Vrije Universiteit Amsterdam
Department of Mathematics

Resit Mathematical Statistics (XB_0049)
February 13, 2024, 12.15–14.30

You are not allowed to use calculators, phones, laptops or other tools. Clearly put your name and student number on all sheets that you submit. Your answer may contain quantiles from a distribution. Expressions for densities, expectation, variance and quantiles for some distributions are given on page 3 of the exam. You can use these expressions if needed for answering the questions below.

Unless stated differently, always add an explanation to your answer.

1. Let (X_i, Y_i), $i = 1, \ldots, n$ be independent, identically distributed pairs of random variables. Assume that the joint density of (X_i, Y_i) is given by

$$f_\theta(x,y) = \begin{cases} e^{-\theta x - y/\theta} & \text{if} \quad x > 0 \text{ and } y > 0 \\ 0 & \text{else} \end{cases}.$$

   Here, $\theta > 0$ is an unknown parameter.

   (a) [3 PT]. Derive the maximum likelihood estimator (MLE) for θ. Denote it by $\hat{\Theta}_n$.

   (b) [1 PT]. Derive a sufficient statistic for θ of dimension (strictly) less than n.

   (c) [2 PT]. Compute the Fisher-information for (X_1, Y_1).

   (d) [1 PT]. $\sqrt{n}(\hat{\Theta}_n - \theta)$ is asymptotically normal. What are its mean and variance?

2. Let $X_1, \ldots, X_n \mid \Theta = \theta \stackrel{\text{iid}}{\sim} Unif(0, \theta)$. Suppose we specify the prior distribution to Θ to be $Par(\alpha, \beta)$ (the Pareto distribution, see "Distributions and Notation" for a definition). Suppose we observe realisations $x_1, \ldots, x_n$.

   (a) [2 PT]. Show that the posterior distribution of Θ is of Pareto type and identify the parameters (α_p, β_p) of this posterior distribution.

   (b) [2 PT]. Suppose c solves $(\beta_p/c)^{\alpha_p} = 0.05$. Show that the 95% highest probability density set for θ is given by $[\beta_p, c]$. *Hint: first make a sketch of the posterior density.*

   (c) [2 PT]. Fix a number $k > 0$. We wish to test the hypothesis $H_0 \colon \theta \in [0, k]$ against $H_1 \colon \theta > k$. We do so by choosing the more likely (a posteriori) hypothesis. Assume $X_{(n)} > \beta$. Show that we choose H_1 if $X_{(n)} > 2^{-\alpha_p} k$.

3. Suppose $X_1, \dots, X_n$ are independent random variables, each with density

$$f_{X_i}(x;\theta) = \frac{1}{B(\theta,\theta)} x^{\theta-1}(1-x)^{\theta-1}\mathbf{1}_{(0,1)}(x), \qquad \theta > 0.$$

Here $B(\theta,\theta) = \int_0^1 x^{\theta-1}(1-x)^{\theta-1}dx$. We wish to test $H_0\colon \theta = 1$ against $H_1\colon \theta = 2$.

(a) [2 PT]. Show that the most powerful test leads to rejecting H_0 in favour of H_1 for large values of

$$\sum_{i=1}^{n} (\log X_i + \log(1 - X_i))\,.$$

(b) Now consider $n = 1$ and the decision rule to reject H_0 when $X_1 \in [1/2 - c, 1/2 + c]$.

i. [2 PT]. Determine c such that the corresponding test has significance level 0.05.

ii. [2 PT]. Show that the power of the test equals

$$6(G(c + 1/2) - G(c - 1/2)),$$

where $G(x) = x^2/2 - x^3/3$.

4. Suppose $X_1, \dots, X_n$ are independent, each with the $N(0, \sigma^2)$ distribution.

(a) [2 PT]. Show that $\sigma^{-2}\sum_{i=1}^{n} X_i^2$ is a pivot.

(b) [2 PT]. Use this result to derive an expression for a $(1-\alpha)100\%$ confidence interval for σ.

Distributions and notation

- If $X \sim Unif\,(0, \theta)$, then $\mathbb{E}X = \theta/2$ and $\mathrm{Var}(X) = \theta^2/12$.
- If $X \sim N(\mu, \sigma^2)$, then its density is given by $p(x) = (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{(x-\mu)^2)}{2\sigma^2}\right)$ and in case $\mu - 0$ and $\sigma = 1$ we define $\Phi(x) := \mathbb{P}(X \leq x)$. The lower quantile ξ_α for the standard Normal distribution is defined by $\Phi(\xi_\alpha) = \alpha$, for $\alpha \in (0, 1)$. The upper quantile z_α for the standard Normal distribution is defined by $\Phi(z_\alpha) = 1 - \alpha$.
- If $X \sim \chi^2_k$, then $\chi^2_{k,\alpha}$ is defined by $\mathbb{P}(X \leq \chi^2_{k,\alpha}) = \alpha$.
- If $X \sim t_k$, then $t_{k,\alpha}$ is defined by $\mathbb{P}(X \leq t_{k,\alpha}) = \alpha$.
- If $X \sim Ga(\alpha, \beta)$, then its density is given by $p(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \mathbf{1}_{(0,\infty)}(x)$. $\mathbb{E}X = \alpha/\beta$, $\mathrm{Var}(X) = \alpha/\beta^2$. The special case of $\alpha = 1$ corresponds to the exponential distribution with parameter β.
- If $X \sim Bin(n, p)$, then $\mathbb{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$, for $x \in \{0, 1, \ldots, n\}$. $\mathbb{E}X = np$, $\mathrm{Var}(X) = np(1-p)$.
- If $X \sim Beta(\alpha, \beta)$, then its density is given by $p(x) = B(\alpha, \beta)^{-1} x^{\alpha-1} (1-x)^{\beta-1} \mathbf{1}_{[0,1]}(x)$. $\mathbb{E}X = \alpha/(\alpha+\beta)$, $\mathrm{Var}(X) = \alpha\beta/\left((\alpha+\beta)^2(\alpha+\beta+1)\right)$
- If $X \sim Par(\alpha, \beta)$, then its density is given by $p(x) = \alpha\beta^\alpha x^{-(\alpha+1)} \mathbf{1}_{[\beta,\infty)}(x)$. $\mathbb{P}(X \leq x) = 1 - (\beta/x)^\alpha$ for $x \geq \beta$. If $\alpha > 1$, then $\mathbb{E}X = \alpha\beta/(\alpha - 1)$.