TOPICS IN STATISTICAL THEORY – EXAMPLES 2/2

Part III, Lent 2019, RN (email: r.nickl@statslab.cam.ac.uk)

Ex. 1 a) For $g_k \sim^{i.i.d.} N(0,1)$, denote by P_θ^Y the law in $\mathbb{R}^p$ of the random vector $(Y_k = \theta_k + g_k/\sqrt{n} : k = 1, \dots, p)$. Show that P_0^Y is a common dominating measure for P_θ^Y and any $\theta \in \mathbb{R}^p$, and that the Radon-Nikodym derivative is almost surely equal to

$$\frac{dP_\theta^Y}{dP_0^Y}(y) = \exp\{n\langle\theta, y\rangle_{\mathbb{R}^p} - \frac{n}{2}\|\theta\|^2_{\mathbb{R}^p}\}.$$

b*) Show that the conclusions in a) remain true if P_θ^Y is the law of an observation $(Y_k = \theta_k + g_k/\sqrt{n} : k \in \mathbb{N})$ in the Gaussian sequence space model, for any $\theta \in \ell_2$, and if the inner product and norm of $\mathbb{R}^p$ are replaced by the ones of ℓ_2.

Ex. 2 Consider a product prior distribution Π on ℓ_2 with each coordinate drawn independently as $\theta_k \sim N(0, k^{-1-2\alpha}), k \in \mathbb{N}, \alpha > 0$.

a) Denote the Sobolev norm by $\|\theta\|^2_{H^\beta} = \sum_k k^{2\beta}\theta_k^2$. For which β do we have $E^\Pi\|\theta\|^2_{H^\beta} < \infty$? Show that $\Pi(\theta \in \ell_2) = 1$.

b) Show that the posterior distribution $\Pi(\cdot|Y)$ given observations $Y \sim P_\theta^Y, \theta \in \ell_2$, in the Gaussian sequence space model, is also Gaussian and find its mean vector $E^\Pi[\theta|Y]$ and covariance. Show further that the distribution of $\theta|Y - E^\Pi[\theta|Y]$ does not depend on the data Y.

c) Assume that a penalised least squares estimator $\hat\theta(Y)$ maximising

$$2\langle\theta, Y\rangle_{\ell_2} - \|\theta\|^2_{\ell_2} - \frac{1}{n}\|\theta\|^2_{H^{\alpha+1/2}}$$

over $\theta \in \ell_2$ exists. Show that $\hat\theta(Y) = E^\Pi[\theta|Y]$ almost surely.

Ex. 3 Let $\theta_0, \theta_1 \in \ell_2$ be such that $\|\theta_0 - \theta_1\|_{\ell_2} \geq \epsilon$ for some $\epsilon > 0$. Suppose you are given observations $Y \sim P_\theta^Y$ in the Gaussian sequence space model. Show that the test

$$\Psi_n = \mathbf{1}\{2\langle\theta_1 - \theta_0, Y\rangle_{\ell_2} > \|\theta_1\|^2_{\ell_2} - \|\theta_0\|^2_{\ell_2}\}$$

satisfies, for all $n \in \mathbb{N}, \epsilon > 0$,

$$E^Y_{\theta_0}\Psi_n \leq e^{-n\epsilon^2/8}, \qquad \sup_{\theta:\|\theta-\theta_1\|_{\ell_2}\leq\epsilon/4} E^Y_\theta(1-\Psi_n) \leq e^{-n\epsilon^2/32}.$$

Ex. 4 Consider a random vector $\theta = (\theta_k, k = 1, \dots, N)$ in ℓ_2 with entries drawn independently as uniform $U(-1,1)$ random variables. Let θ_0 be such that $\|\theta_0\|^2_{H^\alpha} = \sum_{k\in\mathbb{N}}\theta_k^2 k^{2\alpha} \leq 1$ for some $\alpha > 0$. Show that for $N = n^{1/(2\alpha+1)}$ and $\varepsilon_n = c\sqrt{\log n} \times n^{-\alpha/(2\alpha+1)}$, any $c > 0$, we can find a constant $C = C(c, \alpha)$ such that for all n large enough

$$\Pr(\|\theta - \theta_0\|_{\ell_2} < \varepsilon_n) \geq e^{-Cn\varepsilon_n^2}.$$

Deduce that the posterior distribution arising from observations in the Gaussian sequence space model contracts about θ_0 at rate ε_n in $P^Y_{\theta_0}$-probability.

