25.02.2021 **Week 1**

Exercise 1 Consider a bivariate Pareto density :

$$f(x,y) = c(x+y-1)^{-p-2}, \text{ for } x,y>1, \text{ and } p>2.$$

1. Show that c is equal to $p(p+1)$.
2. Determine the marginal laws of this density and compute $\mathbb{E}[X]$.
3. Calculate the variance-covariance matrix Σ.
4. Consider a sample $(X_1,Y_1)',\ldots,(X_n,Y_n)'$ of independent and identically distributed random vectors following the Pareto density with parameters p. Estimate the parameter p using the maximum likelihood method.

Solution 1 Consider a bivariate Pareto density :

$$f(x,y) = c(x+y-1)^{-p-2}, \quad \text{for } x,y>1, \text{ and } p>2.$$

1. We have that
$$1 = \int_1^\infty \int_1^\infty c(x+y-1)^{-p-2} dx dy = \int_1^\infty \frac{c}{p+1} y^{-p-1} dy = \frac{c}{p(p+1)},$$
so $c = p(p+1)$.
2. By integration, we obtain
$$f(x) = \int_1^\infty f(x,y) dy = p x^{-p-1}$$
, for $x>1$, and the same for $f(y)$.
The calculation of the expected values gives $\mathbb{E}[X] = \mathbb{E}[Y] = \displaystyle\int_1^\infty x p x^{-p-1} dx = \frac{p}{p-1}$.
3. We first compute $\mathbb{E}[XY]$:
$$\begin{aligned}
\mathbb{E}[XY] &= \int_1^\infty \int_1^\infty cxy(x+y-1)^{-p-2} dx dy \\
&= \int_1^\infty \Big(\frac{1}{-p-1} xy(x+y-1)^{-p-1}\Big|_1^\infty\Big) dy + \int_1^\infty \int_1^\infty c\frac{y}{p+1}(x+y-1)^{-p-1} dx dy \\
&= \int_1^\infty \Big(\frac{c}{p+1} y^{-p} dy + \int_1^\infty c\Big(\frac{y}{-p(p+1)}(x+y-1)^{-p}\Big|_1^\infty\Big) dy \\
&= \frac{p}{p-1} + c\int_1^\infty \frac{y^{-p+1}}{p(p+1)} dy \\
&= \frac{p}{p-1} + \frac{1}{p-2} = \frac{p^2-p-1}{(p-1)(p-2)}.
\end{aligned}$$

We then obtain
$$\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \frac{p^2-p-1}{(p-1)(p-2)} - \frac{p^2}{(p-1)^2} = \frac{1}{(p-1)^2(p-2)}.$$
We now compute
$$\mathbb{E}\left[X^2\right] = \int_1^\infty x^2 p x^{-p-1} dx = \frac{p}{p-2},$$
and finally
$$\mathrm{Var}(X) = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = \frac{p}{(p-1)^2(p-2)}.$$
We obtain the variance of Y by following the same steps : $\mathrm{Var}(Y) = \frac{p}{(p-1)^2(p-2)}$. We find that $\Sigma = \dfrac{1}{(p-1)^2(p-2)}\begin{pmatrix} p & 1 \\ 1 & p \end{pmatrix}$.

d) The log-likelihood is :

$$L = n \log p + n \log(p+1) + \sum_{i=1}^{n} (-p-2) \log(x_i + y_i - 1).$$

By derivation with respect to p we have

$$\frac{n}{p} + \frac{n}{p+1} - \sum_{i=1}^{n} \log(x_i + y_i - 1) = 0,$$

and by setting $\bar{\alpha} = \frac{1}{n} \sum_{i=1}^{n} \log(x_i + y_i - 1)$ we have the equation

$$p^2 \bar{\alpha} + p(\bar{\alpha} - 2) - 1 = 0.$$

The product of the roots of this equation $-(\bar{\alpha})^{-1}$ is negative. So we only consider the positive root, $\hat{p} = \frac{1}{\bar{\alpha}} - \frac{1}{2} + \sqrt{\frac{1}{\bar{\alpha}^2} + \frac{1}{4}}$.

Exercise 2 Let X_1 and X_2 be two independent Gamma random variables with common scale parameters : $X_1 \sim \text{Gamma}(\alpha, \lambda)$ and $X_2 \sim \text{Gamma}(\beta, \lambda)$. Define

$$\begin{aligned} Y_1 &= X_1 + X_2 \\ Y_2 &= \frac{X_1}{X_1 + X_2} \end{aligned}$$

1. Write the joint density of (X_1, X_2).
2. Determine the joint density of (Y_1, Y_2).
3. Deduce the marginal distributions of Y_1 and Y_2.

