Consistent least squares estimation in population-size-dependent branching processes

Carmen Minuesa (University of Extremadura)

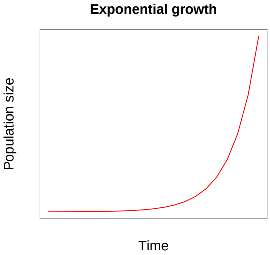
joint work with

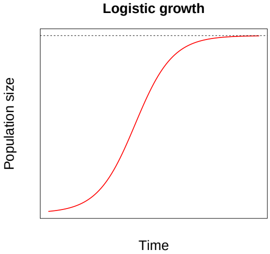
Peter Braunsteins (UNSW Sydney) &
Sophie Hautphenne (University of Melbourne)

Biological motivation

Logistic growth

- The standard assumption of independence between individuals in a population leads to exponential growth.
- This assumption is often not realistic: an individual's lifetime and reproductive parameters usually depend on the availability of a number of resources, such as food, habitat, and breeding opportunities.
- A population can grow only until it reaches the maximum population size a particular habitat can support, named the carrying capacity.





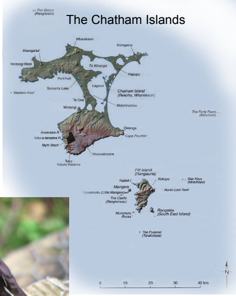
Reindeers at Saint Matthew Island

- In 1944, 29 reindeer were introduced to St. Matthew Island by the United States Coast Guard.
- The Coast Guard abandoned the island a few years later, leaving the reindeer.
- The reindeer population rose to about 6,000 by 1963.
- In the next two years, the number declined to 42 animals (41 females and one male).
- By the 1980s, the reindeer population had died out.

Fig: Source: `https://www.adn.com/features/article/what-wiped-out-st-matthew-islands-reindeer/2010/01/17/`

The black robin population

The endangered Chatham Island black robin population (New Zealand) was saved from the brink of extinction in the early 1980s.



The black robin population

- The current growth of the population does not appear to be exponential.
- The population has not yet reached the carrying capacity of the island.
- A low estimated value of the carrying capacity would highlight the need to find further appropriate habitat.

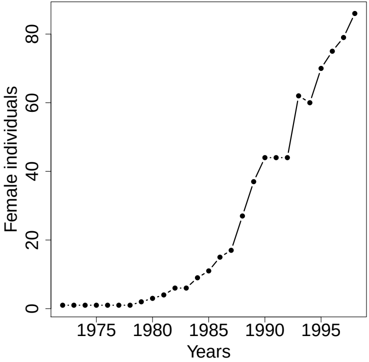


Fig: Number of adult females between 1972 and 1998.

Aim

To estimate the carrying capacity of the black robin population.

The black robin population

- The current growth of the population does not appear to be exponential.
- The population has not yet reached the carrying capacity of the island.
- A low estimated value of the carrying capacity would highlight the need to find further appropriate habitat.

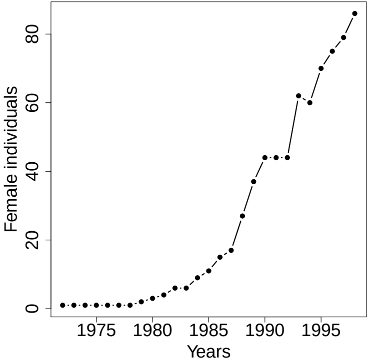


Fig: Number of adult females between 1972 and 1998.

Aim

To estimate the carrying capacity of the black robin population.

Outline

Population-size-dependent branching processes (PSDBPs)

- $\xi(z)$: the offspring distribution at population size z, $\quad z \geq 1$
- Z_n : the population size at generation n,

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_{n,i}(Z_n), \qquad n \geq 0$$

where conditionally on Z_n, the random variables $\xi_{n,i}(Z_n)$ are i.i.d.

