Annals of Mathematics **00** (XXXX), 1–9
http://dx.doi.org/10.4007/annals.XXXX.00.0.0000

LINMA1731 – Project 2019
Fish schools tracking

By Louis Navarre and Gilles Peiffer

Abstract

In this paper we solve the first part of the project for the class "Stochastic processes: Estimation and prediction" given during the Fall term of 2019. The average speed of each fish in a school of fish is approximated by a gamma-distributed random variable with a shape parameter k and a scale parameter s, and various methods for estimating this quantity are given; a numerical simulation is also included.

Contents

Part 1. **Average speed estimation**

1. **Introduction**

In this paper we give a precise form to this connection, showing that WF

For the purpose of this project, we assume that the speed of each fish in a school at time i is a random variable V_i following a Gamma distribution, as suggested in [1]. This distribution is characterized by two parameters: a shape parameter $k > 0$ and a scale parameter $s > 0$. The parameters are the same for every fish and are time invariant. The aim of this first part is to identify these two parameters using empirical observations v_i.

2. **Maximum likelihood estimation**

Let v_i be i.i.d. realisations of a random variable following a Gamma distribution $\Gamma(k, s)$ (with $i = 1, \ldots, N$). We first assume that the shape parameter k is known.

We start by deriving the maximum likelihood estimator of $\theta \coloneqq s$ based on N observations. Since the estimand θ is a deterministic quantity, we use Fisher estimation. In order to do this, let us restate the probability density function of $V_i \sim \Gamma(k, s)$:

$$f_{V_i}(v_i; k, s) = \frac{1}{\Gamma(k) s^k} v_i^{k-1} \mathrm{e}^{-\frac{v_i}{s}} , \quad i = 1, \ldots, N . \tag{2.1}$$

With this in mind, we can find that the likelihood $\mathcal{L}(v_1, \ldots, v_N; k, \theta)$ is given by

$$\mathcal{L}(v_1, \ldots, v_N; k, \theta) = \prod_{i=1}^{N} f_{V_i}(v_i; k, \theta) \tag{2.2}$$

$$= \prod_{i=1}^{N} \frac{1}{\Gamma(k) \theta^k} v_i^{k-1} \mathrm{e}^{-\frac{v_i}{\theta}} . \tag{2.3}$$

In order to alleviate notation, we compute instead the log-likelihood, which is generally easier to work with[1]:

$$\ell(v_1,\dots,v_N;k,\theta) \coloneqq \ln \mathcal{L}(v_1,\dots,v_N;k,\theta) \tag{2.4}$$

$$= \ln\left(\prod_{i=1}^{N} \frac{1}{\Gamma(k)\theta^k} v_i^{k-1} \mathrm{e}^{-\frac{v_i}{\theta}}\right) \tag{2.5}$$

$$= \sum_{i=1}^{N} \ln\left(\frac{1}{\Gamma(k)\theta^k} v_i^{k-1} \mathrm{e}^{-\frac{v_i}{\theta}}\right) \tag{2.6}$$

$$= (k-1)\sum_{i=1}^{N} \ln v_i - \sum_{i=1}^{N} \frac{v_i}{\theta} - N\big(k\ln\theta + \ln\Gamma(k)\big)\,. \tag{2.7}$$

Now, in order to obtain the maximum likelihood estimate $\widehat{\theta}$, we must differentiate the log-likelihood with respect to the estimand θ, and set it equal to zero:

$$\left.\frac{\partial \ell(v_1,\dots,v_N;k,\theta)}{\partial\theta}\right|_{\theta=\widehat{\theta}} = -\frac{kN}{\widehat{\theta}} + \frac{\sum_{i=1}^{N} v_i}{\widehat{\theta}^2} = 0 \tag{2.8}$$

$$\iff \widehat{\theta} = \frac{\sum_{i=1}^{N} v_i}{kN} = \frac{\overline{v}}{k}\,. \tag{2.9}$$

This then allows us to find the maximum-likelihood estimator $\widehat{\Theta}$, given by

$$\widehat{\Theta} = \frac{\sum_{i=1}^{N} V_i}{kN} = \frac{\overline{V}}{k}\,. \tag{2.10}$$

3. **Properties of the estimator**

We now wish to show some of the properties of this estimator.

