Statistics 582, Problem Set 6

Wellner; 2/11/2015

Reading: Chapter 6, sections 1-2.

Due: Wednesday, February 18, 2015.
Reminders: Midterm exam, Friday, February 13;
No office hours Wednesday, February 11;
replaced by Thursday, 8:30-10:30 & 2:30 - 4:30.

1. Continuation of problem 2, problem set 4: Suppose that $X_1, \ldots, X_n$ are i.i.d. Exponential(θ) (so the X's have distribution P_θ and density $p_\theta(x) = \theta e^{-\theta x} 1_{(0,\infty)}(x)$. with respect to Lebesgue measure on $\mathbb{R}$, and that $\theta \sim \Gamma(\alpha, \beta)$:

$$\lambda(\theta) = \beta \frac{(\beta\theta)^{\alpha-1}}{\Gamma(\alpha)} \exp(-\beta\theta) 1_{[0,\infty)}(\theta).$$

In problem set 4 we found the Bayes rules with respect to squared error loss $L(\theta, a) = (\theta - a)^2$ and weighted squared error loss $L(\theta, a) = (\theta - a)^2/\theta$.

(a) Prove a (conditional) limit theorem for the posterior distributions given $\underline{X}$.
(b) What does theorem 5.8.2 say about the limiting distribution of the Bayes rule for squared error loss (assuming that $X_1, \ldots, X_n$ are i.i.d. $P_{\theta_0} \equiv P$ with $\theta_0 \in (0, \infty)$)?
(c) Can you say anything about the asymptotic behavior of the Bayes rule with respect to weighted squared error loss?

2. Specialize the decision rule in Theorem 5.2 of the course notes to the case when P_i is the normal distribution $N_d(\mu_i, I)$, $i = 1, \ldots, k$ where $\mu_1, \ldots, \mu_k$ are distinct vectors in $\mathbb{R}^d$, $\mu_i \neq \mu_j$ for $i \neq j$. What happens if we replace I by Σ?

3. Lehmann and Casella, TPE, Problem 5.17, page 293, parts (a) and (c). (Also note Problems 5.18, 5.19, 5.20, page 293.)

4. **Optional bonus problem 1:** (Birgé). Let $X = (X_0, X_1, \ldots, X_k)$ be a $(k+1)-$dimensional vector, and assume that $X \sim N_{k+1}(\theta, I_{k+1})$ where I_{k+1} denotes the $(k+1) \times (k+1)$ identity matrix. For any vector $\theta \in \mathbb{R}^{k+1}$, let θ' denote the projection of θ onto the $k-$dimensional linear space spanned by the $k-$last coordinates. Consider the subset Θ_0 of $\Theta = \mathbb{R}^{k+1}$ given by

$$\Theta_0 = \{\theta \in \mathbb{R}^{k+1} : |\theta_0| \leq k^{1/4} \text{ and } \|\theta'\| \leq 2(1 - k^{-1/4}|\theta_0|)\}.$$

(a) Show that the MLE of θ over Θ_0 is given by $\widehat{\theta}_0 = 0$ and $\widehat{\theta}' = 2X'/\|X\|$ *on the event*

$$\Omega_0 \equiv \{\|X'\|^2 > 3k/4 \quad \text{and} \quad |X_0| < k^{1/4} + 1.21\}.$$

(b) Show that $P_\theta(\Omega_0) \geq 3/4$ for all $\theta \in \Theta_0$.

(c) Let $\widetilde{\theta} = (X_0, \underline{0})$. Show that for $k \geq 128$ we have

$$\sup_{\theta \in \Theta_0} E_\theta \|\theta - \hat{\theta}\|^2 \geq (3/4)\sqrt{k} + 3, \quad \text{and}$$
$$\sup_{\theta \in \Theta_0} E_\theta \|\theta - \widetilde{\theta}\|^2 \leq 5.$$

Thus the maximal risk of the MLE may be much larger than the minimax risk when k is large.
Hint: A non-central χ_k^2 distribution is stochastically larger than a central χ_k^2 distribution; then use Lemma 1 of Laurent and Massart (2000) [Laurent, B.; Massart, P. Adaptive estimation of a quadratic functional by model selection. *Ann. Statist.* **28** (2000), 13021338].

5. **Optional bonus problem 2:** Problem 3.9, Lehmann and Casella, TPE, page 286.