Some exercices

1 On posterior concentration rates

1.1 Construction of L_1 tests

Let f_0 and f_1 be probability densities on $\mathbb{R}$ and $A = \{x \in \mathbb{R}; f_0(x) \geq f_1(x)\}$ with respect to a measure μ.

1. Show that
$$\|f_0 - f_1\|_1 = 2\int_A (f_0(x) - f_1(x))d\mu(x)$$

2. Let P_0 be the distribution under f_0 and P_f under f and $X^n = (X_1, \cdots, X_n)$ a n sample with $\mathbb{P}_n$ the empirical distribution :
$$\mathbb{P}_n(B) = n^{-1}\sum_{i=1}^n \mathbb{1}_{X_i \in B}$$
Show, using Hoeffding inequality, that
$$P_0^n\left([\mathbb{P}_n - P_0](A) \geq \delta_n\right) \leq 2e^{-2n\delta_n^2}$$

3. How can we choose δ_n so that
$$\sup_{\|f-f_1\|_1 \leq \zeta\|f_0-f_1\|_1} P_f^n\left([\mathbb{P}_n - P_0](A) < \delta_n\right) \leq 2e^{-nc\|f_0-f_1\|_1^2}$$
for some $c > 0$ and $\zeta < 1$?

4. Let $\mathcal{F}_n \subset \mathcal{F} = \{f : \mathbb{R} \to \mathbb{R}^+; \int_{\mathbb{R}} f(x)d\mu(x) = 1\}$ such that
$$N(\zeta\epsilon; \{\|f - f_0\|_1 \in (\epsilon, 2\epsilon)\} \cap \mathcal{F}_n, \|.\|_1) \leq an\epsilon^2$$
for all $\epsilon \geq \epsilon_n$ for some sequence $\epsilon_n = o(1)$, then show that one can construct a test ϕ_n such that
$$E_0^n[\phi_n] \leq e^{-nc_1\epsilon_n^2}, \qquad \sup_{\|f-f_0\|_1 > M\epsilon_n, f\in\mathcal{F}_n} E_f^n[1 - \phi_n] \leq e^{-nc_1\epsilon_n^2}$$

1.2 Regression and sieve priors

Consider the model
$$Y_i = f(x_i) + \epsilon_i, \quad \epsilon_i \overset{iid}{\sim} \mathcal{N}(0, \sigma^2), \quad x_i \in (0,1), \quad i \leq n$$

- Prior for f
$$f := f_\theta = \sum_{j=1}^k \theta_j\phi_j, \quad (\phi_j)_j = \text{BON of} \quad L^2([0,1])$$

$$k \sim \pi_k, \quad (\theta_1, \cdots, \theta_k)|k \overset{iid}{\sim} g$$

- prior for σ: π_σ.

$$f_0 \in L^2[0,1]$$

1. Compute the Kullback-Leibler divergence between P_0^n and P_θ^n
2. Assume that θ_0 is in Sobolev ball

$$\sum_{j=1}^{+\infty} (1+j)^{2\beta} \theta_{0j}^2 \leq L < +\infty$$

   and that $\sigma_0 > 0$ is fixed.

   Show that under random design with appropriate iid distribution for the x_i's

$$S_n = \{KL(\theta_0, \theta) \leq (n/\log n)^{1/(2\beta+1)}; V_2(\theta_0, \theta) \lesssim (n \log n)^{1/(2\beta+1)}\}$$

   contains

$$\{\theta \in \mathbb{R}^{k_n}; \|\theta - \theta_{[0k]}\|_2 \lesssim (n/\log n)^{-\beta/(2\beta+1)} k_n^{-1/2}\}, \quad k_n \asymp (n/\log n)^{1/(2\beta+1)}$$

3. Deduce that if g is positive and continuous on $\mathbb{R}$ and if π_σ is positive and continuous on $\mathbb{R}^+*$ and

$$\pi(S_n) \geq e^{-cn^{1/(2\beta+1)}}$$

   for some $c > 0$ when n is large enough.

4. The construction of tests can be done using individual tests f_0 versus f_{θ_1} using the likelihood ratio statistic.

2 Empirical Bayes

Consider the mixture model for monotone nonincreasing densities

$$f_P(x) = \int_0^\infty \frac{\mathbb{1}_{[0,\theta)}(x)}{\theta} dP(\theta)$$

assume that f_0 is decreasing : $\exists P_0$ s.t.

Consider a prior distribution on $\mathcal{P} = \{P, \text{ proba on } \mathbb{R}^+\}$ as a $DP(M, \Gamma(a,b))$

1. Find the transformation $\psi_{b,b'}$ such that if $P \sim DP(M, \Gamma(a,b))$ then $P \sim DP(M, \Gamma(a,b'))$
2. Same thing with $\psi_{a,a'}$.
3. Consider $X^n = (X_1, \cdots, X_n)$ i.i.d f_P ; study

$$\sup_{b' \in [b, b(1+u_n)]} (\ell_n(\psi_{b,b'}(P)) - \ell_n(P)), \qquad \inf_{b' \in [b, b(1+u_n)]} (\ell_n(\psi_{b,b'}(P) - \ell_n(P))$$

4. Why is it more complicated with $\psi_{a,a'}$?