Name: **Paolo Zacchia / Yaroslav Korobka** Exercise Session: **2**
Course: **Statistics** Date: November 25, 2024

Problem 1

A random variable X is said to have a Pareto distribution with parameter β, denoted as $X \sim \text{Pareto}(\beta)$, if it is continuously distributed with density

$$f_X(x;\beta) = \begin{cases} \beta x^{-\beta-1}, & \text{if } x > 1, \\ 0, & \text{otherwise.} \end{cases}$$

A random sample $x_1, \ldots, x_N$ from the Pareto(β) population is available. Derive the maximum-likelihood estimator of β. Does it maximize the log-likelihood function?

Solution 1

By definition and property of the likelihood function,

$$\begin{aligned} \hat{\beta}_{ML} &= \arg\max_{\beta} \mathcal{L}_N(\beta|x_1, \ldots, x_N) \\ &= \arg\max_{\beta} \log \mathcal{L}_N(\beta|x_1, \ldots, x_N). \end{aligned}$$

Thus, for our case, log-likelihood function looks like

$$\begin{aligned} \log \mathcal{L}_N(\beta|x_1, \ldots, x_N) &= \sum_{i=1}^{N} \log f_X(x_i;\beta) \\ &= \sum_{i=1}^{N} \log(\beta) - (\beta+1)\log(x_i) \\ &= N\log(\beta) - (\beta+1)\sum_{i=1}^{N}\log(x_i). \end{aligned}$$

To obtain the maximum-likelihood estimator of β, we take the derivative w.r.t. β and set it to zero,

$$\frac{\partial \log \mathcal{L}_N}{\partial \beta} = \frac{N}{\hat{\beta}_{ML}} - \sum_{i=1}^{N}\log(x_i) = 0. \tag{1}$$

From (1) we get

$$\hat{\beta}_{ML} = \frac{N}{\sum_{i=1}^{N}\log(x_i)}.$$

To check that $\hat{\beta}_{ML}$ is indeed the optimum, take the second derivative of the log-likelihood w.r.t. β,

$$\frac{\partial^2 \log \mathcal{L}_N}{\partial \beta^2}\bigg|_{\beta=\hat{\beta}_{ML}} = -\frac{N}{\beta^2} < 0.$$

Hence, our estimator maximizes the log-likelihood function.

Problem 2

Suppose that the random variables $Y_1, \dots, Y_n$ satisfy

$$Y_i = \beta x_i + \epsilon_i, \quad i = 1, \dots, n,$$

where $x_1, \dots, x_n$ are fixed constants, and $\epsilon_1, \dots, \epsilon_n$ are i.i.d. $\mathcal{N}(0, \sigma^2)$ with σ^2 unknown. Find the MLE of β. Is it unbiased?

Solution 2

Fix the value of σ^2. Write the likelihood function as

$$\mathcal{L}_N(\beta | Y_1, \dots, Y_N) = (2\pi\sigma^2)^{-N/2} \exp\left(-\frac{\beta^2}{2\sigma^2} \sum_{i=1}^{N} x_i^2\right) \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{N} Y_i^2 + \frac{\beta}{\sigma^2} \sum_{i=1}^{N} Y_i x_i\right),$$

and taking the logarithm, we obtain the log-likelihood function:

$$\log \mathcal{L}_N(\beta | Y_1, \dots, Y_N) = -\frac{N}{2} \log(2\pi\sigma^2) - \frac{\beta^2}{2\sigma^2} \sum_{i=1}^{N} x_i^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{N} Y_i^2 + \frac{\beta}{\sigma^2} \sum_{i=1}^{N} Y_i x_i.$$

Taking the first order derivative w.r.t. β and setting it to zero, we have

$$\frac{\partial \log \mathcal{L}_N}{\partial \beta} = -\frac{\hat{\beta}_{ML}}{\sigma^2} \sum_{i=1}^{N} x_i^2 + \frac{1}{\sigma^2} \sum_{i=1}^{N} Y_i x_i = 0.$$

From it, we obtain the MLE of β,

$$\hat{\beta}_{ML} = \frac{\sum_{i=1}^{N} Y_i x_i}{\sum_{i=1}^{N} x_i^2}.$$

To verify that the log-likelihood is indeed maximized, take the second order derivative w.r.t. β,

$$\frac{\partial^2 \log \mathcal{L}_N}{\partial \beta^2} \Big|_{\beta = \hat{\beta}_{ML}} = -\frac{1}{\sigma^2} \sum_{i=1}^{N} x_i^2 < 0.$$

To check whether the estimator is biased,

$$\mathbb{E}[\hat{\beta}_{ML}] = \mathbb{E}\left[\frac{\sum_{i=1}^{N} Y_i x_i}{\sum_{i=1}^{N} x_i^2}\right] = \frac{\sum_{i=1}^{N} \mathbb{E}[Y_i] x_i}{\sum_{i=1}^{N} x_i^2} = \frac{\sum_{i=1}^{N} \beta x_i x_i}{\sum_{i=1}^{N} x_i^2} = \beta,$$

hence, the ML estimator of β is unbiased.

