
Exercise sheet 3

Exercise 1 Suppose that $\Theta \subseteq \mathbb{R}^d$ and $X_1, \dots, X_n \overset{iid}{\sim} F_{\theta_0}$ for some $\theta_0 \in \Theta$, and

1. θ_0 is the unique maximiser of the **continuous** function ℓ.
2. For all M, $\sup_{\|\theta\| \le M} |\bar{\ell}_n(\theta) - \ell(\theta)| \xrightarrow{p-P_{\theta_0}} 0$.
3. For any $\epsilon > 0$ there exists $M_\epsilon < \infty$ such that $\sup_n \mathbb{P}_{\theta_0}(\|\widehat{\theta}_n^{MLE}\| > M_\epsilon) < \epsilon$.

Show that $\widehat{\theta}_n^{MLE} \xrightarrow{p-P_{\theta_0}} \theta_0$. **Hint:** first show an inequality of the form $P_{\theta_0}(\|\widehat{\theta}_n^{MLE} - \theta_0\| > \epsilon) \le 2\epsilon$ for all $\epsilon > 0$ and all $n \ge N_\epsilon$ large. Then show that this implies the convergence in probability.

Exercise 2 The second part of this question is not for the exam.

(a) **(equivariance of maximum likelihood estimators)**. Consider a model F_θ with $\theta \in \Theta$ and let $h : \Theta \to h(\Theta)$ be injective. Define $\phi = h(\theta)$ and consider the model $G_\phi = F_{h^{-1}(\phi)}$. Show that $\widehat{\phi}^{MLE} = h(\widehat{\theta}^{MLE})$.

(b) **(*invariance of maximum likelihood estimator with respect to the dominating measure)** Recall that f_θ is the Radon–Nikodym derivative of F_θ with respect to a σ-finite measure μ. Suppose that μ' is another measure that dominates μ, and that we replace f_θ by the $g_\theta = \partial dF_\theta / \partial \mu'$. Show that this does not change the maximum likelihood estimator. Deduce that if μ'' is another measure that dominates all the F_θ (but not necessarily μ), then the maximum likelihood estimators with respect to μ and with respect to μ'' are the same. **Hint:** recall that $g_\theta(x) = f_\theta(x)h(x)$, where $h(x) = \partial d\mu(x)/\partial d\mu'(x)$ is the Radon–Nikodym derivative and does not depend on θ, and $\mu(\{x : h(x) = 0\} = 0$.

Exercise 3 We say that a model $f(x;\theta)$ is a k-parameter exponential family in natural parametrisation if $\Theta \subseteq \mathbb{R}^k$ and

$$f(x;\theta) = \exp\left(\sum_{j=1}^k \theta_i T_j(x) - \gamma(\theta) + s(x)\right).$$

Assume that Θ has a nonempty interior and that the covariance matrix of $T(X) = (T_1(X), \dots, T_k(X))$ is nonsingular for all $\theta \in int(\Theta)$.

Find the maximum likelihood estimator for θ based on a sample $X_1, \dots, X_n$ from $f(x;\theta_0)$ with $\theta_0 \in int(\Theta)$ and show that it is consistent and asymptotically normal. You may use without proof that $\mathbb{E}_{\theta_0} T(X) = \nabla\gamma(\theta_0)$ and $var_{\theta_0} T(X) = \nabla^2\gamma(\theta_0)$. **Hint:** use the delta method and the inverse function theorem.

Exercise 4

(a) Let G be any absolutely continuous distribution on the positive real line and let F_λ be the Exp(λ) distribution. Find the value of $\lambda > 0$ that minimises the KL divergence between G and F_λ? Under which condition is λ unique?

(b) Now let G be an arbitrary distribution on $\mathbb{R}$. Find the values of μ and σ^2 that minimise $\mathrm{KL}(G, N(\mu, \sigma^2))$. Under which conditions are these values unique? Can you generalise this to higher dimensions?

Exercise 5 Suppose that $S \sim \mathrm{Exp}(\lambda)$ and $C \sim \mathrm{Exp}(\gamma)$ are independent and define $T = \min(S, C)$ and $D = 1[T = S]$. Assume we have independent and identically distributed observations (T_i, D_i), $i = 1, \ldots, n$. Find the maximum likelihood estimator of the vector $(\lambda, \gamma)^\top$ and show that it is consistent and asymptotically normal. It is **not** required to compute the asymptotic covariance matrix.