Advanced Mathematics and Statistics
Module 2 - Advanced Statistical Methods

Mock general exam

Exercise 1. Let (X, Y) have joint probability density function

$$f_{X,Y}(x, y) = \begin{cases} 10(x^2 - y) & 0 \le x \le 1 \quad \text{and} \quad 0 \le y \le x^2 \\ 0 & \text{otherwise} \end{cases}$$

(a) Determine the conditional probability density function of X given $Y = y$, for any $y \in (0, 1)$.

(b) Evaluate $\mathbb{P}[Y > X^2/3]$

(c) Determine the probability density function of $W = Y/X^2$

Exercise 2. Let $X_1, \ldots, X_n \stackrel{\text{iid}}{\sim} \text{Pareto}(\frac{1}{\theta}, 2)$, with $\theta > 0$, namely

$$f(x|\theta) = \frac{2}{\theta^2 x^3} \mathbb{1}_{(\frac{1}{\theta}, +\infty)}(x)$$

(a) Determine the value of $c \in \mathbb{R}$ that makes $\hat{\theta} = cX_{(1)}$ an unbiased estimator of $1/\theta$, where $X_{(1)} = \min\{X_1, \ldots, X_n\}$.

(b) Identify a sufficient and complete statistic for θ.

(c) Propose a MVUE of $1/\theta$.

Exercise 3. Let $X_1, \ldots, X_n \stackrel{\text{iid}}{\sim} \text{Unif}(0, \theta)$, namely each X_i has density function

$$f(x; \theta) = \frac{1}{\theta} \mathbb{1}_{(0,\theta)}(x).$$

For testing

$$H_0 : \theta = 1 \qquad \text{vs} \qquad H_1 : \theta = 2$$

the critical region CR $\{X_1 > 0.95\}$ has been proposed

(a) Determine the probabilities of type I and type II error associated to this test.

(b) Identify the Neyman–Pearson test of size $\alpha = 0.05$ for this problem.

(c) Show that the Neyman–Pearson test is more powerful than the proposed test with critical region CR.

Exercise 4. Let $X_1, \ldots, X_n | \sigma \overset{\text{iid}}{\sim} f(\cdot|\sigma)$, where

$$f(x|\sigma) = \frac{1}{x}\sqrt{\frac{\sigma}{2\pi}} \exp\left\{-\frac{\sigma}{2}(\log(x))^2\right\}, \quad x > 0$$

and σ is a positive quantity, whose prior distribution over $\mathbb{R}^+$ is a gamma with shape–rate parameters $(1, 2)$, i.e. $p(\sigma) = 2e^{-2\sigma}$ for $\sigma > 0$.

(a) Determine the posterior distribution of σ, given $X_1 = x_1, \ldots, X_n = x_n$.

(b) Determine the Bayes estimator $\hat{\sigma}_p$ of σ under a squared loss function.

(c) Assuming σ fixed, i.e., $X_1, \ldots, X_n \overset{\text{iid}}{\sim} f(\cdot|\sigma)$, determine the MLE $\hat{\sigma}$ of σ and show that $\hat{\sigma}_p/\hat{\sigma} \to 1$ in probability as $n \to +\infty$.

Solutions

Exercise 1.

(a) Since the marginal density of Y is

$$f_Y(y) = \mathbb{1}_{(0,1)}(y) \int_{\sqrt{y}}^{1} 10(x^2 - y)\,\mathrm{d}x = \Big\{\frac{10}{3} - \frac{10}{3}y^{3/2} - 10y + 10y^{3/2}\Big\}\,\mathbb{1}_{(0,1)}(y)$$
$$= \Big\{\frac{10}{3} + \frac{20}{3}y^{3/2} - 10y\Big\}\,\mathbb{1}_{(0,1)}(y)$$

for any $y \in (0,1)$, the conditional density of X, given $Y = y$, turns out to be

$$f_{X|Y}(x|y) = \frac{30(x^2 - y)}{10 + 20y^{3/2} - 30y}\,\mathbb{1}_{(\sqrt{y},\,1)}(x)$$

(b)

$$\mathbb{P}[Y > X^2/3] = \int_0^1 \int_{x^2/3}^{x^2} 10(x^2 - y)\,\mathrm{d}y\,\mathrm{d}x = 10\int_0^1 \Big(1 - \frac{1}{3} - \frac{1}{2} + \frac{1}{18}\Big)\,x^4\,\mathrm{d}x = \frac{4}{9}$$

