Principles of Statistics	**Part II - Michaelmas 2018**

Example Sheet 2

Lecturer: Quentin Berthet

Here, unless specified otherwise, θ is assumed to be a fixed element of the parameter space $\Theta \subset \mathbb{R}^p$. By a 'regular parametric model' we mean a statistical model of probability density/mass functions $\{f(\cdot, \theta) : \theta \in \Theta\}$ that satisfies the regularity conditions from lectures, ensuring asymptotic normality of the maximum likelihood estimator (which in turn may be used without proof in the solution of the respective exercise).

1. Let $\Theta \subseteq \mathbb{R}$ have nonempty interior and let S_n be a sequence of random real-valued continuous functions defined on Θ such that, as $n \to \infty$, $S_n(\theta) \to^P S(\theta)$ $\forall \theta \in \Theta$, where $S : \Theta \to \mathbb{R}$ is nonrandom. Suppose for some θ_0 in the interior of Θ and every $\varepsilon > 0$ small enough we have $S(\theta_0 - \varepsilon) < 0 < S(\theta_0 + \varepsilon)$, and that S_n has *exactly one* zero $\hat{\theta}_n$ for every $n \in \mathbb{N}$. Deduce that $\hat{\theta}_n \to^P \theta_0$ as $n \to \infty$.

2. Give an example of (possibly random) functions Q_n, Q defined on $\Theta \subset \mathbb{R}$ that have unique maximisers $\hat{\theta}_n, \theta_0$, respectively, such $Q_n(\theta) \to Q(\theta)$ (almost surely) for every $\theta \in \Theta$ as $n \to \infty$, but $\hat{\theta}_n \not\to \theta_0$ (almost surely).

3. Consider an i.i.d. sample $X_1, \ldots, X_n$ arising from the model

$$\left\{f(x, \theta) = \theta x^{\theta - 1} \exp\{-x^\theta\}, x > 0, \theta \in (0, \infty)\right\}$$

of *Weibull distributions*. Show that the MLE exists and is consistent.

4. Consider the maximum likelihood estimator $\hat{\theta}$ from $X_1, \ldots, X_n$ i.i.d. $\mathcal{N}(\theta, 1)$ where $\theta \in \Theta = [0, \infty)$. Show that $\sqrt{n}(\hat{\theta} - \theta)$ is asymptotically normal whenever $\theta > 0$. What happens when $\theta = 0$? Comment on your findings in light of the general asypmtotic theory for maximum likelihood estimators.

5. Let $X_1, \ldots, X_n$ be i.i.d. random variables from a uniform $U(0, \theta), \theta \in \Theta = (0, \infty)$ distribution. Calculate the Fisher information for this model. Find the maximum likelihood estimator $\hat{\theta}$ of θ and show that $\tilde{\theta}_n = \frac{n+1}{n}\hat{\theta}$ is unbiased for θ. Find the variance of $\tilde{\theta}_n$, compare it to what the Cramèr-Rao inequality predicts, and discuss your findings. Finally find the asymptotic distribution of $n(\hat{\theta} - \theta)$.

6. Suppose one is given a regular parametric model $\{f(\cdot, \theta) : \theta \in \Theta\}$ with likelihood function $L(\theta)$ and corresponding maximum likelihood estimator $\hat{\theta}_{MLE}$, and consider a mapping $\Phi : \Theta \to F$, where Θ, F are subsets of $\mathbb{R}$.

a) Assuming that Φ is one to one, show that the maximum likelihood estimator of ϕ in the model $\{f(\cdot,\phi) : \phi = \Phi(\theta) \text{ for some } \theta \in \Theta\}$ equals $\Phi(\hat\theta_{MLE})$.

b) Now consider a mapping Φ that is not necessarily one-to-one. Define the induced likelihood function $L^*(\phi) = \sup_{\theta:\Phi(\theta)=\phi} L(\theta)$ and show that $\Phi(\hat\theta_{MLE})$ is a maximum likelihood estimator of ϕ (that is, show that $\Phi(\hat\theta_{MLE})$ maximises $L^*(\phi)$).

c) Based on n repeated observations of a random variable X from one of the following parametric models, find the maximum likelihood estimator of the parameter ϕ: i) $\phi = Var(X)$ in a Poisson-θ model. ii) $\phi = Var(X)$ in a Bernoulli-p-model, iii) $\phi = (EX)^2$ in a $\mathcal{N}(\mu,\sigma^2)$ model. Which of these MLEs are unique?