- $m(z) := \mathbb{E}[\xi(z)]$ the offspring mean at population size z
- $\sigma^2(z) := \mathrm{Var}[\xi(z)]$ the offspring variance at population size z
- The conditional mean and variance of the process are given by

$$\mathbb{E}[Z_n \mid Z_{n-1}] = Z_{n-1}\, m(Z_{n-1}), \quad \mathrm{Var}[Z_n \mid Z_{n-1}] = Z_{n-1}\, \sigma^2(Z_{n-1})$$

Mathematical carrying capacity

The carrying capacity K of a population in a certain environment is defined such that

$$m(z) > 1, \quad \text{if } z < K, \quad \text{and} \quad m(z) < 1, \quad \text{if } z > K.$$

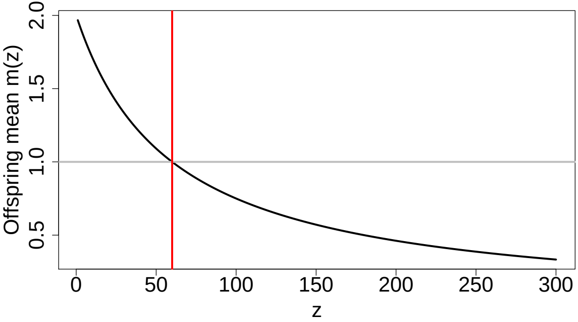


Fig: Mean offspring function $m(z)$ of a PSDBP with carrying capacity $K = 60$.

Almost sure extinction!

Carrying capacity + random nature of the process = a.s. extinction.

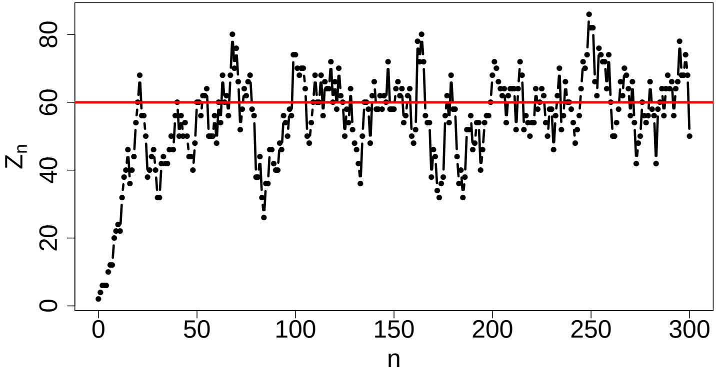


Fig: Portion of a trajectory of a PSDBP with carrying capacity $K = 60$.

A PSDBP model for the black robin population

Between generations n and $n+1$, if there are z female birds,

- A female bird makes a successful attempt to reproduce with probability $r := r(z, K, v)$, where

$$r(z,K,v) = \frac{v\,K}{(\mu-1)z+K} \qquad \text{(Beverton-Holt model).}$$

$$r(z,K,v) = v\left(\frac{1}{\mu}\right)^{z/K} \qquad \text{(Ricker model).}$$

  with $\mu = 5pv/d$.

- If reproduction is successful, the mother produces daughters according to a binomial distribution with $n = 5$ and $p = 0.1988$.

- The mother survives to the next generation with probability $1 - d = 0.6861$.

Outline

Parametric estimation in PSDBPs

Given that our goal is to estimate the carrying capacity, we will need a parametric PSDBP model.

We assume that the offspring distribution belongs to some parametric family, that is,

$$p_k(z) := \mathbb{P}[\xi(z) = k] = p_k(z, \boldsymbol{\theta}_0),$$

for some $\boldsymbol{\theta}_0 \in int(\Theta) \subseteq \mathbb{R}^b$. Then, $m(z) = m(z, \boldsymbol{\theta}_0)$.

In the black robin population example the parameter is $\boldsymbol{\theta}_0 = (K, v)$.

We aim at finding C-consistent estimators for θ_0 based on the observation of the population sizes $Z_0, Z_1, Z_2, \ldots, Z_n$.

Parametric estimation in PSDBPs

Given that our goal is to estimate the carrying capacity, we will need a parametric PSDBP model.

We assume that the offspring distribution belongs to some parametric family, that is,

$$p_k(z) := \mathbb{P}[\xi(z) = k] = p_k(z, \boldsymbol{\theta}_0),$$

for some $\boldsymbol{\theta}_0 \in int(\Theta) \subseteq \mathbb{R}^b$. Then, $m(z) = m(z, \boldsymbol{\theta}_0)$.

In the black robin population example the parameter is $\boldsymbol{\theta}_0 = (K, v)$.

We aim at finding C-consistent estimators for $\boldsymbol{\theta}_0$ based on the observation of the population sizes $Z_0, Z_1, Z_2, \ldots, Z_n$.

A first approach: MLE of the offspring mean

First aim: to find a *good* estimator for the offspring means $m(z)$ based on the observation of the population sizes $Z_0, Z_1, Z_2, \ldots, Z_n$ in a general framework.