(c) Setting

$$W = \frac{Y}{X^2} \qquad V = Y$$

we have

$$X = \sqrt{\frac{V}{W}}, \quad Y = V$$

and the Jacobian is

$$\Big| -\frac{\sqrt{v}}{2w^{3/2}} \Big|$$

and the density of (W, V) is

$$f_{W,V}(w,v) = 10\Big(\frac{v}{w} - v\Big)\,\frac{\sqrt{v}}{2w^{3/2}}\,\mathbb{1}_{(0,1)}(w)\,\mathbb{1}_{(0,w)}(v) = \frac{5v^{3/2}(1-w)}{w^{5/2}}\,\mathbb{1}_{(0,1)}(w)\,\mathbb{1}_{(0,w)}(v).$$

Indeed

$$0 \le x \le 1 \Rightarrow 0 \le v \le w, \quad \text{and} \quad 0 \le y \le x^2 \Rightarrow 0 \le w \le 1.$$

It, then, follows that

$$f_W(w) = \frac{5(1-w)}{w^{5/2}}\,\mathbb{1}_{(0,1)}(w) \int_0^w v^{3/2}\,\mathrm{d}v = 2(1-w)\,\mathbb{1}_{(0,1)}(w)$$

so that $W \sim \mathrm{Beta}(1,2)$.

Exercise 2.

(a) Since

$$F(x) = \mathbb{1}_{(1/\theta,+\infty)}(x) \int_{1/\theta}^{x} \frac{2}{\theta^2 s^3}\,\mathrm{d}s = \left\{1 - \frac{1}{\theta^2\,x^2}\right\} \mathbb{1}_{(1/\theta,+\infty)}(x)$$

the density function of the first order statistic $X_{(1)}$ is

$$f_{X_{(1)}}(x) = nf(x)\,(1-F(x))^{n-1} = n\frac{2}{\theta^2 x^3}\,\mathbb{1}_{(1/\theta,+\infty)}(x)\left(\frac{1}{\theta^2\,x^2}\right)^{n-1} = \frac{2n}{\theta^{2n}\,x^{2n+1}}\,\mathbb{1}_{(1/\theta,+\infty)}(x)$$

namely $X_{(1)} \sim \text{Pareto}(1/\theta, 2n)$. Hence

$$\mathbb{E}\hat{\theta} = c\,\mathbb{E}X_{(1)} = c\int_{1/\theta}^{+\infty} x\,\frac{2n}{\theta^{2n}x^{2n+1}}\,\mathrm{d}x = c\,\frac{2n}{\theta^{2n}}\int_{1/\theta}^{+\infty}\frac{1}{x^{2n}}\,\mathrm{d}x = c\,\frac{2n}{2n-1}\,\frac{1}{\theta}.$$

Hence, if we set $c = (2n-1)/(2n)$, one has

$$\mathbb{E}\,\hat{\theta} = \frac{2n-1}{2n}\,\mathbb{E}X_{(1)} = \frac{1}{\theta}.$$

(b) In order to show sufficiency, one may resort to the factorization theorem. Since

$$f(x_1,\ldots,x_n|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n}\frac{2}{\theta^2\,x_i^3}\,\mathbb{1}_{(1/\theta,+\infty)}(x_i) = \frac{2^n}{\theta^{2n}\prod_{i=1}^{n}x_i^3}\prod_{i=1}^{n}\mathbb{1}_{(1/\theta,+\infty)}(x_i).$$

Since

$$\prod_{i=1}^{n}\mathbb{1}_{(1/\theta,+\infty)}(x_i) = \mathbb{1}_{(1/\theta,+\infty)}(x_{(1)})$$

one has

$$f(x_1,\ldots,x_n|\theta) = \frac{2^n}{\theta^{2n}}\,\mathbb{1}_{(1/\theta,+\infty)}(x_{(1)})\,\frac{1}{\prod_{i=1}^{n}x_i^3} = \nu(t;\theta)\,w(x_1,\ldots,x_n)$$

where $t = x_{(1)}$,

$$\nu(t;\theta) = \frac{2^n}{\theta^{2n}}\,\mathbb{1}_{(1/\theta,+\infty)}(t) \quad \text{and} \quad w(x_1,\ldots,x_n) = \frac{1}{\prod_{i=1}^{n}x_i^3}$$

Hence, $T = X_{(1)}$ is sufficient for estimating θ.

