Topics in statistics III/IV (MATH3361/4071) Michaelmas term, 2020-2021

Problem class handout 3: Likelihood methods

Lecturer: Georgios P. Karagiannis georgios.karagiannis@durham.ac.uk

Exercise 1. Consider random sample $x_1, ..., x_n \overset{IID}{\sim} \mathrm{G}(a, b)$, $a > 0$, $b > 0$ with PDF

$$f(x|a, b) = \frac{1}{\Gamma(a)b^a} x^a e^{-x\frac{1}{b}} 1(x > 0)$$

1. Find the moment estimator $\tilde{\theta}$ of $\theta = (a, b)^T$ by using the first raw moment and the first central moment

2. Is the moment estimator $\tilde{\theta}$ consistent and asymptotically Normal?

3. Find the one step estimator by Fisher scoring algorithm.

Hint-1 Digamma function $\psi(x) = \frac{\mathrm{d}}{\mathrm{d}x} \log \Gamma(x)$

Hint-2 Trigamma function $\psi_1(x) = \frac{\mathrm{d}^2}{\mathrm{d}x^2} \log \Gamma(x)$

Hint-3 $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

Solution.

1. The first raw moment is the expected value/mean, and the first central moment is the variance.

   The first raw moment is

   $$\mathrm{E}(x) = \int_0^1 x \frac{1}{\Gamma(a)b^a} x^{a-1} e^{-x/b} \mathrm{d}x = \int_0^1 \frac{1}{\frac{1}{a}\Gamma(a+1)\frac{1}{b}b^{a+1}} x^{(a+1)-1} e^{-x/b} \mathrm{d}x = ab$$

   and the sample one

   $$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

   The first central moment is

   $$\mathrm{var}(x) = \mathrm{E}(x^2) - (\mathrm{E}(x))^2$$

   So

   $$\mathrm{E}(x^2) = \int_0^1 x^2 \frac{1}{\Gamma(a)b^a} x^{a-1} e^{-x/b} \mathrm{d}x = \int_0^1 \frac{1}{\frac{1}{a(a+1)}\Gamma(a+2)\frac{1}{b^2}b^{a+2}} x^{(a+2)-1} e^{-x/b} \mathrm{d}x = a(a+1)b^2$$

and hence

$$\mathrm{var}(x) = \mathrm{E}(x^2) - (\mathrm{E}(x))^2 = ab^2$$

The sample first central moment is

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

From the method of moments I get

$$\begin{cases} \mathrm{E}(x|\tilde{a},\tilde{b}) = & \bar{x} \\ \mathrm{var}(x|\tilde{a},\tilde{b}) = & s^2 \end{cases} \implies \begin{cases} \tilde{a} = & \frac{\bar{x}^2}{s^2} \\ \tilde{b} = & \frac{\bar{x}^2}{s^2} \end{cases} \implies \begin{cases} \tilde{a} = & \frac{(\mathrm{E}(x))^2}{\mathrm{var}(x)} = \frac{\bar{x}^2}{s^2} \\ \tilde{b} = & \frac{\mathrm{var}(x)}{\mathrm{E}(x)} = \frac{\bar{x}^2}{s^2} \end{cases}$$

So the moment estimator is

$$\tilde{\theta} = \begin{bmatrix} \tilde{a} \\ \tilde{b} \end{bmatrix} = \begin{bmatrix} \frac{\bar{x}^2}{s^2} \\ \frac{s^2}{\bar{x}} \end{bmatrix} \tag{0.1}$$

2. It is consistent because $\tilde{\theta} \xrightarrow{as} \theta$. This is because of the following.

   It is

$$\begin{cases} \mathrm{E}(x) = & ab \\ \mathrm{var}(x) = & ab^2 \end{cases} \implies \begin{cases} a = & \frac{(\mathrm{E}(x))^2}{\mathrm{var}(x)} \\ b = & \frac{\mathrm{var}(x)}{\mathrm{E}(x)} \end{cases} \implies \begin{cases} a = & \frac{(\mathrm{E}(x))^2}{\mathrm{var}(x)} \\ b = & \frac{\mathrm{var}(x)}{\mathrm{E}(x)} \end{cases}$$

   From SLLN, $\bar{x} \xrightarrow{as} \mathrm{E}(x)$. From SLLN, $\overline{x^2} \xrightarrow{as} \mathrm{E}(x^2)$. From Slutsky Theorem, $s^2 = \overline{x^2} - (\overline{x})^2 \xrightarrow{as} \mathrm{E}(x^2) - \mathrm{E}(x)^2 = \mathrm{var}(x)$

   So From Slutsky theorem

$$\tilde{\theta} = \begin{bmatrix} \frac{\bar{x}^2}{s^2} \\ \frac{s^2}{\bar{x}} \end{bmatrix} \xrightarrow{as} \begin{bmatrix} \frac{(\mathrm{E}(x))^2}{\mathrm{var}(x)} \\ \frac{\mathrm{var}(x)}{\mathrm{E}(x)} \end{bmatrix} = \theta$$

   It is asymptotically Normal because of the following.