- MLE for the offspring mean $m(z)$ at population size z, based on the observation of the population sizes $Z_0, Z_1, \ldots, Z_n$:

$$\hat{m}_n(z) := \frac{\sum_{i=1}^{n} Z_i \mathbb{1}_{\{Z_{i-1}=z\}}}{z \sum_{i=1}^{n} \mathbb{1}_{\{Z_{i-1}=z\}}}.$$

A first approach: MLE of the offspring mean

First aim: to find a *good* estimator for the offspring means $m(z)$ based on the observation of the population sizes $Z_0, Z_1, Z_2, \ldots, Z_n$ in a general framework.

- MLE for the offspring mean $m(z)$ at population size z, based on the observation of the population sizes $Z_0, Z_1, \ldots, Z_n$:

$$\hat{m}_n(z) := \frac{\sum_{i=1}^{n} Z_i \, \mathbb{1}_{\{Z_{i-1}=z\}}}{z \sum_{i=1}^{n} \mathbb{1}_{\{Z_{i-1}=z\}}}.$$

A first approach: MLE of the offspring mean

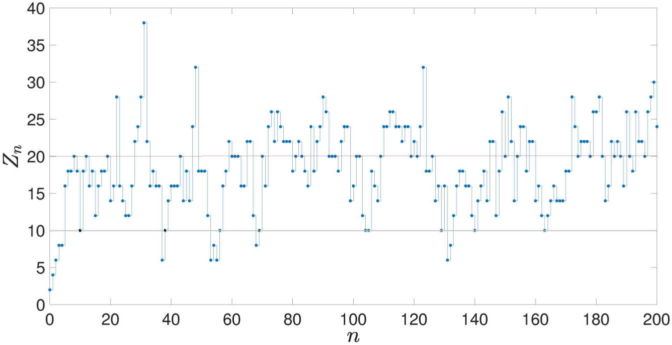


Fig: Example with carrying capacity $K = 20$

$$\hat{m}_n(z) := \frac{\sum_{i=1}^{n} Z_i \, \mathbb{1}_{\{Z_{i-1}=z\}}}{z \sum_{i=1}^{n} \mathbb{1}_{\{Z_{i-1}=z\}}}$$

Example:

$\hat{m}_{200}(10) = (18+14+16+20+10+18+16+14+12)/(10 \cdot 9) = 138/90 = 1.53$

MLE of the offspring mean: asymptotic properties

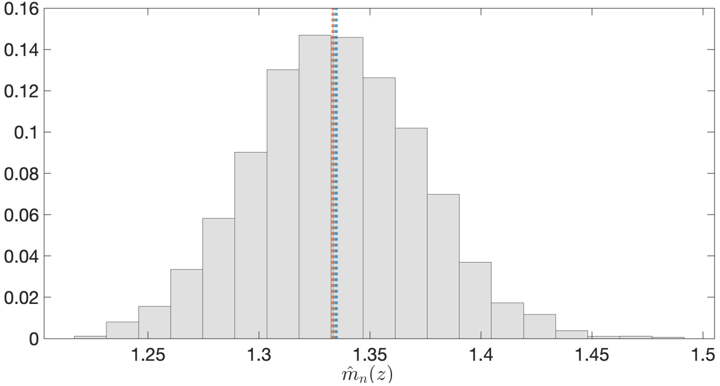


Fig: Histogram of $\hat{m}_n(z)$ for $z = 10$ and $n = 2000$, based on 5000 simulations, $K = 20$.

Real value of $m(z) = 1.3333$; Empirical mean of $\hat{m}_n(z) = 1.3349$

Conditional on $Z_n > 0$, $\hat{m}_n(z) \to m^{\uparrow}(z) = 1.3334 \neq m(z)$ as $n \to \infty$

MLE of the offspring mean: asymptotic properties

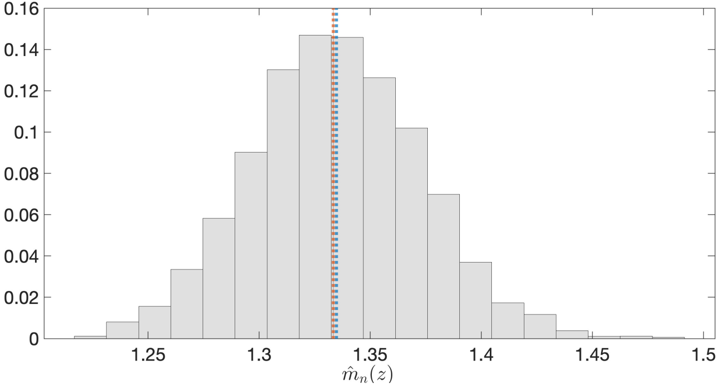


Fig: Histogram of $\hat{m}_n(z)$ for $z = 10$ and $n = 2000$, based on 5000 simulations, $K = 20$.

