

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	2
1. Introduction	2
2. Maximum likelihood estimation	2
3. Properties of the estimator	3
3.1. Asymptotically unbiased	3
3.2. Efficiency	4
3.3. Best asymptotically normal	5
3.4. Consistent	5
4. Joint maximum likelihood estimation	5
5. Numerical simulation	6
6. Fisher information matrix	7
7. Numerical proof	8
Appendix A.	
Analytical derivations for d	8
A.1. Derivation of the Fisher information matrix coefficients	9
References	9

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 $i = 1, \dots, N$). We first assume that the shape parameter k is known.

We start by deriving the maximum likelihood estimator of $\theta := s$ 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 $V_i \sim \Gamma(k, s)$:

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

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

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

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

In order to alleviate notation, we compute instead the log-likelihood, which is generally easier to work with¹:

$$(2.4) \quad \ell(v_1, \dots, v_N; k, \theta) := \ln \mathcal{L}(v_1, \dots, v_N; k, \theta)$$

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

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

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

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

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

$$(2.9) \quad \iff \hat{\theta} = \frac{\sum_{i=1}^N v_i}{kN} = \frac{\bar{v}}{k}.$$

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

$$(2.10) \quad \hat{\Theta} = \frac{\sum_{i=1}^N V_i}{kN} = \frac{\bar{V}}{k}.$$

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 $\hat{\Theta} = g(Z)$ of θ is *unbiased* if

$$(3.1) \quad m_{\hat{\Theta};\theta} := \mathbb{E}[g(Z); \theta] = \theta, \quad \text{for all } \theta,$$

where

$$(3.2) \quad \mathbb{E}[g(Z); \theta] := \int_{\text{dom } Z} g(Z) f_Z(z; \theta) \, dz.$$

¹This 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,*

$$(3.3) \quad \lim_{N \rightarrow +\infty} \mathbb{E}[g(V_1, \dots, V_n); \theta] = \theta.$$

Proof. We wish to prove that

$$(3.4) \quad \lim_{N \rightarrow +\infty} \mathbb{E} \left[\frac{\bar{V}}{k} \right] = \theta.$$

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

$$(3.5) \quad \mathbb{E} \left[\frac{\bar{V}}{k} \right] = \frac{\mathbb{E} \left[\frac{1}{N} \sum_{i=1}^N V_i \right]}{k} = \frac{\frac{1}{N} \sum_{i=1}^N \mathbb{E}[V_i]}{k} = \frac{\frac{1}{N} N k \theta}{k} = \theta.$$

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

3.2. Efficiency.

THEOREM 3.2 (Cramr–Rao inequality). *If $Z = (Z_1, \dots, Z_N)^T$ 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:*

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

then the covariance of any unbiased estimator $\hat{\Theta}$ satisfies the Cramr–Rao inequality

$$(3.7) \quad \text{cov } \hat{\Theta} \geq \mathcal{I}^{-1}(\theta),$$

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

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

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

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

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

Proof. \square

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. \square

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. \square

4. Joint maximum likelihood estimation

We now consider $V_i \sim \Gamma(k, s)$ (for $i = 1, \dots, N$) 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