   $\bar{x}$ and s^2 are asymptotically Normal by the CLT, as averages of IID quantities. Hence, by Delta method, (0.1) is asymptotically Normal.

3. Recall the the one-step estimators

$$\text{Newton alg.} \qquad \breve{\theta}_n = \tilde{\theta}_n - \ddot{\ell}_n(\tilde{\theta}_n)^{-1}\dot{\ell}_n(\tilde{\theta}_n) \tag{0.2}$$

$$\text{Fisher scoring alg.} \qquad \breve{\theta}_n = \tilde{\theta}_n + \frac{1}{n}\mathcal{I}(\tilde{\theta}_n)^{-1}\dot{\ell}_n(\tilde{\theta}_n) \tag{0.3}$$

For the Fisher algorithm, I need to find $\mathcal{I}(\theta)^{-1}$. It is

$$\begin{aligned}
\log f(x|\theta) &= -\log\Gamma(a) - a\log(b) - \frac{1}{b}x + (a-1)\log(x) \\
\frac{\mathrm{d}}{\mathrm{d}\theta}\log f(x|\theta) &= \begin{bmatrix} -\psi(a) - \log(b) + \log(x) \\ -\frac{a}{b} + \frac{1}{b^2}x \end{bmatrix} \\
\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\log f(x|\theta) &= \begin{bmatrix} -\psi_1(a) & -\frac{1}{b} \\ -\frac{1}{b} & -\frac{2x-ab}{b^3} \end{bmatrix} \\
\mathcal{I}(\theta) &= \begin{bmatrix} \psi_1(a) & \frac{1}{b} \\ \frac{1}{b} & \frac{a}{b^2} \end{bmatrix} \\
\mathcal{I}(\theta)^{-1} &= \frac{1}{a\psi_1(a) - 1}\begin{bmatrix} a & -b \\ -b & b^2\psi_1(a) \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\ell_n(\theta) &= -n\log\Gamma(a) - na\log(b) - \frac{1}{b}\sum_{i=1}^n x_i + (a-1)\sum_{i=1}^n \log(x_i) \\
\dot{\ell}_n(\theta) &= \begin{bmatrix} -n\psi(a) - n\log(b) + \sum_{i=1}^n \log(x_i) \\ -n\frac{a}{b} + n\frac{1}{b^2} \end{bmatrix}
\end{aligned}$$

The Fisher recursion is

$$\begin{aligned}
\breve{\theta}_n &= \tilde{\theta}_n + \frac{1}{n}\mathcal{I}(\tilde{\theta}_n)^{-1}\dot{\ell}_n(\tilde{\theta}_n) \\
\breve{\theta}_n &= \begin{bmatrix} \tilde{a} \\ \tilde{b} \end{bmatrix} + \frac{1}{n}\frac{1}{\tilde{a}\psi_1(\tilde{a}) - 1}\begin{bmatrix} \tilde{a} & -\tilde{b} \\ -\tilde{b} & \tilde{b}^2\psi_1(\tilde{a}) \end{bmatrix}\begin{bmatrix} -n\psi(\tilde{a}) - n\log(\tilde{b}) + \sum_{i=1}^n \log(x_i) \\ -n\frac{\tilde{a}}{\tilde{b}} + n\frac{1}{\tilde{b}^2} \end{bmatrix} \\
&= \begin{bmatrix} \tilde{a} \\ \tilde{b} \end{bmatrix} + \frac{1}{\tilde{a}\psi_1(\tilde{a}) - 1}\begin{bmatrix} \tilde{a} & -\tilde{b} \\ -\tilde{b} & \tilde{b}^2\psi_1(\tilde{a}) \end{bmatrix}\begin{bmatrix} -\psi(\tilde{a}) - \log(\tilde{b}) + \frac{1}{n}\sum_{i=1}^n \log(x_i) \\ -\frac{\tilde{a}}{\tilde{b}} + \frac{1}{\tilde{b}^2} \end{bmatrix} \\
&= \begin{bmatrix} \tilde{a} \\ \tilde{b} \end{bmatrix} + \frac{1}{\tilde{a}\psi_1(\tilde{a}) - 1}\begin{bmatrix} -\tilde{a}\psi(\tilde{a}) - \frac{1}{\tilde{b}}(\bar{x} - \tilde{a}\tilde{b}) - \tilde{a}\log(\tilde{b}) + \frac{\tilde{a}}{n}\sum_{i=1}^n \log(x_i) \\ \tilde{b}\psi(\tilde{a}) - \psi_1(\tilde{a})(\bar{x} - \tilde{a}) + \tilde{b}\log(\tilde{b}) - \frac{\tilde{b}}{n}\sum_{i=1}^n \log(x_i) \end{bmatrix}
\end{aligned}$$

