PRINCIPLES OF STATISTICS Part II
Example Sheet 2 (of 4) Michaelmas 2025
By courtesy of Prof Rajen Shah

1. Suppose $X_1, \dots, X_n$ are independent copies of a random variable X taking values in some interval $[a, b]$. Let $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$ be the order statistics of the data (assume for simplicity that there are no ties almost surely). The sample median $\hat{\theta}$ is given by

$$\hat{\theta} = \begin{cases} X_{((n+1)/2)} & \text{if } n \text{ is odd,} \\ \frac{1}{2}(X_{(n/2)} + X_{(n/2+1)}) & \text{if } n \text{ is even.} \end{cases}$$

   (a) Show that $\hat{\theta}$ minimises
   $$M_n(\theta) := \frac{1}{n}\sum_{i=1}^{n} |X_i - \theta|$$
   over $\theta \in [a, b]$.

   (b) Suppose θ_0 is such that $\mathbb{P}(X \le \theta_0) = 1/2$ and assume additionally that for all $\epsilon > 0$, we have
   $$\mathbb{P}(X \le \theta_0 - \epsilon) < 1/2 < \mathbb{P}(X \le \theta_0 + \epsilon).$$
   Show that θ_0 is the unique minimiser over θ of $\theta \mapsto \mathbb{E}|X - \theta|$. [*Hint: First aim to show that when* $\theta > \theta_0$, $\mathbb{E}|X - \theta| - \mathbb{E}|X - \theta_0| = 2\mathbb{E}\{\mathbb{1}_{\{\theta_0 < X \le \theta\}}(\theta - X)\}$.]

   (c) Show that $\hat{\theta} \xrightarrow{p} \theta_0$. [*Hint: Apply a (version of) a result from lectures.*]

2. Let $\Theta \subseteq \mathbb{R}$ have non-empty interior, and let $(\psi_n)_{n\in\mathbb{N}}$ be a sequence of real-valued continuous functions defined on Θ such that $\psi_n(\theta) \xrightarrow{p} \psi(\theta)$ for all $\theta \in \Theta$, where ψ is some deterministic function $\psi : \Theta \to \mathbb{R}$. Suppose that for some $\theta_0 \in \operatorname{int}\Theta$ and every $\delta > 0$ sufficiently small, we have $\psi(\theta_0 \pm \delta) < 0 < \psi(\theta_0 \mp \delta)$. Suppose moreover that ψ_n has exactly one zero $\hat{\theta}_n \in \Theta$ for every $n \in \mathbb{N}$. Deduce that $\hat{\theta}_n \xrightarrow{p} \theta_0$ as $n \to \infty$.

3. Let $X_1, \dots, X_n$ be i.i.d. observations from the model
$$\{f(x, \theta) = \theta x^{\theta-1}\exp(-x^{\theta})\mathbb{1}_{\{x>0\}} : \theta \in (0, \infty)\}$$
of *Weibull* distributions. Show that the MLE exists almost surely and is consistent. [*Hint: Use the previous question. You may interchange differentiation and integration without proof in your answer.*]

4. Let $X_1, \dots, X_n \overset{\text{i.i.d.}}{\sim} N(\theta, 1)$ where $\theta \in \Theta = [0, \infty)$. Find the MLE $\hat{\theta}$. Show that when the true parameter $\theta_0 > 0$ then $\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, 1)$. What happens when $\theta_0 = 0$? Comment on your findings in the light of the general asymptotic theory of MLEs.

5. Let $X_1, \dots, X_n \overset{\text{i.i.d.}}{\sim} U[0, \theta]$ with $\theta \in \Theta = (0, \infty)$.

   (a) Find the MLE $\hat{\theta}$ and show that $\tilde{\theta} = (n+1)\hat{\theta}/n$ is unbiased for estimating θ.

   (b) Find the variance of $\tilde{\theta}$ and comment on this in relation to the form of the Cramér–Rao lower bound in regular models.

   (c) Finally, find the asymptotic distribution of $n(\hat{\theta} - \theta)$.