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

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

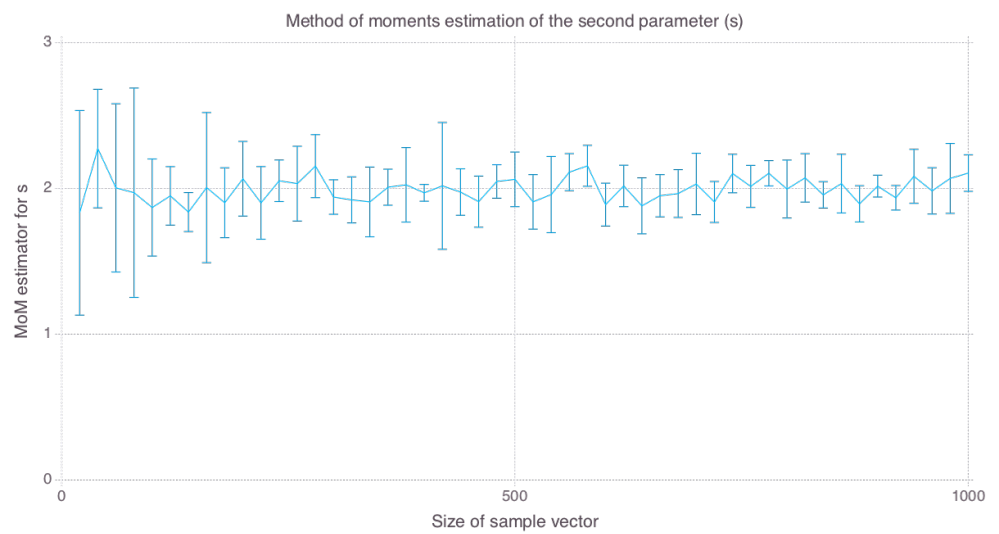
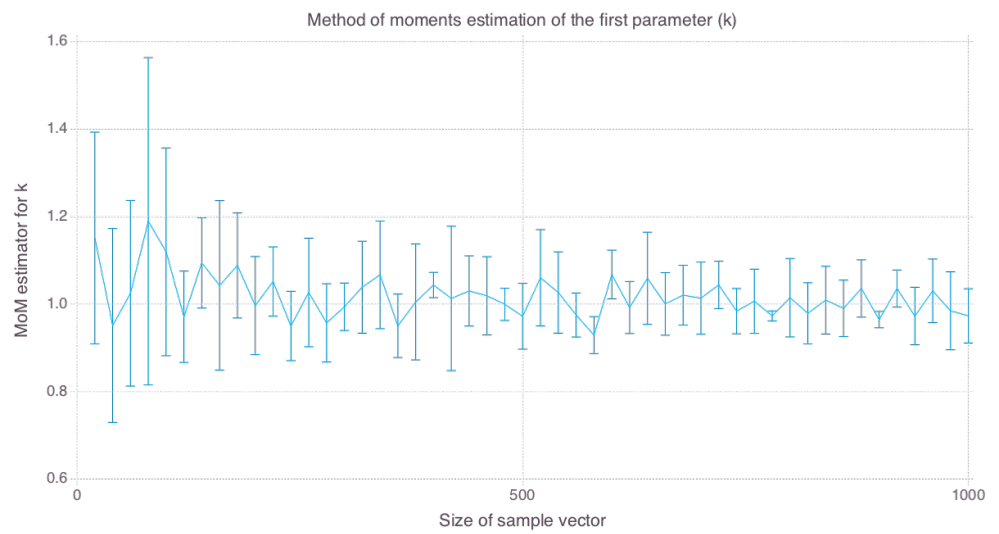
$$\begin{aligned}
 (4.2) \quad \left. \frac{\partial \ell(v_1, \dots, v_N; \theta)}{\partial k} \right|_{k=\hat{k}} &= \sum_i \ln(V_i) - N - N \ln\left(\sum_i V_i\right) + N \ln(k) + \frac{Nk}{k} + N \ln(N) - N \frac{\Gamma(k)}{\Gamma(k)} \phi^{(0)}(k) \\
 &= \sum_i \ln(V_i) - N \ln\left(\sum_i V_i\right)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

