23.11.2021

Seminar W9 - 831

$X_1, X_2, \ldots, X_n$ i.i.d. , $X_i \sim X$

We want to find estimators for the parameters of the law of X

method of moments: $\overline{\nu}_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$

$$\nu_k = \begin{cases} \sum_{i=1}^{n} x^k \cdot f_X(x) \\ \int_{\mathbb{R}} x^k \cdot f_X(x)\, dx, \text{ if } X \text{ continuous} \end{cases}$$

method of maximal likelihood

$\hookrightarrow$ minimising the likelihood function

Exercise 1. Using the method of moments, find estimators for the parameter(s) of the distribution of the characteristic X, if:

(a) $X \sim Geo(p)$, $p \in (0,1)$;

(b) $X \sim Exp(\lambda)$, $\lambda > 0$;

(c) $X \sim Bino(n,p)$, $n \in \mathbb{N}^*$, $p \in (0,1)$;

(d) $X \sim Unif[a,b]$, $a < b$;

(e) $X \sim NBin(n,p)$, $n \in \mathbb{N}^*$, $p \in (0,1)$;

(f) $X \sim Gamma(a,b)$, $a, b > 0$;

(g) $X \sim Pareto(\alpha,\beta)$, $\alpha > 2$, $\beta > 0$;

- *If $X \sim Gamma(a,b)$, $a,b > 0$, then:*

$$f_X(x) = \frac{1}{\Gamma(a) b^a} \cdot x^{a-1} e^{\frac{-x}{b}} \cdot 1_{(0,\infty)}(x)$$

where:

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx, \; a > 0$$

$E(X) = ab$, $V(X) = ab^2$

Sol. (f) $X \sim Gamma(a,b)$.

$$\begin{cases} \overline{\nu}_1 = \nu_1 \\ \overline{\nu}_2 = \nu_2 \end{cases}$$

$\nu_1 = E(X) = ab$

$V(X) = ab^2$

$V(X) = E(X^2) - E(X)^2 = \nu_2 - \nu_1^2$

$V(X) = ab^2 \Rightarrow \bar{\nu}_2 = V(X) + \bar{\nu}_1^2 = ab^2 + (ab)^2$

$$\Rightarrow \begin{cases} \bar{\nu}_1 = ab \\ \bar{\nu}_2 = ab^2 + (ab)^2 \end{cases}$$

We solve the system, considering a and b as our variables

$$\Rightarrow \begin{cases} a = \frac{\bar{\nu}_1}{b} \\ \bar{\nu}_2 = ab^2 + (ab)^2 \end{cases} \Rightarrow \begin{cases} a = \frac{\bar{\nu}_1}{b} \\ \bar{\nu}_2 = \frac{\bar{\nu}_1}{b} \cdot b^2 + \bar{\nu}_1^2 \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} a = \frac{\bar{\nu}_1}{b} \\ \bar{\nu}_2 = b\bar{\nu}_1 + \bar{\nu}_1^2 \end{cases} \Rightarrow \begin{cases} b = \frac{\bar{\nu}_2 - \bar{\nu}_1^2}{\bar{\nu}_1} \\ a = \frac{\bar{\nu}_1}{b} \end{cases} \Rightarrow \begin{cases} b = \frac{\bar{\nu}_2 - \bar{\nu}_1^2}{\bar{\nu}_1} \\ a = \frac{\bar{\nu}_1^2}{\bar{\nu}_2 - \bar{\nu}_1^2} \end{cases}$$

Therefore the estimators for a and b are:

$$\bar{a} = \frac{\bar{\nu}_1^2}{\bar{\nu}_2 - \bar{\nu}_1^2} \qquad \bar{b} = \frac{\bar{\nu}_2 - \bar{\nu}_1^2}{\bar{\nu}_1}$$

(a) $X \sim Geo(p)$

$\bar{\nu}_1 = \nu_1 \qquad \nu_1 = E(X) = \frac{1-p}{p}$

$\bar{\nu}_1 = \frac{1-p}{p} \Rightarrow p\bar{\nu}_1 = 1-p \Rightarrow p\bar{\nu}_1 + p = 1 \Rightarrow p(\bar{\nu}_1 + 1) = 1 \Rightarrow$

$\Rightarrow p = \frac{1}{\bar{\nu}_1 + 1} \Rightarrow \bar{p} = \frac{1}{\bar{\nu}_1 + 1}$