Real value of $m(z) = 1.3333$; Empirical mean of $\hat{m}_n(z) = 1.3349$

Conditional on $Z_n > 0$, $\hat{m}_n(z) \to m^{\uparrow}(z) = 1.3334 \neq m(z)$ as $n \to \infty$

How do $m(z)$ and $m^{\uparrow}(z)$ differ?

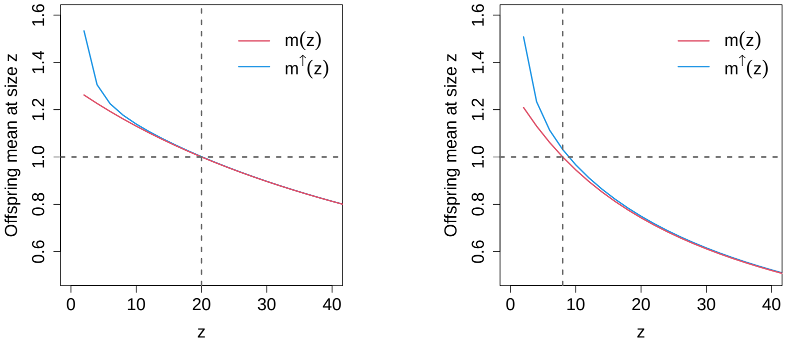


Fig: Comparison between the functions $z \mapsto m(z)$ (red line) and $z \mapsto m^{\uparrow}(z)$ (blue line). Left: $K = 20$. Right: $K = 8$.

Conditioning on $Z_n > 0$ — The Q-process

- Q : the sub-stochastic transition probability matrix of $\{Z_n\}$ restricted to the transient states $\{1, 2, \ldots\}$
- We set the following conditions:
  - (A1) The matrix Q is irreducible
  - (A2) $\limsup_{z\to\infty} m(z) < 1$
  - (A3) For each $\nu \in \mathbb{N}$, $\sup_{z\in\mathbb{N}} E[\xi_{01}(z)^{\nu}] < \infty$.
- Under these conditions,
  - $\mathbb{P}[Z_n \to 0] = 1$ (almost sure extinction)
  - $Q^n \sim \rho^n \boldsymbol{v}\boldsymbol{u}^\top$, where $\rho := \lim_{n\to\infty}(Q^n)_{ij}^{1/n}$, and $\boldsymbol{u}, \boldsymbol{v} > \mathbf{0}$ such that

$$\boldsymbol{u}^\top Q = \rho \boldsymbol{u}^\top, \quad Q\boldsymbol{v} = \rho\boldsymbol{v}, \quad \boldsymbol{u}^\top\mathbf{1} = 1, \quad \text{and} \quad \boldsymbol{u}^\top\boldsymbol{v} = 1.$$

Conditioning on $Z_n > 0$ — The Q-process

- For n fixed: the process $\{Z_\ell\}_{0\leq\ell\leq n}$ conditioned on $Z_n > 0$ is a time-inhomogeneous Markov chain:

$$\begin{aligned}\mathbb{P}[Z_\ell^{(n)} = j \mid Z_{\ell-1}^{(n)} = i] &:= \mathbb{P}[Z_\ell = j \mid Z_{\ell-1} = i,\ Z_n > 0] \\ &= Q_{ij}\,\frac{\boldsymbol{e}_j^\top Q^{n-\ell}\mathbf{1}}{\boldsymbol{e}_i^\top Q^{n-\ell+1}\mathbf{1}}.\end{aligned}$$

- As $n \to \infty$:

$$\begin{aligned}\mathbb{P}[Z_\ell^{\uparrow} = j \mid Z_{\ell-1}^{\uparrow} = i] &:= \lim_{n\to\infty} \mathbb{P}[Z_\ell = j \mid Z_{\ell-1} = i,\ Z_n > 0] \\ &= \lim_{n\to\infty} Q_{ij}\,\frac{\boldsymbol{e}_j^\top \rho^{n-\ell}\boldsymbol{v}}{\boldsymbol{e}_i^\top \rho^{n-\ell+1}\boldsymbol{v}} \\ &= Q_{ij}\,\frac{v_j}{\rho v_i}.\end{aligned}$$