Problem 3

Let $x_1, \dots, x_N$ be a random sample from a gamma(α, β) population.

(a) Find the MLE of β, assuming α is known.

(b) If α and β are both unknown, there is no explicit formula for the MLE of α, but the maximum can be found numerically. How can we use the result in part (a) to reduce the problem to the maximization of a univariate function?

Solution 3

a) Write the likelihood function as

$$\mathcal{L}_N(\beta|x_1,\dots,x_N) = \prod_{i=1}^{N} \frac{1}{\Gamma(\alpha)\beta^\alpha} x_i^{\alpha-1} e^{-x_i/\beta} = \frac{1}{\Gamma(\alpha)^N \beta^{N\alpha}} \left[\prod_{i=1}^{N} x_i\right]^{\alpha-1} e^{-\sum_i x_i/\beta},$$

and taking the logarithm,

$$\log \mathcal{L}_N(\beta|x_1,\dots,x_N) = -\log \Gamma(\alpha)^N - N\alpha \log \beta + (\alpha-1)\log\left[\prod_{i=1}^{N} x_i\right] - \frac{\sum_i x_i}{\beta}.$$

Taking the first order derivative w.r.t. β and setting to zero, we have

$$\frac{\partial \log \mathcal{L}_N}{\partial \beta} = -\frac{N\alpha}{\hat{\beta}_{ML}} + \frac{\sum_i x_i}{\hat{\beta}_{ML}^2} = 0.$$

From the expression above, we get the ML estimator of β,

$$\hat{\beta}_{ML} = \frac{\sum_{i=1}^{N} x_i}{N\alpha} = \frac{\bar{x}}{\alpha}.$$

To check that this is a maximum, calculate

$$\frac{\partial^2 \log \mathcal{L}_N}{\partial \beta^2}\Big|_{\beta=\hat{\beta}_{ML}} = \frac{N\alpha}{\beta^2} - \frac{2\sum_i x_i}{\beta^3}\Big|_{\beta=\hat{\beta}_{ML}} = \frac{(N\alpha)^3}{(\sum_i x_i)^2} - \frac{2(N\alpha)^3}{(\sum_i x_i)^2} = -\frac{(N\alpha)^3}{(\sum_i x_i)^2} < 0.$$

Because $\hat{\beta}_{ML}$ is the unique point where the derivative is zero and it is a local maximum, it is a global maximum. That is, $\hat{\beta}_{ML}$ is the MLE.

b) Now the likelihood function is

$$\mathcal{L}_N(\alpha,\beta|x_1,\dots,x_N) = \frac{1}{\Gamma(\alpha)^N \beta^{N\alpha}} \left[\prod_{i=1}^{N} x_i\right]^{\alpha-1} e^{-\sum_i x_i/\beta},$$

the same as in part (a) except α and β are both variables now. There is no closed form for the MLEs $\hat{\alpha}_{ML}$ and $\hat{\beta}_{ML}$. One approach to finding $\hat{\alpha}_{ML}$ and $\hat{\beta}_{ML}$ would be to numerically maximize the function of two arguments. But it is usually best to do as much as possible analytically, first, and perhaps reduce the complexity of the numerical problem. From part (a), for each fixed value of α, the value of β that maximizes $\mathcal{L}(\alpha,\beta|x_1,\dots,x_N)$ is $\sum_i x_i/N\alpha$. Substituting this into the likelihood function, we are left with one variable α,

$$\begin{aligned}\mathcal{L}(\alpha|x_1,\dots,x_N) &= \frac{1}{\Gamma(\alpha)^N (\sum_i x_i/N\alpha)^{N\alpha}} \left[\prod_{i=1}^{N} x_i\right]^{\alpha-1} e^{-\sum_i x_i/(\sum_i x_i/N\alpha)} \\ &= \frac{1}{\Gamma(\alpha)^N (\sum_i x_i/N\alpha)^{N\alpha}} \left[\prod_{i=1}^{N} x_i\right]^{\alpha-1} e^{-N\alpha} \to \max_{\alpha}.\end{aligned}$$

From the problem defined above, we get the ML estimator $\hat{\alpha}_{ML}$ which can be used to compute $\hat{\beta}_{ML}$.