(b) $X \sim Exp(\lambda)$

$$E(X) = \frac{1}{\lambda} \qquad \left(V(X) = \frac{1}{\lambda^2} \quad \text{to whom it may concern} \right)$$

$$\overline{\nu_1} = E(X) = \frac{1}{\lambda} \qquad \text{"new"} \quad \text{"nu"}$$

$$\Rightarrow \bar{\lambda} = \frac{1}{\overline{\nu_1}} = \frac{n}{X_1 + X_2 + \dots + X_n}$$

(g) $X \sim \text{Pareto}(\alpha, \beta)$

$$\nu_2 = E(X^2) = V(X) + (E(X))^2$$

$$\begin{cases} \overline{\nu_1} = \nu_1 = E(X) = \dfrac{\alpha\beta}{\alpha - 1} \\ \overline{\nu_2} = \nu_2 = V(X) + E(X)^2 = \dfrac{\alpha\beta^2}{(\alpha-1)^2 \cdot (\alpha - 2)} + \dfrac{\alpha^2 \beta^2}{(\alpha-1)^2} = \end{cases}$$

$$= \frac{\alpha\beta^2}{(\alpha-1)^2} \left(\frac{1}{\alpha - 2} + \alpha \right)$$

$$\begin{cases} \beta = \dfrac{\overline{\nu_1} \cdot (\alpha - 1)}{\alpha} \\ \overline{\nu_2} = \dfrac{\alpha}{(\alpha-1)^2} \cdot \dfrac{\alpha^2 - 2\alpha + 1}{\alpha - 2} \cdot \dfrac{\overline{\nu_1}^2 \cdot (\alpha-1)^2}{\alpha^2} \end{cases}$$

$$\Rightarrow \begin{cases} \beta = \dfrac{\overline{\nu_1} \cdot (\alpha - 1)}{\alpha} \\ \overline{\nu_2} = \dfrac{\alpha \cdot (\alpha - 1)^2 \cdot \overline{\nu_1}^2}{\alpha^2 \cdot (\alpha - 2)} \end{cases}$$

$$\Rightarrow \begin{cases} \beta = \dfrac{\overline{\nu_1} \cdot (\alpha - 1)}{\alpha} \\ \overline{\nu_2} = \dfrac{(\alpha - 1)^2}{\alpha(\alpha - 2)} \cdot \overline{\nu_1}^2 \end{cases} \Rightarrow \frac{\alpha^2 - 2\alpha + 1}{\alpha^2 - 2\alpha} = \frac{\overline{\nu_2}}{\overline{\nu_1}^2} \Rightarrow$$

$$\Rightarrow \frac{\alpha^2 - 2\alpha + 1}{(\alpha^2 - 2\alpha + 1) - (\alpha^2 - 2\alpha)} = \frac{\overline{\nu_2}}{\overline{\nu_2} - \overline{\nu_1}^2}$$

$$\Rightarrow (\alpha-1)^2 = \frac{\overline{\nu_2}}{\overline{\nu_2}-\overline{\nu_1}^2} \Rightarrow \alpha = 1 + \sqrt{\frac{\overline{\nu_2}}{\overline{\nu_2}-\overline{\nu_1}^2}}$$

$$\Rightarrow \beta = \overline{\nu_1} \cdot \frac{\alpha-1}{\alpha} = \overline{\nu_1} \cdot \frac{\sqrt{\frac{\overline{\nu_2}}{\overline{\nu_2}-\overline{\nu_1}^2}}}{1+\sqrt{\frac{\overline{\nu_2}}{\overline{\nu_2}-\overline{\nu_1}^2}}} =$$

$$= \frac{\overline{\nu_1}\sqrt{\overline{\nu_2}}}{\sqrt{\overline{\nu_2}-\overline{\nu_1}^2}+\sqrt{\overline{\nu_2}}}$$