$\{Z_\ell^{\uparrow}\}_{\ell\geq 0}$ is a positive recurrent time-homogeneous Markov chain called the Q-process

Conditioning on $Z_n > 0$ — The Q-process

- For n fixed: the process $\{Z_\ell\}_{0\leq\ell\leq n}$ conditioned on $Z_n > 0$ is a time-inhomogeneous Markov chain:

$$\mathbb{P}[Z_\ell^{(n)} = j \mid Z_{\ell-1}^{(n)} = i] \;:=\; \mathbb{P}[Z_\ell = j \mid Z_{\ell-1} = i,\; Z_n > 0] \;=\; Q_{ij}\,\frac{\mathbf{e}_j^\top Q^{n-\ell}\mathbf{1}}{\mathbf{e}_i^\top Q^{n-\ell+1}\mathbf{1}}.$$

- As $n \to \infty$:

$$\begin{aligned}\mathbb{P}[Z_\ell^{\uparrow} = j \mid Z_{\ell-1}^{\uparrow} = i] \;&:=\; \lim_{n\to\infty}\mathbb{P}[Z_\ell = j \mid Z_{\ell-1} = i,\; Z_n > 0]\\ &=\; \lim_{n\to\infty} Q_{ij}\,\frac{\mathbf{e}_j^\top \rho^{n-\ell}\,\boldsymbol{v}}{\mathbf{e}_i^\top \rho^{n-\ell+1}\,\boldsymbol{v}}\\ &=\; Q_{ij}\,\frac{v_j}{\rho v_i}.\end{aligned}$$

$\{Z_\ell^{\uparrow}\}_{\ell\geq 0}$ is a positive recurrent time-homogeneous Markov chain called the Q-process

The Q-process and 'Q-consistency' of the MLE

$m(z) := \mathbb{E}[\xi(z)]$ the mean offspring at population size z

$$\hat{m}_n(z) := \frac{\sum_{i=1}^{n} Z_i \, \mathbb{1}_{\{Z_{i-1}=z\}}}{z \sum_{i=1}^{n} \mathbb{1}_{\{Z_{i-1}=z\}}}$$

- In $\{Z_n\}$: $\quad m(z) = z^{-1} \sum_{j \geq 1} j \, Q_{zj}$
- In $\{Z_n^{\uparrow}\}$: $\quad m^{\uparrow}(z) = z^{-1} \sum_{j \geq 1} j \, Q_{zj}^{\uparrow} \qquad$ with $Q_{ij}^{\uparrow} := Q_{ij} \frac{v_j}{\rho v_i}$

Theorem (Braunsteins, Hautphenne, M., 2022a)

Under (A1)–(A3), for any $z \in \mathbb{N}$, initial state i, and $\varepsilon > 0$, $\hat{m}_n(z)$ satisfies

$$\lim_{n \to \infty} \mathbb{P}_i[|\hat{m}_n(z) - m^{\uparrow}(z)| > \varepsilon \,|\, Z_n > 0] = 0 \qquad \text{'Q-consistency'}$$

The Q-process and 'Q-consistency' of the MLE

$m(z) := \mathbb{E}[\xi(z)]$ the mean offspring at population size z

$$\hat{m}_n(z) := \frac{\sum_{i=1}^{n} Z_i \, \mathbb{1}_{\{Z_{i-1}=z\}}}{z \sum_{i=1}^{n} \mathbb{1}_{\{Z_{i-1}=z\}}}$$

- In $\{Z_n\}$: $m(z) = z^{-1} \sum_{j \geq 1} j \, Q_{zj}$
- In $\{Z_n^{\uparrow}\}$: $m^{\uparrow}(z) = z^{-1} \sum_{j \geq 1} j \, Q_{zj}^{\uparrow}$ with $Q_{ij}^{\uparrow} := Q_{ij} \frac{v_j}{\rho v_i}$

Theorem (Braunsteins, Hautphenne, M., 2022a)

Under (A1)–(A3), for any $z \in \mathbb{N}$, initial state i, and $\varepsilon > 0$, $\hat{m}_n(z)$ satisfies

$$\lim_{n \to \infty} \mathbb{P}_i[|\hat{m}_n(z) - m^{\uparrow}(z)| > \varepsilon \,|\, Z_n > 0] = 0 \qquad \text{'}Q\text{-consistency'}$$

Asymptotic normality of the MLE

- $(u_i v_i)_{i\geq 1}$: stationary distribution of $\{Z_n^{\uparrow}\}$
- $\sigma^{2\uparrow}(z) = \dfrac{\sum_{k=1}^{\infty} k^2 Q_{zk}^{\uparrow}}{z^2} - (m^{\uparrow}(z))^2$