Ex. 5 Suppose a posterior distribution on $\Theta \subseteq \ell_2$ satisfies

$$\Pi(\theta \in \Theta : \|\theta - \theta_0\|_{\ell_2} > \varepsilon_n | Y) \to^P 0 \quad \text{as } n \to \infty$$

for some sequence $\varepsilon_n \to 0$. Show that if $\hat{\theta}_n$ is the centre of the smallest ball that contains posterior mass at least 1/2, then $\|\hat{\theta}_n - \theta_0\|_{\ell_2} = O_P(\varepsilon_n)$ as $n \to \infty$. [For random variables Z_n we write $Z_n = O_P(\varepsilon_n)$ if $\forall \delta > 0$ there exists M_δ such that $P(\varepsilon_n^{-1}|Z| > M_\delta) < \delta$ for all n lage enough.]

Ex. 6 Let X be a Poisson random variable with parameter $\lambda = e$. Show that for every $k \in \mathbb{N}$ large enough and a numerical constant c_0,

$$\Pr(X > k) \leq e^{-c_0 k \log k}, \quad \Pr(X = k) \geq e^{-k \log k}.$$

Ex. 7 Let D_n be a sequence of measurable sets whose posterior probability $\Pi(D_n|Y) \to^P 1$ as $n \to \infty$. [Here P is the 'frequentist' law of the data Y and it is assumed that the posterior distribution arises from a dominated likelihood model]. Let $\Pi^{D_n} = \Pi(\cdot \cap D_n)/\Pi(D_n)$ denote the prior restricted to D_n and renormalised, and let $\Pi^{D_n}(\cdot|Y)$ be the new posterior distribution arising from the prior Π^{D_n} and observations Y. Show that

$$\sup_{A \text{ measurable}} |\Pi^{D_n}(A|Y) - \Pi(A|Y)| \to^P 0 \quad \text{as } n \to \infty.$$

Ex. 8 Let $X_1, \ldots, X_n$ be i.i.d. from some law P_0 of probability density function p_0 on $[0, 1]$ and let $\mathcal{P}$ be a set of probability densities p that are all absolutely continuous with respect to p_0. Let ν be a probability measure on

$$B = \{p \in \mathcal{P} : E_{P_0} \log(p_0/p) \leq \varepsilon^2, E_{P_0}(\log p/p_0)^2 \leq \varepsilon^2\}, \quad \varepsilon > 0.$$

Show that for all $c > 0, \varepsilon > 0$ and $n \in \mathbb{N}$ we have

$$P_0^n \left(\int_B \prod_{i=1}^n \frac{p(X_i)}{p_0(X_i)} d\nu(p) \leq e^{-(c+1)n\varepsilon^2} \right) \leq \frac{1}{c^2 n \varepsilon^2}.$$

Ex. 9 * Let $(P_n : n \in \mathbb{N})$ be a sequence of *random* probability measures on $\mathbb{R}$ (defined on a suitable probability space $(\Omega, \mathcal{A}, \Pr)$ such that $P_n(\omega)$ are Borel probability measures on $\mathbb{R}$ for each fixed $\omega \in \Omega$ and $n \in \mathbb{N}$), and let P be a fixed (non-random) probability measure on $\mathbb{R}$. Show that if

$$E_{P_n} e^{tX} \to^{\Pr} E_P e^{tX} \quad \forall t \in \mathbb{R},$$

as $n \to \infty$ then also

$$E_{P_n} f(X) \to^{\Pr} E_P f(X)$$

for every bounded continuous function $f : \mathbb{R} \to \mathbb{R}$ and deduce further that, if P is Gaussian, then

$$\sup_{s \in \mathbb{R}} |P_n((-\infty, s]) - P((-\infty, s])| \to^{\Pr} 0 \quad \text{as } n \to \infty.$$