Solution 2

1. X_1 and X_2 are independent so

$$f_{\mathbf{X}}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2) = \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \lambda^{\alpha+\beta} x_1^{\alpha-1} x_2^{\beta-1} \exp[-\lambda(x_1 + x_2)], \text{ for } x_1, x_2 > 0$$

2. We look for the function $\mathbf{h}$ such that $\mathbf{Y} = \mathbf{h}(\mathbf{X})$.

$$\mathbf{h}(\mathbf{x}) = \begin{pmatrix} x_1 + x_2 \\ x_1/(x_1 + x_2) \end{pmatrix}$$

and

$$\mathbf{h^{-1}}(\mathbf{y}) = \begin{pmatrix} y_1 y_2 \\ y_1(1 - y_2) \end{pmatrix}.$$

The Jacobian of $\mathbf{h^{-1}}$ is $J_{h^{-1}}(y_1, y_2) = -y_1$. Now,

$$\begin{aligned} f_{\mathbf{Y}}(y_1, y_2) &= f_{\mathbf{X}}\left(\mathbf{h^{-1}}\left(\mathbf{Y}\right)\right) \mid J_{h^{-1}}\left(\mathbf{Y}\right) \mid \\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \lambda^{\alpha+\beta} (y_1 y_2)^{\alpha-1} \left(y_1 \left(1 - y_2\right)\right)^{\beta-1} \exp\left[-\lambda \left(y_1 y_2 + y_1 \left(1 - y_2\right)\right)\right] y_1, \end{aligned}$$

for $\mathbf{y} \in]0, \infty[\times]0, 1[$.

3. We see that we can separate the joint density $f_{\mathbf{Y}}$ into two terms $g_1(y_1)$ and $g_2(y_2)$ where

$$g_1(y_1) = \frac{\lambda^{\alpha+\beta} y_1^{\alpha+\beta-1} \exp[-\lambda y_1]}{\Gamma(\alpha + \beta)}$$

and

$$g_2(y_2) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} y_2^{\alpha-1} (1 - y_2)^{\beta-1},$$

for $0 < y_1 < \infty$ and $0 < y_2 < 1$. Hence, Y_1 is independent of Y_2. Moreover, $Y_1 \sim \text{Gamma}(\alpha + \beta, \lambda)$ and $Y_2 \sim$Beta(α, β).

Exercise 3 Suppose that $\mathbf{X}_1, \ldots, \mathbf{X}_n \in \mathbb{R}^p$ are independent and identically distributed random vectors following a multivariate Gaussian distribution $N_p(\boldsymbol{\mu}, \Sigma)$. We consider the sample mean

$$\bar{\mathbf{X}} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i$$

and the sample variance-covariance matrix

$$S = \frac{1}{n} \sum_{i=1}^{n} \left(\mathbf{X}_i - \bar{\mathbf{X}}\right) \left(\mathbf{X}_i - \bar{\mathbf{X}}\right)'$$

1. Show that $\bar{\mathbf{X}}$ is an unbiased estimate of $\boldsymbol{\mu}$. (i.e. $E\left[\bar{\mathbf{X}}\right] = \boldsymbol{\mu}$).
2. Show that $E\left[S\right] = \dfrac{n}{n-1}\Sigma$. Propose another estimate of Σ which is not biased.

Solution 3

1.

$$\begin{aligned} E\left[\bar{\mathbf{X}}\right] &= E\left[\frac{1}{n}\sum_i \mathbf{X}_i\right] \\ &= \frac{1}{n}\sum_i E\left[\mathbf{X}_i\right] \\ &= \frac{1}{n}\sum_i \boldsymbol{\mu} = \boldsymbol{\mu}. \end{aligned}$$

2.

$$\begin{aligned} S &= \frac{1}{n}\sum_i \mathbf{X}_i\mathbf{X}_i' - \bar{\mathbf{X}}\bar{\mathbf{X}}' \\ &= \frac{1}{n}\sum_i (\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_i - \boldsymbol{\mu})' - (\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' \\ &= \left(\frac{1}{n} - \frac{1}{n^2}\right)\sum_i (\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_i - \boldsymbol{\mu})' - \frac{1}{n^2}\sum_{i\neq j}(\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})'. \end{aligned}$$

Since $E\left[(\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})'\right] = 0$ for $i \neq j$, we have

$$E[S] = \frac{n-1}{n}\Sigma.$$

Hence, S is a biased estimate of Σ. If we consider $\tilde{S} = \dfrac{n}{n-1}S$ then $E[\tilde{S}] = \Sigma$.

Exercise 4 We consider a matrix $\Sigma \in \mathbb{R}^{p\times p}$ and we write

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \text{ and } \Sigma^{-1} = \Psi = \begin{pmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{pmatrix}.$$

Show the following equations :

(a) $\Sigma_{12}\Sigma_{22}^{-1} = -\Psi_{11}^{-1}\Psi_{12}$

(b) $\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = \Psi_{11}^{-1}$

Solution 4 We assume that $\dim(\Sigma_{11}) = m \times m$ and $\dim(\Sigma_{22}) = (p-m)\times(p-m)$. Since $\Psi\Sigma = \mathrm{I}_p$, we can write

$$\begin{aligned}
\Psi_{11}\Sigma_{11} + \Psi_{12}\Sigma_{21} &= \mathrm{I}_m \\
\Psi_{21}\Sigma_{12} + \Psi_{22}\Sigma_{22} &= \mathrm{I}_{p-m} \\
\Psi_{21}\Sigma_{11} + \Psi_{22}\Sigma_{21} &= 0_{p-m,m} \\
\Psi_{11}\Sigma_{12} + \Psi_{12}\Sigma_{22} &= 0_{m,p-m}
\end{aligned}$$

The matrices Σ_{ii} and Ψ_{ii} are invertible for $i = 1, 2$. So,

(a) $\Sigma_{12}\Sigma_{22}^{-1} = -\Psi_{11}^{-1}\Psi_{12}$ using the last equation,

(b) $\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = \Psi_{11}^{-1}$ by replacing in the first equation the Ψ_{12} as obtained from the last one.