As for completeness, if $g : \mathbb{R} \to \mathbb{R}$ is any function such that $\mathbb{E}g(X_{(1)}) = 0$, we have to show that $\mathbb{P}[g(X_{(1)}) = 0] = 1$. Because of (a)

$$0 = \mathbb{E}g(X_{(1)}) = \int_{1/\theta}^{+\infty} g(x)\frac{2n}{\theta^{2n}x^{2n+1}}\,\mathrm{d}x = \frac{2n}{\theta^{2n}}\int_{1/\theta}^{+\infty}\frac{g(x)}{x^{2n+1}}\,\mathrm{d}x$$

If we decompose g into its positive $g^+ = \max\{0, g\}$ and negative $g^- = -\min\{0, g\}$, one has that $\mathbb{E}g(X_{(1)}) = 0$ if and only if

$$\int_{1/\theta}^{+\infty} \frac{g^+(x)}{x^{2n+1}}\,\mathrm{d}x = \int_{1/\theta}^{+\infty} \frac{g^-(x)}{x^{2n+1}}\,\mathrm{d}x$$

Each of the two integrals on the right and left hand side are non–increasing functions of θ since $g^+ \geq 0$ and $g^- \geq 0$. Hence, the derivatives must coincide, which entails

$$-\frac{1}{\theta^2}\,\frac{g^+\left(\frac{1}{\theta}\right)}{(1/\theta)^{2n+1}} = -\frac{1}{\theta^2}\,\frac{g^-\left(\frac{1}{\theta}\right)}{(1/\theta)^{2n+1}}$$

and this is equivalent to $g^+(1/\theta) = g^-(1/\theta)$ for every θ. Hence, *a fortiori* $\mathbb{P}[g(X_{(1)}) = 0] = 1$ and $X_{(1)}$ is also complete.

(c) Since

$$\hat{\theta} = \frac{2n-1}{2n}\,X_{(1)}$$

is an unbiased estimator that is also a function of the complete and sufficient statistic $X_{(1)}$, by the Lehmann-Scheffè theorem is is also a MVUE of $1/\theta$.

Exercise 3.

(a) Since under H_0 one has $X \sim \text{Unif}(0,1)$, the probability of type I error is

$$\alpha = \mathbb{P}_{\theta=1}(X_1 > 0.95) = 0.05$$

so that the proposed test has size $\alpha = 0.05$. As for the probability of type II error, under H_1 one has $X_1 \sim \text{Unif}(0,2)$ and

$$\beta = \mathbb{P}_{\theta=2}(X_1 \le 0.95) = \frac{0.95}{2} = 0.475.$$

Note that this implies that the power function at $\theta = 2$ is $p(2) = 1 - \beta = 0.525$.

(b) The Neyman–Pearson lemma suggests to use a test with critical region such that the decision rule is

$$\text{reject } H_0 \text{ whenever } \frac{L(x_1,\dots,x_n;2)}{L(x_1,\dots,x_n;1)} > k$$

where k is such that

$$\mathbb{P}_{\theta=1}\Big(\frac{L(X_1,\dots,X_n;2)}{L(X_1,\dots,X_n;1)} > k\Big) = \alpha = 0.05.$$

In this case, it can be seen that

$$L(x_1,\dots,x_n;\theta) = \frac{1}{\theta^n}\,\mathbb{1}_{(0,+\infty)}(x_{(1)})\,\mathbb{1}_{(-\infty,\theta)}(x_{(n)}),$$

which implies that

$$\frac{L(x_1,\dots,x_n;2)}{L(x_1,\dots,x_n;1)} = \begin{cases} 2^{-n} & \text{if } 0 \le x_{(n)} < 1 \\ +\infty & \text{if } 1 \le x_{(n)} < 2 \end{cases}$$