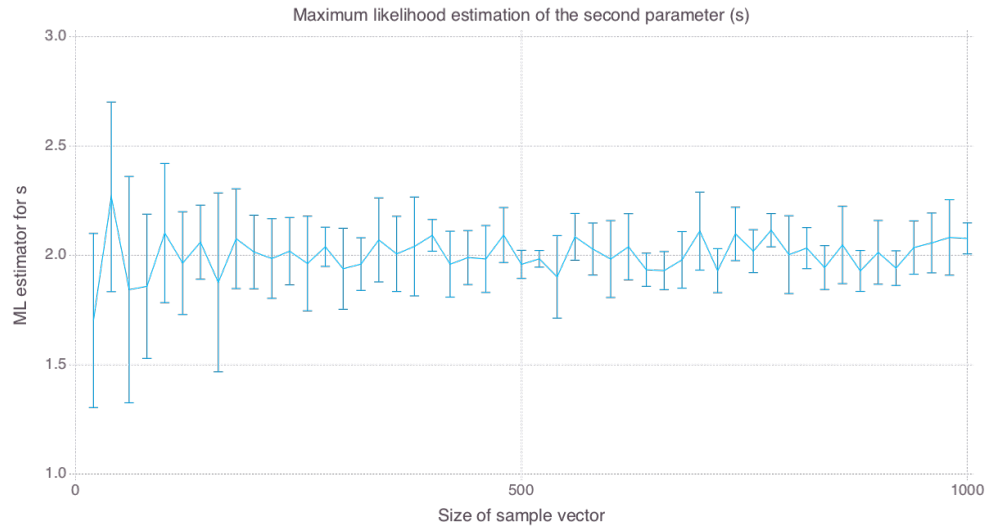
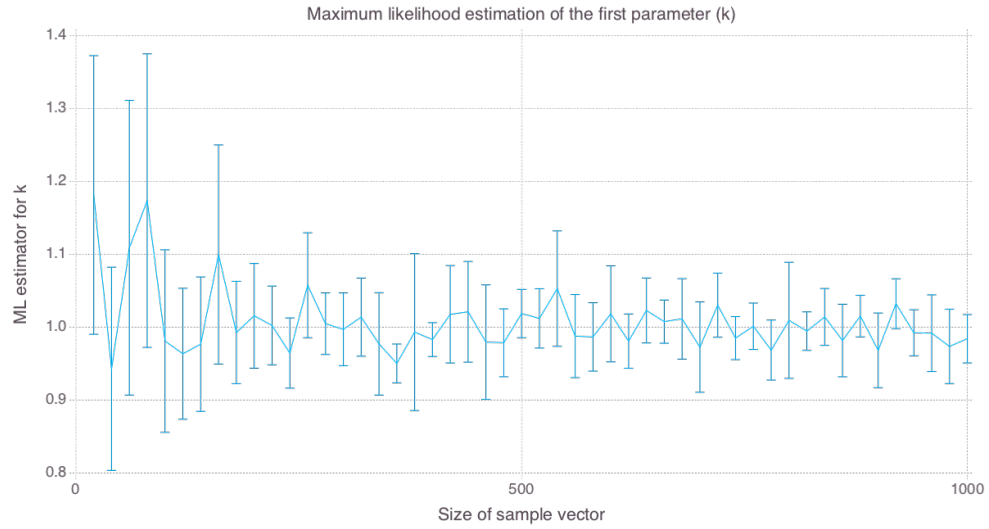
And looking for the root of this derivative, we have

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

We can't find an analytical solution to this equation, but we can approximate it by 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

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

Therefore, before calculating the entries of the information matrix, we have to compute the partial derivatives before taking the expectation of it. Two examples of derivation are given in equations 6.4 and 6.5. The procedure is similar for the other entries and the calculations can be found in Appendix A.1

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

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

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

$$(6.5) \qquad \frac{\partial^2 h(z; \theta)}{\partial k^2} = -\frac{d\phi^{(0)}(k)}{dk} = -\frac{d^2\Gamma(k)}{dk^2} = -\phi^{(1)}(k).$$

Then, we take the expectation of these computed values.

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

Finally, the Fisher information matrix is

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

Then we have

$$\mathcal{I}^{-1}(\theta) = \begin{pmatrix} -\phi^{(1)}(k)s^2 & -s \\ -s & 0 \end{pmatrix}$$

7. Numerical proof

Appendix A. Analytical derivations for d

A.1. *Derivation of the Fisher information matrix coefficients.*

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

$$(A.2) \quad \mathcal{I}_{10} = \mathcal{I}_{01} = \frac{1}{s}.$$

$$\begin{aligned}
 \mathcal{I}_{11} &= -\mathbb{E} \left\{ \frac{\partial^2 h(z; \theta)}{\partial k^2} \right\} \\
 (A.3) \quad &= -\mathbb{E} \left\{ \phi^{(1)}(k) \right\} \\
 &= -\phi^{(1)}(k).
 \end{aligned}$$

References

UNIVERSIT CATHOLIQUE DE LOUVAIN, OTTIGNIES-LOUVAIN-LA-NEUVE, BELGIUM
E-mail: navarre.louis@student.uclouvain.be

UNIVERSIT CATHOLIQUE DE LOUVAIN, OTTIGNIES-LOUVAIN-LA-NEUVE, BELGIUM
E-mail: gilles.peiffer@student.uclouvain.be