Principles of Statistics **Part II - Michaelmas 2017**

Example Sheet 1

Lecturer: Quentin Berthet

Throughout, the abbreviations 'i.i.d.', 'p.d.f./p.m.f.' and 'MLE' stand for 'independent and identically distributed', 'probability density/mass function' and 'maximum likelihood estimator', respectively. A normal distribution in $\mathbb{R}^d$ with mean vector μ and covariance matrix Σ is denoted by $\mathcal{N}(\mu, \Sigma)$.

1. Consider an i.i.d. sample $X_1, \ldots, X_n$ of random variables. For each of the following parametric models of p.m.f/p.d.f.'s, find the MLE of the unknown parameter, the score equation and the Fisher information.

a) $X_i \sim^{i.i.d.} Bernoulli(\theta), \theta \in [0, 1]$,

b) $X_i \sim^{i.i.d.} \mathcal{N}(\theta, 1), \theta \in \mathbb{R}$,

c) $X_i \sim^{i.i.d.} \mathcal{N}(0, \theta), \theta \in (0, \infty)$,

d) $X_i \sim^{i.i.d.} \mathcal{N}(\mu, \sigma^2), \theta = (\mu, \sigma^2)^T \in \mathbb{R} \times (0, \infty)$,

e) $X_i \sim^{i.i.d.} Poisson(\theta)$, $\theta \in (0, \infty)$,

f) $X_i \sim^{i.i.d.}$ from model $\{f(\cdot, \theta) : \theta \in (0, \infty)\}$ with pdf $f(x, \theta) = (1/\theta)e^{-x/\theta}$, $x \geq 0$.

g) $X_i \sim^{i.i.d.}$ from model $\{f(\cdot, \theta) : \theta \in (0, \infty)\}$ with pdf $f(x, \theta) = \theta e^{-\theta x}$, $x \geq 0$.

2. a) Show that the proof of the Cramèr-Rao lower bound in the lectures for $n = 1$ extends to general $n \geq 1$.

b) In which of the examples of the previous exercise is the MLE unbiased (i.e., does one have $\mathbf{E}_\theta[\hat{\theta}] = \theta$ for all $\theta \in \Theta$)? When unbiased, deduce whether the variance of the MLE attains the Cramèr-Rao lower bound or not.

3. Let $X_1, \ldots, X_n$ be i.i.d. Poisson random variables with parameter $\theta > 0$, and let $\bar{X}_n = (1/n) \sum_{i=1}^n X_i, S_n^2 = (n-1)^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$. Show that $Var(\bar{X}_n) \leq Var(S_n^2)$.

4. Find the MLE for an i.i.d. sample $X_1, \ldots, X_n$ arising from the models a) $\mathcal{N}(\theta, 1)$ where $\theta \in \Theta = [0, \infty)$ and b) $\mathcal{N}(\theta, \theta)$ where $\theta \in \Theta = (0, \infty)$.

5. Consider an i.i.d. sample $X_1, \dots, X_n$ arising from the model

$$\{f(\cdot, \theta) : \theta \in \mathbb{R}\}, \quad f(x, \theta) = \frac{1}{2} e^{-|x-\theta|}, x \in \mathbb{R},$$

of *Laplace distributions*. Assuming n to be odd for simplicity, show that the MLE is equal to the sample median. Discuss what happens when n is even. Can you calculate the Fisher information?

6. Consider observing an $n \times 1$ random vector $Y \sim \mathcal{N}(X\theta, I)$ where X is a non-stochastic $n \times p$ matrix of full column rank, where $\theta \in \Theta = \mathbb{R}^p$ for $p \leq n$, and where I is the $n \times n$ identity matrix. Compute the MLE and find its distribution. Calculate the Fisher information for this model and compare it to the variance of the MLE. Deduce, as a special case, the form of the MLE and Fisher information in the case when $p = n$ and $X = I$.

7. Let $(X, X_n : n \in \mathbb{N})$ be random vectors in $\mathbb{R}^k$.

a) Prove that $X_n \to^P X$ as $n \to \infty$ if and only if each vector component $X_{n,j}$, for $j = 1, \dots, k$, of X_n converges in probability to the corresponding vector component X_j of X as $n \to \infty$. Formulate and prove an analogous result for random symmetric $k \times k$-matrices.

b) Suppose $\mathbf{E}\|X_n - X\| \to 0$ as $n \to \infty$ where $\| \cdot \|$ is the Euclidean norm on $\mathbb{R}^k$. Deduce that $X_n \to^P X$ as $n \to \infty$.

c) Show that the converse in b) is false, that is, give an example of real random variables $X_n \to^P X$ as $n \to \infty$ but $\mathbf{E}|X_n - X| \not\to 0$.

8. Given $X_1, \dots, X_n$ i.i.d. random variables such that $\mathbf{E}[X_i] = 0, \mathbf{E}[X_i^2] \in (0, \infty)$, the *Student t-statistic* is given by

$$t_n = \frac{\sqrt{n}\bar{X}_n}{S_n}, \quad \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i, \; S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

Show that $t_n \to^d \mathcal{N}(0, 1)$ as $n \to \infty$. Assuming now $\mathbf{E}[X_i] = \mu \in \mathbb{R}$, deduce an asymptotic level $1 - \alpha$ confidence interval for $\mathbf{E}[X_i]$.

9. For the examples from Exercise 1, derive directly (without using the general asymptotic theory for MLEs) the asymptotic distribution of $\sqrt{n}(\hat{\theta}_{MLE} - \theta)$ as $n \to \infty$.

10. Suppose one observes *one* random vector $X = (X_1, X_2)^T$ from a bivariate normal distribution $N_2(\theta, \Sigma)$ where $\theta = (\theta_1, \theta_2)^T$ and where Σ is an arbitrary but *known* 2×2 positive definite covariance matrix.

i) Compute the Cramèr-Rao lower bound for estimating the first coefficient θ_1 if a) θ_2 is known and b) if θ_2 is unknown.

ii) Show that the two bounds in i) coincide when Σ is a diagonal matrix.

iii) Show that the bound in i)a) is always less than or equal to the bound in i)b), and give an information-theoretic interpretation of this result.