(c) $X \sim Bin(n,p)$

$$E(X) = np \qquad V(X) = np(1-p)$$

$$\begin{cases} \overline{\nu_1} = \nu_1 \\ \overline{\nu_2} = \nu_2 \end{cases} \qquad \nu_2 = V(X) + E(X)^2$$

$$\nu_1 = np \qquad \nu_2 = np(1-p) + (np)^2$$

$$\Rightarrow \begin{cases} \overline{\nu_1} = np \\ \overline{\nu_2} = np(1-p) + (np)^2 \end{cases} \Leftrightarrow \begin{cases} n = \frac{\overline{\nu_1}}{p} \\ \overline{\nu_2} = \frac{\overline{\nu_1}}{p} \cdot p \cdot (1-p) + \frac{\overline{\nu_1}^2}{p^2} \cdot p^2 \end{cases}$$

$$\Leftrightarrow \begin{cases} n = \frac{\overline{\nu_1}}{p} \\ \overline{\nu_2} = \overline{\nu_1}(1-p) + \overline{\nu_1}^2 \end{cases} \Leftrightarrow \begin{cases} n = \frac{\overline{\nu_1}}{p} \\ 1-p = \frac{\overline{\nu_2}-\overline{\nu_1}^2}{\overline{\nu_1}} \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} n = \frac{\overline{\nu_1}}{p} \\ p = 1 - \frac{\overline{\nu_2}-\overline{\nu_1}^2}{\overline{\nu_1}} \end{cases} \Rightarrow \overline{p} = \frac{\overline{\nu_1}-\overline{\nu_2}+\overline{\nu_1}^2}{\overline{\nu_1}}$$

$$\overline{n} = \frac{\overline{\nu_1}^2}{\overline{\nu_1}-\overline{\nu_2}+\overline{\nu_1}^2}$$

Method of maximum likelihood

$X_1, \dots, X_n$ i.i.d. $\qquad \theta = (\theta_1, \dots, \theta_k)$

$$L(X_1, X_2, \dots, X_n ; \theta) = \prod_{i=1}^{n} f_{X_i}$$

θ maximum likelihood estimator if it achieves the maximum for the likelihood function L

In order to do this, it suffices to find the value of θ for which $\ln L$ achieves its maximum.

We will find this value of θ by imposing the condition

$$\frac{\partial \ln L}{\partial \theta_i} = 0$$

Exercise 2. Using the maximum likelihood method, estimate the parameters of the distribution of the characteristic X, if:

(a) $X \sim Unif[a, b]$, $a < b$, for the sample data: $x_1, \dots, x_n \in [a, b]$.

(b) $X \sim Bino(10, p)$, $p \in (0, 1)$, for the sample data: $x_1, \dots, x_n \in \{0, 1, \dots, 10\}$ not all zero and not all ten.

(c) $X \sim Exp(\lambda)$, $\lambda > 0$, for the sample data: $x_1, \dots, x_n > 0$.

(d) $X \sim Geo(p)$, $p \in (0, 1)$, for the sample data: $x_1, \dots, x_n \in \mathbb{N}$ not all zero.

- *If $X \sim Bino(n,p)$, $n \in \mathbb{N}^*$, $p \in (0,1)$, then:*

$$X\begin{pmatrix} k \\ C_n^k p^k (1-p)^{n-k} \end{pmatrix}_{k\in\overline{0,n}}$$

$E(X) = np$, $V(X) = np(1-p)$

(b) $X \sim Bino(10, p)$

$$L(x_1,\ldots,x_m; p) = \prod_{i=1}^{m} f_{x_i} = \prod_{i=1}^{m} \underbrace{C_{10}^{x_i} \cdot p^{x_i} \cdot (1-p)^{10-x_i}}_{f_{x_i}} =$$

$$= \underbrace{\left(\prod_{i=1}^{m} C_{10}^{x_i}\right)}_{=:K} \cdot p^{x_1+\ldots+x_n} \cdot (1-p)^{10m-(x_1+\ldots+x_m)}$$

$$\ln L(x_1,\ldots,x_m; p) = \ln K + (x_1+\ldots+x_m)\cdot \ln p +$$

$$+ \left(10m - (x_1+\ldots+x_m)\right)\cdot \ln(1-p)$$

$$\frac{\partial \ln L}{\partial p} = \frac{x_1+\ldots+x_m}{p} - \frac{10m-(x_1+\ldots+x_m)}{1-p}$$

Using maximal likelihood we get:

$$\frac{x_1+\ldots+x_m}{p} = \frac{10m-(x_1+\ldots+x_n)}{1-p} \quad \Big| /m$$

$$\Rightarrow \frac{\overline{\nu}_1}{p} = \frac{10-\overline{\nu}_1}{1-p} \Rightarrow \frac{\overline{\nu}_1}{10-\overline{\nu}_1} = \frac{p}{1-p} \Rightarrow$$

$$\Rightarrow \frac{\overline{\nu}_1}{10-\overline{\nu}_1+\overline{\nu}_1} = \frac{p}{1-p+p} \Rightarrow p = \frac{\overline{\nu}_1}{10} \Rightarrow \overline{p} = \frac{1}{10}\cdot\overline{\nu}_1$$