Theorem (Braunsteins, Hautphenne, M., 2022a)

Under (A1)–(A3), *for any* $z \in \mathbb{N}$, *initial state* i, *and* $x \in \mathbb{R}$, $\hat{m}_n(z)$ *satisfies*

$$\lim_{n\to\infty} \mathbb{P}_i[\{n/\gamma(z)\}^{1/2}\left(\hat{m}_n(z) - m^{\uparrow}(z)\right) \leq x \,|\, Z_n > 0] = \Phi(x),$$

where $\Phi(x)$ *is the standard normal distribution, and*

$$\gamma(z) := \frac{\sigma^{2\uparrow}(z)}{u_z v_z}.$$

Proof approach: coupling techniques and martingale CLT.

Asymptotic normality of the MLE

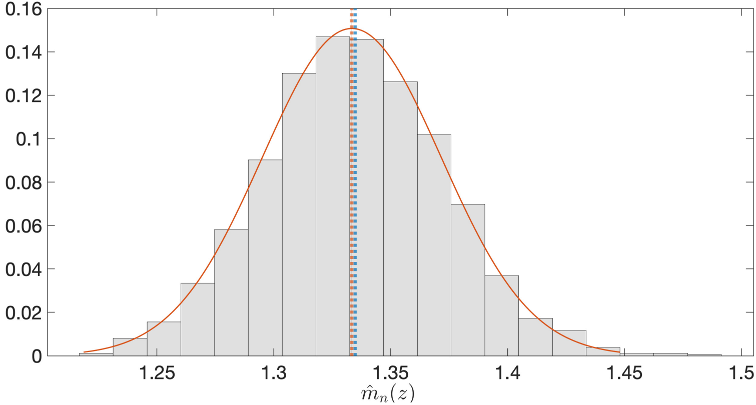


Fig: Histogram of $\hat{m}_n(z)$ for $z = 10$ and $n = 2000$, based on 5000 simulations, $K = 20$.

'Q-consistency' versus C-consistency

- 'Q-consistency': for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}[|\hat{m}_n(z) - m^{\uparrow}(z)| > \varepsilon \mid Z_n > 0] = 0$$

- C-consistency: for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}[|\tilde{m}_n(z) - m(z)| > \varepsilon \mid Z_n > 0] = 0$$

The estimator $\hat{m}_n(z)$ is Q-consistent but not C-consistent

- In our PSDBP, $m^{\uparrow}(z) \approx m(z)$ because $\{Z_\ell^{\uparrow}\} \approx \{Z_\ell\}$.

- The properties of the estimator $\hat{m}_n(z)$ are the key to obtain C-consistent estimators for the parameter θ_0.

'Q-consistency' versus C-consistency

- 'Q-consistency': for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}[|\hat{m}_n(z) - m^{\uparrow}(z)| > \varepsilon \,|\, Z_n > 0] = 0$$

- C-consistency: for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}[|\tilde{m}_n(z) - m(z)| > \varepsilon \,|\, Z_n > 0] = 0$$

The estimator $\hat{m}_n(z)$ is Q-consistent but not C-consistent

- In our PSDBP, $m^{\uparrow}(z) \approx m(z)$ because $\{Z_\ell^{\uparrow}\} \approx \{Z_\ell\}$.
- The properties of the estimator $\hat{m}_n(z)$ are the key to obtain C-consistent estimators for the parameter $\boldsymbol{\theta}_0$.

C-consistent estimators for $\boldsymbol{\theta}_0$

We propose the following weighted least squares estimator of the

$$\hat{\boldsymbol{\theta}}_n = \arg\min_{\boldsymbol{\theta}\in\Theta} \sum_{z=1}^{\infty} \hat{w}_n(z) \left[\hat{m}_n(z) - m^{\uparrow}(z,\boldsymbol{\theta})\right]^2,$$

where the weights $\{\hat{w}_n(z)\}_{z\geq 1}$ are computed from the observations $Z_0, Z_1, \ldots, Z_n$, and are assumed to form an empirical distribution such that for any $z, i \geq 1$ and $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}_i[|\hat{w}_n(z) - w_z| > \varepsilon \mid Z_n > 0] = 0,$$

for some limiting distribution $\{w_z := w_z(\boldsymbol{\theta}_0)\}_{z\geq 1}$.

