2)$ decreases faster than any power of $x-\xi$. Thus convergence of integrals is no problem. It is also not hard to see the quadratic mean differentiability with respect to the μ, σ^2 variables (by bounding the Taylor series remainders in the density), so we show it only for the ξ variable. The expression to be proved $o(h^2)$ is

$$\int \Big\{ \Big(\frac{1}{x-\xi-h}\, I_{[x>\xi+h]}\, \exp(-\frac{1}{2\sigma^2}(\log(x-\xi-h)-\mu)^2 \Big)^{1/2} -$$

$$(1+\frac{h}{2(x-\xi)}(1+\frac{\log(x-\xi)-\mu}{\sigma^2})\Big(\frac{1}{x-\xi}\, I_{[x>\xi]}\, \exp(-\frac{1}{2\sigma^2}(\log(x-\xi-h)-\mu)^2\Big)^{1/2}\Big\}^2 dx$$

Without loss of generality, we consider it only for $h>0$; perform the change of variable $y = \log(x-\xi)$; and split the resulting integral into one on $(-\infty, C\sqrt{\log h})$ and one on $(C\sqrt{\log h}, \infty)$ for a large constant C to be specified. The first of these three integrals is easily seen to be

$$O\Big(\int_{-\infty}^{C\sqrt{\log h}} \exp(-\frac{(y-\mu)^2}{2\sigma^2} - 3y)\,dy\Big) = O((\log h)^{-2}\, e^{-(C\sqrt{\log h}-\mu)^2/(2\sigma^2)}) = o(h^2)$$

as long as $C > 10\sigma^2$, and the integrand in the second of these integrals involves Taylor series terms which can readily be bounded $o(h^2)$ multiplied by a normal density.

This problem required some very detailed analysis for a fully rigorous proof, but the result is that quadratic mean differentiability *does hold* for the 3-parameter lognormal.

#8.3. For fixed x, the usual Central Limit Theorem implies for a general distribution F that

$$\sqrt{n}\,(F_n(x) - F(x)) \;\xrightarrow{\mathcal{D}}\; \mathcal{N}(0,\, F(x)(1-F(x)) \qquad \text{as} \qquad n\to\infty$$

For the specific case of $X_i \sim \mathcal{N}(\theta, 1)$, we have $F(x) = \Phi(x-\theta)$. In this case, the delta method immediately implies (with $\phi(\cdot)$ denoting the standard normal density)

$$\sqrt{n}\,\Big(\Phi(x-\bar{X}) \;-\; \Phi(x-\theta)\Big) \;\xrightarrow{\mathcal{D}}\; \mathcal{N}(0,\, \phi^2(x-\theta)) \qquad \text{as} \qquad n\to\infty$$

As a result, the Asymptotic Relative Efficiency of the estimators is

$$ARE(F_n(x),\, \Phi(x - \bar{X})) \;=\; \phi^2(x-\theta)/(\Phi(x-\theta)(1-\Phi(x-\theta)))$$

Although it is analytically not easy to bound, a graph of the very smooth function $G(z) = \phi(z)/(\Phi(z)(1 - \Phi(z))$ shows it to be even and unimodal (a fact which *can* be analytically proved without much trouble), peaked in the middle with a maximum value of .6366 at 0. So if one wants to estimate a distribution function value not too far from the mean, the empirical d.f. does so with a loss of effciency equivalent to throwing away about 1/3 of the data in a normal-distribution setting, although the situation becomes much worse far from the mean.