PRINCIPLES OF STATISTICS – EXAMPLES 2/4

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

Throughout, the term "asymptotic" refers to a large-sample limit as $n \to \infty$ under a sampling distribution $P_\theta = P_\theta^{\mathbb{N}}$ where, unless otherwise specified, θ is assumed to be a fixed element of the parameter space $\Theta \subseteq \mathbb{R}^p$.

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 that for some θ_0 in the interior of Θ and every $\varepsilon > 0$ small enough, we have $S(\theta_0 \pm \varepsilon) < 0 < S(\theta_0 \mp \varepsilon)$, and also 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. Consider an i.i.d. sample $X_1, \dots, X_n$ 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 with probability one and is consistent. [Hint: Use the previous exercise. You may interchange differentiation $d/d\theta$ and dx-integration without justification in your argument.]

3. Give an example of functions $\{Q_n\}$ and Q defined on $\Theta \subseteq \mathbb{R}$ which have unique maximizers $\{\hat{\theta}_n\}$ and θ_0, respectively, such that $Q_n(\theta) \to Q(\theta)$ for every $\theta \in \Theta$ as $n \to \infty$, but $\hat{\theta}_n \not\to \theta_0$.

4. Consider the maximum likelihood estimator $\hat{\theta}$ from $X_1, \dots, X_n$ i.i.d. $N(\theta, 1)$ 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 asymptotic theory for maximum likelihood estimators.

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

6. Suppose one is given a 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 and F are subsets of $\mathbb{R}$.

(a) Assuming Φ is injective, show that a 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 injective. 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 $\Phi(\hat{\theta}_{MLE})$ maximizes $L^*(\phi)$).

(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(θ) model, and (iii) $\phi = (EX)^2$ in a $N(\theta, 1)$ model.

7. Consider the parameter $\phi = EX^4$ equal to the fourth moment of a $N(0, \theta)$ distribution. Find the MLE $\hat{\phi}$ of ϕ and derive the asymptotic distribution of $\sqrt{n}(\hat{\phi} - \phi)$ as $n \to \infty$.

8. Let $\hat{\theta}$ be the maximum likelihood estimator in a model $\{f(\cdot, \theta) : \Theta \subset \mathbb{R}^p\}$ arising from an i.i.d. sample $X_1, \ldots, X_n$. Assuming the model satisfies the regularity conditions from lectures, ensuring in particular the asymptotic normality of $\sqrt{n}(\hat{\theta} - \theta)$ under P_θ, derive the asymptotic distribution of the random variable

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

under P_θ, where i_n equals either $i_n(\theta)$ or $i_n(\hat{\theta})$, and $i_n(\theta)$ denotes the observed Fisher information matrix at θ. From this limiting result: (i) derive a test for the hypothesis $H_0 : \theta = \theta_0$ vs. $H_1 = \Theta \backslash \{\theta_0\}$ which has asymptotic type I error at most α; and (ii) deduce that the confidence ellipsoid

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

has asymptotic coverage level $1 - \alpha$, where z_α denotes the $(1-\alpha)$-quantile of the limit distribution derived above.

9. Consider the parametric models from Exercise 1 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 test of $H_0 : \theta = \theta_0$, $\theta_0 \in \Theta$, vs. $H_1 = \Theta \setminus \{\theta_0\}$.

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

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

11. Consider $X_1, \ldots, X_n | \mu, \sigma^2$ i.i.d. $N(\mu, \sigma^2)$ 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 thus, the distribution of $\mu | \sigma^2, X_1, \ldots, X_n$ is $N(\bar{X}, \sigma^2/n)$, where $\bar{X} = (1/n) \sum_{i=1}^n X_i$. When $0 < \alpha < 1$ 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.