3.1. *Asymptotically unbiased.*

Definition 3.1 (Unbiased estimator). The Fisher estimator $\widehat{\Theta} = g(Z)$ of θ is *unbiased* if

$$m_{\widehat{\Theta};\theta} \coloneqq \mathbb{E}\left[g(Z);\theta\right] = \theta\,, \quad \text{for all } \theta\,, \tag{3.1}$$

where

$$\mathbb{E}\left[g(Z);\theta\right] \coloneqq \int_{\operatorname{dom} Z} g(Z) f_Z(z;\theta)\,\mathrm{d}z\,. \tag{3.2}$$

[1]This is possible because the values of θ which maximize the log-likelihood also maximize the likelihood.

PROPERTY 3.1. *The maximum likelihood estimator derived in* (2.10) *is asymptotically unbiased, that is,*

$$\lim_{N\to+\infty} \mathbb{E}\left[g(V_1,\ldots,V_n);\theta\right] = \theta\,. \tag{3.3}$$

Proof. We wish to prove that

$$\lim_{N\to+\infty} \mathbb{E}\left[\frac{\overline{V}}{k}\right] = \theta\,. \tag{3.4}$$

We recall that $\mathbb{E}\left[V_i\right] = k\theta$ for $V_i \sim \Gamma(k,\theta)$ and that the expected value operator is linear to obtain that

$$\mathbb{E}\left[\frac{\overline{V}}{k}\right] = \frac{\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N} V_i\right]}{k} = \frac{\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[V_i\right]}{k} = \frac{\frac{1}{N}Nk\theta}{k} = \theta\,. \tag{3.5}$$

This proves that the maximum likelihood estimator of (2.10) is unbiased, hence it is also asymptotically unbiased. □

3.2. *Efficiency.*

THEOREM 3.2 (Cramér–Rao inequality). *If* $Z = (Z_1,\ldots,Z_N)^T$ *with i.i.d. random variables* Z_k *and if its probability density function given by* $f_Z(z;\theta) = \prod_{k=1}^{N} f_{Z_k}(z_k;\theta)$ *satisfies the following regularity condition*:

$$\mathbb{E}\left[\frac{\partial f_Z(z;\theta)}{\partial\theta}\right] = \int_{-\infty}^{+\infty} \frac{\partial f_Z(z;\theta)}{\partial\theta} f_Z(z;\theta)\,\mathrm{d}z\,, \quad \forall\theta\,, \tag{3.6}$$

then the covariance of any unbiased estimator $\widehat{\Theta}$ *satisfies the* Cramér–Rao inequality

$$\operatorname{cov}\widehat{\Theta} \geq \mathcal{I}^{-1}(\theta)\,, \tag{3.7}$$

where $\mathcal{I}(\theta)$ *is the* $N\times N$ Fisher information matrix, *defined by*

$$\left[\mathcal{I}(\theta)\right]_{i,j} \coloneqq -\mathbb{E}\left[\frac{\partial^2 \ln f_Z(z;\theta)}{\partial\theta_i\partial\theta_j}\right]\,. \tag{3.8}$$

Definition 3.2 (Efficient estimator). An estimator is said to be *efficient* if it reaches the Cramér–Rao bound for all values of θ, that is,

$$\operatorname{cov}\widehat{\Theta} = \mathcal{I}^{-1}(\theta)\,, \quad \forall\theta\,. \tag{3.9}$$

PROPERTY 3.3. *The maximum likelihood estimator derived in* (2.10) *is efficient.*

Proof. □

3.3. *Best asymptotically normal.*

PROPERTY 3.4. *The maximum likelihood estimator is best asymptotically normal.*

Proof. The proof is trivial and left as an exercise to the reader. □

3.4. *Consistent.*

PROPERTY 3.5. *The maximum likelihood estimator is consistent.*

Proof. The proof is trivial and left as an exercise to the reader. □

4. Joint maximum likelihood estimation

We now consider $V_i \sim \Gamma(k, s)$ (for $i = 1, \ldots, N$) with both k and s unknown. Before, we assumed k known, so we could maximize the log-likelihood function with respect to s. Now, we have to maximize this function with respect to s and k at the same time. We know that the maximum likelihood estimator of s, $\hat{s} = f(k)$. Therefore, in the log-likelihood function, we can replace all the occurrences of s by the found estimator, $\hat{s}$. We then get a function of k only; and we could seek for the maximum likelihood estimator of k by derivate this function and equal it to 0. Remember from equation 2.7 the log-likelihood function. If we replace all the occurrences of s by its estimator, we get

