PRINCIPLES OF STATISTICS – EXAMPLES 1/4

Part II, Michaelmas 2021, Po-Ling Loh (email: pll28@cam.ac.uk)
Questions by courtesy of Richard Nickl

Throughout, the abbreviations "i.i.d.," "pdf/pmf," 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 $N_d(\mu, \Sigma)$, and $N(\mu, \sigma^2)$ corresponds to the univariate case $d = 1$.

1. Consider an i.i.d. sample $X_1, \ldots, X_n$ of random variables. For each of the following parametric models of pmf/pdfs, 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) $X_i \sim^{i.i.d.} N(\theta, 1), \theta \in \mathbb{R}$,
c) $X_i \sim^{i.i.d.} N(0, \theta), \theta \in (0, \infty)$,
d) $X_i \sim^{i.i.d.} N(\mu, \sigma^2), \theta = (\mu, \sigma^2)^T \in \mathbb{R} \times (0, \infty)$,
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. In which of the examples in the previous exercise is the MLE unbiased (i.e., does one have $E_\theta \hat{\theta} = \theta$ for all $\theta \in \Theta$)? When unbiased, deduce whether or not the variance of the MLE attains the Cramér-Rao lower bound.

3. Let $X_1, \ldots, X_n$ be i.i.d. Poisson random variables with parameter $\theta > 0$, and let $\bar{X}_n = (1/n)\sum_{i=1}^n X_i$ and $S_n^2 = (n-1)^{-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2$. 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) $N(\theta, 1)$, where $\theta \in \Theta = [0, \infty)$, and (b) $N(\theta, \theta)$, where $\theta \in \Theta = (0, \infty)$.

5. Consider an i.i.d. sample $X_1, \ldots, X_n$ 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 $Y \sim N_n(X\theta, I)$, where X is a non-stochastic $n \times p$ matrix of full column rank, $\theta \in \Theta = \mathbb{R}^p$ for $p \leq n$, and 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 P_θ be the law of a $N_p(\theta, I)$ distribution on $\mathbb{R}^p$ with identity covariance matrix I and mean vector θ. For fixed $\theta_0 \in \mathbb{R}^p$, compute the Kullback-Leibler divergence $KL(P_{\theta_0}, P_\theta)$.

8. 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 $X_{n,j}, j = 1, \ldots, k$, 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 $k \times k$ matrices.

(b) Suppose $E\|X_n - X\| \to 0$ 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, i.e., give an example of real random variables such that $X_n \to^P X$ as $n \to \infty$, but $E|X_n - X| \not\to 0$.

9. Given $X_1, \dots, X_n$ i.i.d. random variables such that $EX_1 = 0$ and $EX_1^2 \in (0, \infty)$, the *Student's t-statistic* is given by

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

Show that $t_n \to^d N(0,1)$ as $n \to \infty$. Assuming now that $EX_1 = \mu \in \mathbb{R}$, deduce an asymptotic level $1 - \alpha$ confidence interval for EX_1.

10. 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 $n \to \infty$.

11. 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 Σ is an arbitrary but *known* 2×2 positive definite covariance matrix.

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

(b) Show that the two bounds in (a) coincide when Σ is a diagonal matrix.

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