6. Consider the setup of Question 5 from Example Sheet 1 where $(Y_1, X_1), \dots (Y_n, X_n) \overset{\text{i.i.d.}}{\sim} N_2(\mu, \Sigma)$ where Σ is known and $\mu \in \mathbb{R}^2$. Show that the Wald test, score test and generalised likelihood ratio test for testing the composite null hypothesis $H_0 : \mu_2 = \mu_0$ are identical and deliver exact (non-asymptotic) size control.

7. Suppose $X_1, \dots, X_n \overset{\text{i.i.d.}}{\sim} \text{Bern}(\theta)$ where $\theta \in (0, 1)$.

   (a) Write down a $(1-\alpha)$-level Wald confidence interval for θ (estimating $I_1(\theta_0)$, where θ_0 is the true parameter, by $I_1(\bar{X})$). What happens when $\min(\bar{X}, (1-\bar{X})) < z_\alpha^2(n+z_\alpha^2)^{-1}$, where z_α is the upper $\alpha/2$ point of a standard Gaussian?

   (b) Now let $\phi(u) := \log(u/(1-u))$ for $u \in (0, 1)$. By considering the asymptotic distribution of $\sqrt{n}(\phi(\bar{X}) - \phi(\theta))$, find a different confidence interval for θ. Briefly comment on why this second approach might be preferred.

8. We revisit Question 3 from Example Sheet 1. Suppose we have n i.i.d. copies $X_1, \dots, X_n$ of $X^{(1)} \sim N(\mu_1, 1)$ and also n i.i.d. observations $W_1, \dots, W_n$, independent of $X_1, \dots, X_n$, each having distribution equal to that of $\text{sgn}(X^{(2)})$ where $X^{(2)} \sim N(\mu_2, 1)$. Without using the general asymptotic theory concerning MLEs, find an asymptotic $(1 - \alpha)$-level confidence interval for $\mu_0 := \mu_1 - \mu_2$.

9. Consider the model $X_1, \dots, X_n \overset{\text{i.i.d.}}{\sim} \text{Exp}(\theta)$ where $\theta > 0$ (so the density is $\theta e^{-\theta x}$). We wish to test the null hypothesis $\theta = \theta_0$. Recall that the power of a test $\phi_n = \phi_n(X_1, \dots, X_n)$ where $\phi_n = 1$ indicates rejection (and 0 acceptance) at the alternative $\theta_1 \neq \theta_0$ is $\mathbb{P}_{\theta_1}(\phi_n = 1)$. Consider a sequence $\theta_n = \theta_0 + h/\sqrt{n}$ (where $h \neq 0$ is some fixed real number) of *local alternatives*. Show that the limit

$$\lim_{n\to\infty} \mathbb{P}_{\theta_n}(\phi_n = 1)$$

   is the same for the score and Wald tests with asymptotic size α and find this limit. [*Hint: First argue that if* $X \sim \text{Exp}(\theta)$ *then* $a^{-1}X \sim \text{Exp}(a\theta)$ *for* $a > 0$.]

10. Let $\sigma^2 > 0$ be deterministic and suppose $X_1, \dots, X_n \,|\, \theta \overset{\text{i.i.d.}}{\sim} N(\theta, \sigma^2)$ 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 is given by

$$\theta \,|\, X_1, \dots, X_n \sim N\left(\frac{\frac{n\bar{X}}{\sigma^2} + \frac{\mu}{v^2}}{\frac{n}{\sigma^2} + \frac{1}{v^2}}, \frac{1}{\frac{n}{\sigma^2} + \frac{1}{v^2}}\right)$$

11. Consider $X_1, \dots, X_n \,|\, \mu, \sigma^2 \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$ 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 thus the distribution of $\mu \,|\, \sigma^2, X_1, \dots, X_n$ is $N(\bar{X}, \sigma^2/n)$. Assuming σ^2 is known, for $\alpha \in (0, 1)$ construct a $(1-\alpha)$-level credible set for the posterior distribution $\mu \,|\, \sigma^2, X_1, \dots, X_n$ that is also an exact $(1 - \alpha)$-level (frequentist) confidence set.