(4.1)
$$\begin{aligned}\ell(z;\theta) &= (k-1)\sum_i ln(V_i) - \frac{kN}{\sum_i V_i}\sum_i V_i - Nkln(\frac{\sum_i}{kN}) - Nln(\Gamma(k)) \\ &= (k-1)\sum_i ln(V_i) - kN - Nkln(\sum_i V_i) + Nkln(k) + Nkln(N) - Nln(\Gamma(k)).\end{aligned}$$

Taking the derivative of this function with respect to k, we get

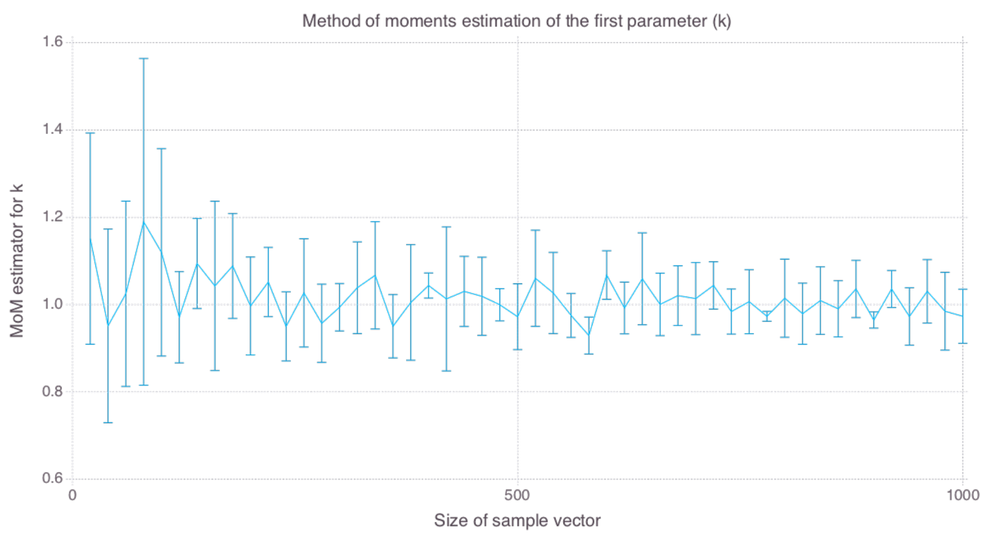
(4.2)
$$\begin{aligned}\left.\frac{\partial \ell(v_1,\ldots,v_N;\theta)}{\partial k}\right|_{k=\hat{k}} &= \sum_i ln(V_i) - N - Nln(\sum_i V_i) + Nln(k) + \frac{Nk}{k} + Nln(N) - N\frac{\Gamma(k)}{\Gamma(k)}\phi^{(0)}(k) \\ &= \sum_i ln(V_i) - Nln(\sum_i V_i) + Nln(k) + Nln(n) - N\phi^{(0)}(k).\end{aligned}$$

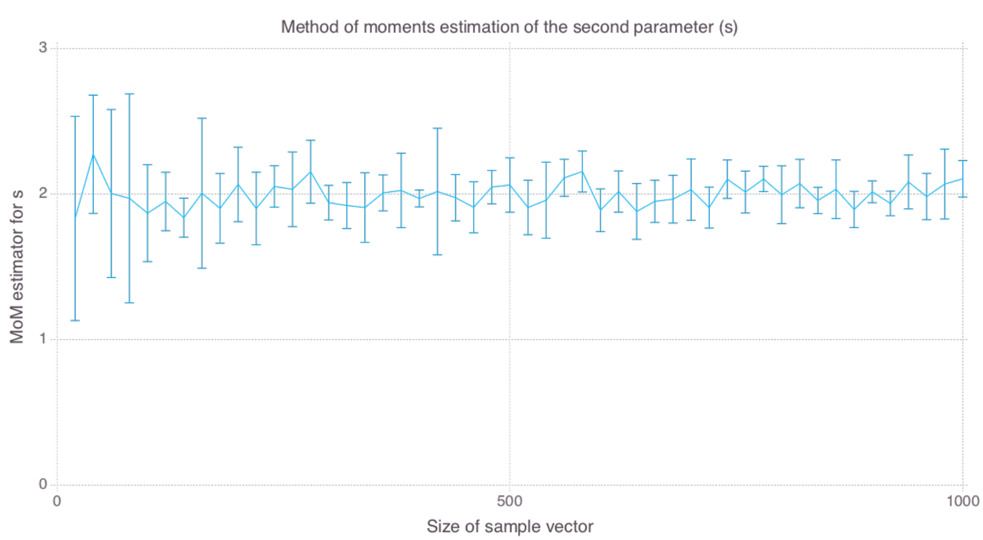
And looking for the root of this derivative, we have