C-consistent estimators for $\boldsymbol{\theta}_0$

Under additional regularity assumptions on the parametric family

$$\mathcal{F}_\Theta = \{m^{\uparrow}(z,\boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta, z \in \mathbb{N}\},$$

we proved that $\hat{\boldsymbol{\theta}}_n$ is C-consistent and asymptotically normal.

Theorem (Braunsteins, Hautphenne, M., 2022b)

For any initial state i, and $\varepsilon > 0$, $\hat{\boldsymbol{\theta}}_n$ satisfies

$$\lim_{n\to\infty} \mathbb{P}_i\left[\left\|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\right\| > \epsilon | Z_n > 0\right] = 0 \qquad C\text{-consistency.}$$

C-consistent estimators for $\boldsymbol{\theta}_0$

Theorem (Braunsteins, Hautphenne, M., 2022b)

For any initial state i, and for any $\boldsymbol{x} = (x_1, \ldots, x_b) \in \mathbb{R}^b$,

$$\lim_{n\to\infty} \mathbb{P}_i \left[\sqrt{n} \left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0 \right) \in (-\infty, x_1] \times \ldots \times (-\infty, x_b] \,|Z_n > 0 \right] = \Psi_{\beta(\boldsymbol{\theta}_0)}(\boldsymbol{x}),$$

where $\Psi_{\beta(\boldsymbol{\theta}_0)}(\cdot)$ is the distribution function of a b-dimensional normal r.v. with mean vector $\mathbf{0}$ and positive semi-definite covariance matrix

$$\beta(\boldsymbol{\theta}_0) := \eta(\boldsymbol{\theta}_0)^{-1}\zeta(\boldsymbol{\theta}_0)\eta(\boldsymbol{\theta}_0)^{-1},$$

where $\eta(\boldsymbol{\theta}_0)$ and $\zeta(\boldsymbol{\theta}_0)$ are the b-dimensional matrices given by

$$\begin{aligned}
\eta(\boldsymbol{\theta}_0) &= 2\sum_{z=1}^{\infty} w_z(\boldsymbol{\theta}_0)\nabla m^{\uparrow}(z,\boldsymbol{\theta}_0)\nabla m^{\uparrow}(z,\boldsymbol{\theta}_0)^{\top}, \\
\zeta(\boldsymbol{\theta}_0) &= 4\sum_{z=1}^{\infty} w_z(\boldsymbol{\theta}_0)^2\gamma(z,\boldsymbol{\theta}_0)\nabla m^{\uparrow}(z,\boldsymbol{\theta}_0)\nabla m^{\uparrow}(z,\boldsymbol{\theta}_0)^{\top},
\end{aligned}$$

Weight functions

We consider different weight functions:

(1) The proportion of generations with population size z:

$$\hat{w}_n^{(1)}(z) = \frac{\sum_{i=0}^{n-1} \mathbb{1}_{\{Z_i=z\}}}{n}.$$

Lemma (Braunsteins, Hautphenne, M., 2022b)

For any $z \geq 1$ and $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}_i\left[\left|\hat{w}_n^{(1)}(z) - w_z^{(1)}\right| > \epsilon | Z_n > 0\right] = 0,$$

where

$$w_z^{(1)} = u_z v_z.$$

Conditionally on $Z_n > 0$, the weights $\hat{w}_n^{(1)}(z)$ converge to the stationary distribution of the Q-process.

Weight functions

(2) The proportion of individuals who are alive when the population size is z:

$$\hat{w}_n^{(2)}(z) = \frac{z \sum_{i=0}^{n-1} \mathbb{1}_{\{Z_i = z\}}}{\sum_{i=0}^{n-1} Z_i}.$$

Lemma (Braunsteins, Hautphenne, M., 2022b)

For any $z \geq 1$ and $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}_i \left[\left| \hat{w}_n^{(2)}(z) - w_z^{(2)} \right| > \epsilon \middle| Z_n > 0 \right] = 0,$$

where

$$w_z^{(2)} = \frac{z\, u_z v_z}{\sum_{k=1}^{\infty} k u_k v_k}.$$

Conditionally on $Z_n > 0$, the weights $\hat{w}_n^{(2)}(z)$ converge to the size-biased distribution of the stationary distribution of the Q-process.

Comparison with the classical least squares estimator

- Classical least squares estimators:

$$\widetilde{\boldsymbol{\theta}}_n^* := \arg\min_{\boldsymbol{\theta}\in\Theta} \sum_{k=1}^{n} w_k \left\{Z_k - Z_{k-1}\, m(Z_{k-1}, \boldsymbol{\theta})\right\}^2,$$

where $\{w_k\}$ is an appropriately chosen weighting function.