(c) $X \sim \mathcal{E}xp(\lambda)$

(d) $X \sim Geo(p)$

- *If $X \sim Exp(\lambda)$, $\lambda > 0$ then:*

$$f_X(x) = \lambda e^{-\lambda x} \cdot 1_{[0,\infty)}(x)$$

$$E(X) = \frac{1}{\lambda},\ V(X) = \frac{1}{\lambda^2}$$

(c) $X \sim \mathcal{E}xp(\lambda)$

$$L(x_1, \dots, x_n; \lambda) = \prod_{i=1}^{n} f_{x_i} = \prod_{i=1}^{n} \left(\lambda \cdot e^{-\lambda x_i} \cdot 1_{[0,\infty)}(x_i) \right) =$$

$$= \lambda^n \cdot \underbrace{\left(\prod_{i=1}^{n} e^{-\lambda x_i} \right)}_{= e^{-\lambda(x_1 + x_2 + \dots + x_n)}} \cdot \underbrace{\prod_{i=1}^{n} 1_{[0,\infty)}(x_i)}_{=: K} =$$

$$= K \lambda^n \cdot e^{-\lambda(x_1 + \dots + x_n)}$$

$$\Rightarrow \ln L(x_1; \dots, x_n; \lambda) = \ln K + n \ln(\lambda) - \lambda(x_1 + \dots + x_n)$$

$$\Rightarrow \frac{\partial \ln L}{\partial \lambda} = -(x_1 + \dots + x_n) + \frac{n}{\lambda}$$

$$\Rightarrow \frac{n}{\lambda} = x_1 + \dots + x_n \Rightarrow \lambda = \frac{n}{x_1 + \dots + x_n}$$

$$\Rightarrow \bar{\lambda} = \frac{n}{x_1 + \dots + x_n} = \frac{1}{\bar{x}} = \frac{1}{\bar{\nu}_1}$$

(d) $X \sim Geo(p)$

- *If $X \sim Geo(p)$, $p \in (0,1)$, then*

$$X \begin{pmatrix} k \\ p(1-p)^k \end{pmatrix}_{k \in \mathbb{N}}$$

$$E(X) = \frac{1-p}{p},\ V(X) = \frac{1-p}{p^2}$$

$$L(x_1, \dots, x_n; p) = \prod_{i=1}^{n} f_{x_i} = \prod_{i=1}^{n} p(1-p)^{x_i} = p^n \cdot \prod_{i=1}^{n} (1-p)^{x_i} =$$

$$= p^n \cdot (1-p)^{x_1+\dots+x_n}$$

$$\ln L = n \ln(p) + \ln(1-p) \cdot (x_1+\dots+x_n)$$

$$\frac{\partial \ln L}{\partial p} = \frac{n}{p} - \frac{1}{1-p} \cdot (x_1+\dots+x_n)$$

$$\Rightarrow \frac{n}{p} - \frac{1}{1-p}(x_1+\dots+x_n) = 0$$

$$\Rightarrow \frac{n}{p} = \frac{1}{1-p}(x_1+\dots+x_n)$$

$$\Rightarrow n(1-p) = p(x_1+\dots+x_n)$$

$$\Rightarrow p(n+x_1+\dots+x_n) = n$$

$$\Rightarrow p = \frac{n}{n+x_1+\dots+x_n}$$

$$\Rightarrow \bar{p} = \frac{n}{n+x_1+\dots+x_n} = \frac{1}{1+\bar{X}} = \frac{1}{1+\bar{x}_n}$$

(a) $X \sim \text{Unif}[a,b] \Rightarrow f_X = \frac{1}{b-a} \cdot 1_{[a,b]}$

$$L(x_1,\dots,x_n; a,b) = \prod_{i=1}^{n} \frac{1}{b-a} = \frac{1}{(b-a)^n}$$

The maximum of L is achieved when $b-a$ achieves its minimum

$x_1,\dots,x_n \in [a,b] \Rightarrow$ $\bar{a} = \min\{x_1,\dots,x_n\}$

$\bar{b} = \max\{x_1,\dots,x_n\}$