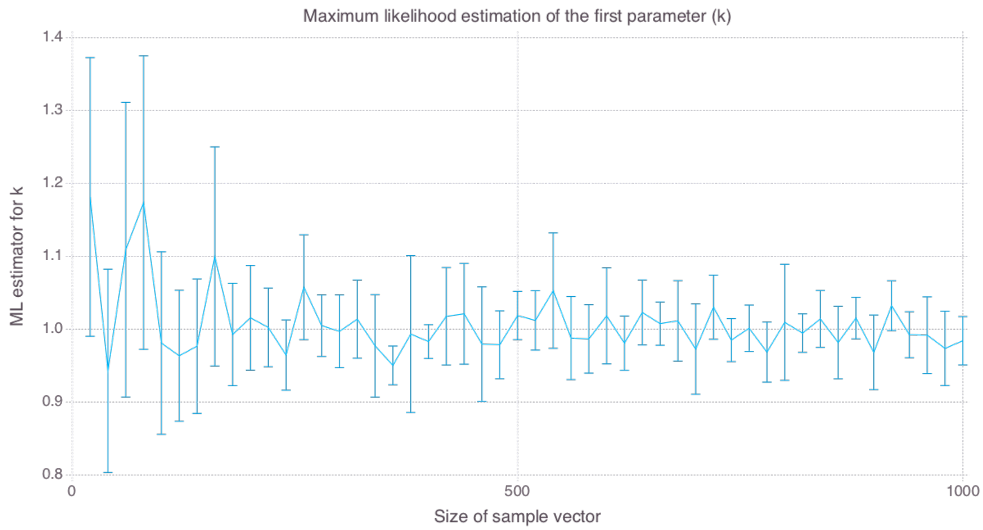
$$\begin{aligned} ln(k) - \phi^{(0)}(k) &= ln(\sum_i V_i) - ln(N) - \frac{\sum_i ln(V_i)}{N} \\ ln(k) - \phi^{(0)}(k) &= ln(\frac{\sum_i V_i}{N}) - \frac{\sum_i ln(V_i)}{N}.\end{aligned} \tag{4.3}$$

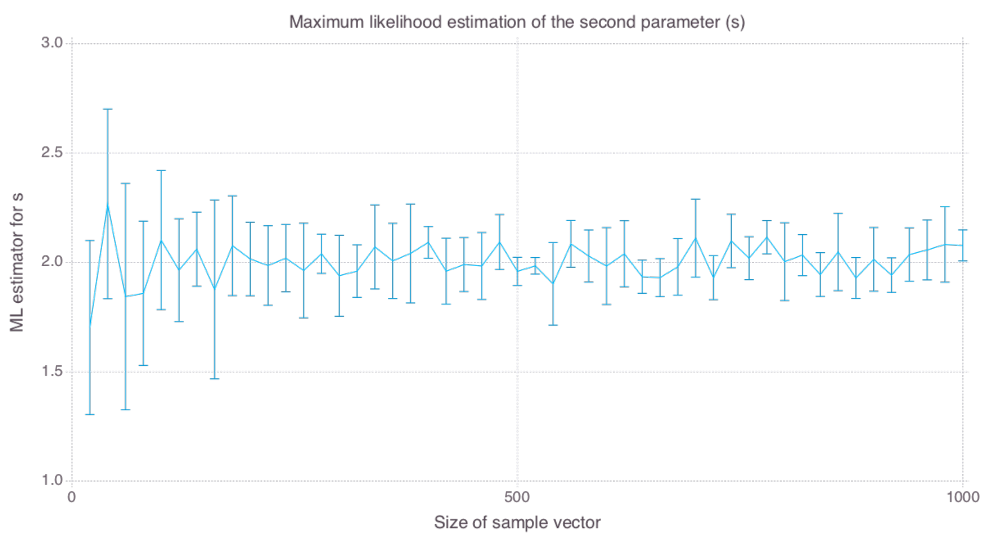
We can't find an analytical solution to this equation, but we can approximate it my numerical methods.