This clearly shows that the likelihood ratio is monotone increasing in $x_{(n)}$ and the Neyman–Pearson test of size α boils down to

$$\text{reject } H_0 \text{ whenever } x_{(n)} > k'$$

where $k' > 0$ is such that

$$\mathbb{P}_{\theta=1}\Big[X_{(n)} > k'\Big] = \alpha = 0.05$$

If $\theta = 1$, then the density of $X_{(n)}$ is

$$f_{X_{(n)}}(x;\theta) = n\,x^{n-1}\,\mathbb{1}_{(0,1)}(x).$$

In view of this

$$0.05 = \int_{k'}^{1} nx^{n-1}\,\mathrm{d}x = 1 - (k')^n$$

from which we get $k' = (0.95)^{1/n}$. Hence, the Neyman–Pearson test is

$$\text{reject } H_0 \text{ whenever } x_{(n)} > (0.95)^{1/n}$$

(c) The power function of the Neyman and Pearson test above, evaluated at $\theta = 2$, is

$$p^*(2) = \mathbb{P}_{\theta=2}\Big[X_{(n)} > (0.95)^{1/n}\Big]$$

If $\theta = 2$, the density function of the largest order statistic $X_{(n)}$ is

$$f_{X^{(n)}}(x) = n\,\frac{x^{n-1}}{2^{n-1}}\,\frac{1}{2}\,\mathbb{1}_{(0,2)}(x)$$

and the power function of the Neyman–Pearson test is

$$p^*(2) = \frac{1}{2^n}\int_{(0.95)^{1/n}}^{2} n\,x^{n-1}\,\mathrm{d}x = 1 - \frac{0.95}{2^n}$$

If $n = 1$, then $p^*(2) = 0.525 = p(2)$ and the power is the same as before. But, if $n \geq 2$, then

$$p^*(2) = 1 - \frac{0.95}{2^n} > 0.525 = p(2)$$

and, as expected, the Neyman–Pearson test is more powerful.

Exercise 4.

(a) Recall that the likelihood function of the data is

$$f(x_1, \dots, x_n|\sigma) = \frac{1}{\prod_{i=1}^n x_i} (\sigma/(2\pi))^{n/2} \exp\left\{-\frac{\sigma}{2}\sum_{i=1}^n \log(x_i)^2\right\}.$$

We may apply the Bayes theorem to determine the posterior:

$$\begin{aligned} p(\sigma|x_1, \dots, x_n) &\propto f(x_1, \dots, x_n|\sigma) \cdot p(\sigma) \\ &\propto \sigma^{n/2} \exp\left\{-\sigma\left(2^{-1}\sum_{i=1}^n \log(x_i)^2 + 2\right)\right\} \end{aligned}$$

therefore

$$\tilde{\sigma}|X_1 = x_1, \dots, X_n = x_n \sim \text{Gamma}\left(n/2 + 1, \frac{1}{2}\sum_{i=1}^n \log(x_i)^2 + 2\right).$$

(b) The Bayes estimator of σ under a squared loss function is the posterior mean:

$$\hat{\sigma}_p = \int_0^\infty \sigma p(\sigma|x_1, \dots, x_n) \mathrm{d}\sigma = \frac{n+2}{4 + \sum_{i=1}^n \log(x_i)^2}$$

where we used the fact that the mean of a gamma with parameters (a, b) equals a/b.

(c) Now we have to maximize the likelihood function:

$$L(\sigma) = f(x_1, \dots, x_n|\sigma) = \frac{1}{\prod_{i=1}^n x_i} (\sigma/(2\pi))^{n/2} \exp\left\{-\frac{\sigma}{2}\sum_{i=1}^n \log(x_i)^2\right\}.$$

For simplicity, we consider the log-likelihood function

$$\ell(\sigma) = \log(L(\sigma)) = -\sum_{i=1}^n \log(x_i) + \frac{n}{2}\log(\sigma/(2\pi)) - \frac{\sigma}{2}\sum_{i=1}^n \log(x_i)^2.$$

It is easy to see that

$$\frac{\partial}{\partial\sigma}\ell(\sigma) \geq 0 \quad \text{iff} \quad \sigma \leq \frac{n}{\sum_{i=1}^n \log(x_i)^2}$$

therefore

$$\hat{\sigma} = \frac{n}{\sum_{i=1}^n \log(X_i)^2}$$

is the MLE of σ. In order to prove the convergence in probability, we use the consistency of the MLE, indeed one has $\hat{\sigma} \xrightarrow{\text{P}} \sigma$ as $n \to +\infty$. As a consequence we obtain:

$$\frac{\hat{\sigma}_p}{\hat{\sigma}} = \frac{n+2}{4 + \sum_{i=1}^n \log(X_i)^2} \cdot \frac{1}{\hat{\sigma}} = \frac{n+2}{4 + n/\hat{\sigma}} \cdot \frac{1}{\hat{\sigma}} \xrightarrow{\text{P}} \frac{\sigma}{\sigma} = 1$$

as $n \to +\infty$.