So bu substituting

$$\breve{\theta}_n = \begin{bmatrix} \frac{\bar{x}^2}{s^2} \\ \frac{s^2}{\bar{x}} \end{bmatrix} + \frac{1}{\frac{\bar{x}^2}{s^2}\psi_1(\frac{\bar{x}^2}{s^2}) - 1}\begin{bmatrix} -\frac{\bar{x}^2}{s^2}\psi(\frac{\bar{x}^2}{s^2}) - \frac{\bar{x}^2}{s^2}\log(\frac{s^2}{\bar{x}}) + \frac{1}{n}\frac{\bar{x}^2}{s^2}\sum_{i=1}^n \log(x_i) \\ \frac{s^2}{\bar{x}}\psi(\frac{\bar{x}^2}{s^2}) - \psi_1(\frac{\bar{x}^2}{s^2})(\bar{x} - \frac{\bar{x}^2}{s^2}) + \frac{s^2}{\bar{x}}\log(\frac{s^2}{\bar{x}}) - \frac{\tilde{b}}{n}\sum_{i=1}^n \log(x_i) \end{bmatrix}$$

Additionally for the Newton recursion I need

$$
\begin{aligned}
\ddot{\ell}_n(\theta) &= -n \begin{bmatrix} \psi_1(a) & \frac{1}{b} \\ \frac{1}{b} & \frac{2\bar{x}-ab}{b^3} \end{bmatrix} \\
(\ddot{\ell}_n(\theta))^{-1} &= -\frac{1}{n}\frac{1}{\psi_1(a)\frac{2\bar{x}-ab}{b}-1} \begin{bmatrix} \frac{2\bar{x}-ab}{b} & -b \\ -b & b^2\psi_1(a) \end{bmatrix}
\end{aligned}
$$

The Newton recursion is

$$
\begin{aligned}
\breve{\theta}_n &= \tilde{\theta}_n - (\ddot{\ell}_n(\theta))^{-1}\dot{\ell}_n(\tilde{\theta}_n) \\
\breve{\theta}_n &= \begin{bmatrix} \tilde{a} \\ \tilde{b} \end{bmatrix} + \frac{1}{n}\frac{1}{\psi_1(\tilde{a})\frac{2\bar{x}-\tilde{a}\tilde{b}}{\tilde{b}}-1} \begin{bmatrix} 2\bar{x}-\tilde{a}\tilde{b} & \tilde{b} \\ \tilde{b} & \tilde{b}^2\psi_1(\tilde{a}) \end{bmatrix} \begin{bmatrix} -n\psi(\tilde{a}) - n\log(\tilde{b}) + \sum_{i=1}^n \log(x_i) \\ -n\frac{\tilde{a}}{\tilde{b}} + n\frac{1}{\tilde{b}^2} \end{bmatrix} \\
&= \begin{bmatrix} \tilde{a} \\ \tilde{b} \end{bmatrix} + \frac{1}{\psi_1(\tilde{a})\frac{2\bar{x}-\tilde{a}\tilde{b}}{\tilde{b}}-1} \begin{bmatrix} 2\bar{x}-\tilde{a}\tilde{b} & \tilde{b} \\ \tilde{b} & \tilde{b}^2\psi_1(\tilde{b}) \end{bmatrix} \begin{bmatrix} -\psi(\tilde{a}) - \log(\tilde{b}) + \frac{1}{n}\sum_{i=1}^n \log(x_i) \\ -\frac{\tilde{a}}{\tilde{b}} + \frac{1}{\tilde{b}^2} \end{bmatrix} \\
&= \text{...calculations}
\end{aligned}
$$