5. Numerical simulation









6. **Fisher information matrix**

We can compute the Fisher information matrix. The entry (i, j) of this matrix is given by equation 3.8. Since we have two estimators, this matrix is a 2×2 matrix. Remember that

$$\begin{aligned} lnf_Z(z,\theta) = h(z,\theta) &= ln\left(\frac{1}{\Gamma(k)s^k}x^{k-1}e^{-x/s}\right) \\ &= ln(x^{k-1}) + ln(e^{-x/s}) - ln(\Gamma(k)) - ln(s^k) \\ &= (k-1)ln(x) - \frac{x}{s} - ln(\Gamma(k)) - kln(s). \end{aligned} \tag{6.1}$$

Therefore, before calculating the entries of the information matrix, we have to compute the partial derivatives before taking the expectation of it. These are given by equations 6.4, 6.5, 6.6 and 6.7

$$\frac{\partial h(z;\theta)}{\partial s} = \frac{x}{s^2} - \frac{k}{s} \tag{6.2}$$

$$\begin{aligned} \frac{\partial h(z;\theta)}{\partial k} &= ln(x) - \frac{1}{\Gamma(k)}\Gamma(k)\phi^{(0)}(k) - ln(s) \\ &= ln(x) - \phi^{(0)}(k) - ln(s). \end{aligned} \tag{6.3}$$

$$\frac{\partial^2 h(z;\theta)}{\partial s^2} = -\frac{x}{s^3} \cdot \frac{1}{2} + \frac{k}{s^2}. \tag{6.4}$$

$$\frac{\partial^2 h(z;\theta)}{\partial k \partial s} = -\frac{1}{s}. \tag{6.5}$$

$$\frac{\partial h(z;\theta)}{\partial s \partial k} = -\frac{1}{s}. \tag{6.6}$$

$$\begin{aligned} \frac{\partial h(z;\theta)}{\partial k^2} &= -\frac{d\phi^{(0)}(k)}{dk} \\ &= -\frac{d^2\Gamma(k)}{dk^2} \\ &= -\phi^{(1)}(k). \end{aligned} \tag{6.7}$$

Then, we take the expectation of these computed values.

$$\begin{aligned} \mathcal{I}_{00} &= -\mathbb{E}\left\{\frac{\partial^2 h(z;\theta)}{\partial s^2}\right\} \\ &= -\mathbb{E}\left\{\frac{x}{s^2} - \frac{k}{s}\right\} \\ &= \frac{k}{s} - \frac{\mathbb{E}\{x\}}{s^2} \\ &= \frac{k}{s} - \frac{ks}{s^2} \\ &= 0. \end{aligned} \tag{6.8}$$

$$\begin{aligned}\mathcal{I}_{01} &= -\mathbb{E}\left\{\frac{\partial^2 h(z;\theta)}{\partial k \partial s}\right\} \\ &= -\mathbb{E}\left\{-\frac{1}{s}\right\} \\ &= \frac{1}{s}.\end{aligned} \tag{6.9}$$

$$\mathcal{I}_{10} = \mathcal{I}_{01} = \frac{1}{s}. \tag{6.10}$$

$$\begin{aligned}\mathcal{I}_{11} &= -\mathbb{E}\left\{\frac{\partial^2 h(z;\theta)}{\partial k^2}\right\} \\ &= -\mathbb{E}\left\{\phi^{(1)}(k)\right\} \\ &= -\phi^{(1)}(k).\end{aligned} \tag{6.11}$$

Finally, the Fisher information matrix is

$$\mathcal{I}(\theta) = \begin{pmatrix} 0 & \frac{1}{s} \\ \frac{1}{s} & -\phi^{(1)}(k) \end{pmatrix}$$

7. Numerical proof

References

Université catholique de Louvain, Ottignies-Louvain-la-Neuve, Belgium
E-mail: navarre.louis@student.uclouvain.be

Université catholique de Louvain, Ottignies-Louvain-la-Neuve, Belgium
E-mail: gilles.peiffer@student.uclouvain.be