7. Consider the parameter $\phi = \mathbf{E}[X^4]$ equal to the fourth moment of a $\mathcal{N}(0,\theta)$ distribution. Find the MLE $\hat\phi$ of ϕ and derive the asymptotic distribution of $\sqrt{n}(\hat\phi - \phi)$ as $n \to \infty$. Conjecture a corresponding result for higher moments $\phi_m = \mathbf{E}[X^m]$ where $m > 4$ is an even integer.

8. In a regular parametric model with parameter space $\Theta \subset \mathbb{R}^d$, let $\hat\theta$ be the maximum likelihood estimator arising from an i.i.d. sample $X_1, \dots, X_n$. Derive the asymptotic distribution of

$$W_n = n(\hat\theta - \theta)^T i_n(\hat\theta - \theta)$$

under P_θ, where i_n equals either $i_n(\theta)$ or $i_n(\hat\theta_n)$ and where $i_n(\theta)$ is the observed Fisher information matrix at θ. Deduce from this limiting result i) a test for the hypothesis $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$ that has type-one-errors of level at most α and ii) that the confidence ellipsoid

$$C_n = \{\theta \in \mathbb{R}^d : (\hat\theta - \theta)^T i_n(\hat\theta)(\hat\theta - \theta) \le z_\alpha/n\}$$

has asymptotic coverage level $1-\alpha$ for z_α the $1-\alpha$-quantile constants of the limit distribution derived above.

9. Consider the parametric models from Exercise on Sheet 1 with corresponding parameter space Θ. For all these models, derive explicit expressions for the likelihood ratio test statistic of a simple hypothesis $H_0 : \theta = \theta_0$, $\theta_0 \in \Theta$ vs. $H_1 : \theta \in \Theta$, and deduce the corresponding test statistics.

10. For σ^2 a fixed positive constant, consider $X_1, \dots, X_n|\theta \sim^{i.i.d} \mathcal{N}(\theta,\sigma^2)$ with prior distribution $\theta \sim \mathcal{N}(\mu, v^2), \mu \in \mathbb{R}, v^2 > 0$. Show that the posterior distribution of θ given the observations is

$$\theta|X_1,\dots,X_n \sim N\left(\frac{\frac{n\bar X}{\sigma^2} + \frac{\mu}{v^2}}{\frac{n}{\sigma^2} + \frac{1}{v^2}}, \frac{1}{\frac{n}{\sigma^2} + \frac{1}{v^2}}\right), \quad \text{where } \bar X = \frac{1}{n}\sum_{i=1}^n X_i.$$

11. Consider $X_1, \dots, X_n|\mu,\sigma^2$ i.i.d. $\mathcal{N}(\mu,\sigma^2)$ with *improper prior* density $\pi(\mu,\sigma)$ proportional to σ^{-2} (constant in μ). Argue that the resulting 'posterior distribution' has a density proportional to

$$\sigma^{-(n+2)} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n (X_i-\mu)^2\right\},$$

and that the distribution of $\mu|\sigma^2, X_1, \ldots, X_n$ is $\mathcal{N}(\bar{X}, \sigma^2/n)$, where $\bar{X} = (1/n)\sum_{i=1}^n X_i$. For $0 < \alpha < 1$ and assuming σ^2 is known, construct a level $1-\alpha$ credible set for the posterior distribution $\mu|\sigma^2, X_1, \ldots, X_n$ that is also an exact level $1-\alpha$ (frequentist) confidence set.

12. Consider the maximum likelihood estimator $\hat{\theta}_n$ of a sample of size n from a $\mathcal{N}(\theta, 1)$ model, $\theta \in \mathbb{R}$. Define the ('Hodges'-) estimator

$$\tilde{\theta}_n = \hat{\theta}_n 1_{|\hat{\theta}_n| \geq n^{-1/4}}.$$

Show that, under $P_\theta, \theta \neq 0$, one has $\sqrt{n}(\tilde{\theta}_n - \theta) \to^d \mathcal{N}(0,1)$, but that $\tilde{\theta}$ is *superefficient* at $\theta = 0$, that is, under $P_\theta, \theta = 0$, one has $\sqrt{n}(\tilde{\theta}_n - 0) \to^d \mathcal{N}(0,0)$, improving upon the maximum likelihood estimator.