Proposition (Braunsteins, Hautphenne, M., 2022b)

The least squares estimator $\widehat{\boldsymbol{\theta}}_n$ with weight $\{w_z^{(1)}\}$ or $\{w_z^{(2)}\}$ is equal to the previous estimator modified such that $m(\cdot)$ is replaced by $m^{\uparrow}(\cdot)$,

$$\widehat{\boldsymbol{\theta}}_n^* := \arg\min_{\boldsymbol{\theta}\in\Theta} \sum_{k=1}^{n} w_k \left\{Z_k - Z_{k-1}\, m^{\uparrow}(Z_{k-1}, \boldsymbol{\theta})\right\}^2,$$

with respective weight $w_k^{(1)} = Z_{k-1}^{-2}$ or $w_k^{(2)} = Z_{k-1}^{-1}$.

Comparison with the classical least squares estimator

- Classical least squares estimators:

$$\widetilde{\boldsymbol{\theta}}_n^* := \arg\min_{\boldsymbol{\theta}\in\Theta} \sum_{k=1}^{n} w_k \left\{Z_k - Z_{k-1}\, m(Z_{k-1}, \boldsymbol{\theta})\right\}^2,$$

where $\{w_k\}$ is an appropriately chosen weighting function.

Proposition (Braunsteins, Hautphenne, M., 2022b)

The least squares estimator $\widehat{\boldsymbol{\theta}}_n$ with weight $\{w_z^{(1)}\}$ or $\{w_z^{(2)}\}$ is equal to the previous estimator modified such that $m(\cdot)$ is replaced by $m^{\uparrow}(\cdot)$,

$$\widehat{\boldsymbol{\theta}}_n^* := \arg\min_{\boldsymbol{\theta}\in\Theta} \sum_{k=1}^{n} w_k \left\{Z_k - Z_{k-1}\, m^{\uparrow}(Z_{k-1}, \boldsymbol{\theta})\right\}^2,$$

with respective weight $w_k^{(1)} = Z_{k-1}^{-2}$ or $w_k^{(2)} = Z_{k-1}^{-1}$.

Outline

Simulated example: estimation in the quasi-stationary phase

We consider the PSDBP with a carrying capacity and binary fission reproduction given by a modified BH model:

$$p_2(z,K) = \frac{vK}{K+(2v-1)z}, \quad p_0(z,K) = 1 - p_2(z,K), \quad z \in \{2i : i \in \mathbb{N}\}.$$

The offspring parameter is $\boldsymbol{\theta} = (K, v) \in \Theta = (0,\infty) \times (1/2, 1]$.

We fixed $\boldsymbol{\theta}_0 = (K_0, v_0) = (50, 0.70)$.

We simulated 2000 non-extinct trajectories, and we are interested in estimating the parameters based on the observation of $Z_0, Z_1, \ldots, Z_{2000}$.

Estimation in the quasi-stationary phase

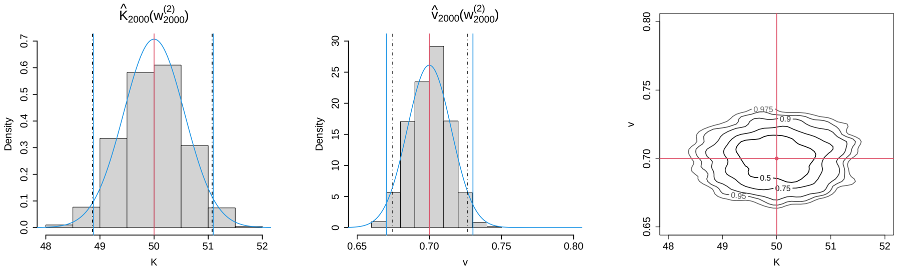


Fig: Left: marginal distribution of K, together with the empirical and theoretical marginal 95% confidence intervals. Centre: marginal distribution of the estimator of v, together with the empirical and theoretical marginal 95% confidence intervals. Right: confidence regions for levels $50\%, 75\%, 90\%, 95\%, 97.5\%$.

Simulated example: estimation in the growth phase

We consider the PSDBP with a carrying capacity and binary fission reproduction given by a modified BH model:

$$p_2(z,K)=\frac{vK}{K+(2v-1)z},\quad p_0(z,K)=1-p_2(z,K),\quad z\in\{2i:i\in\mathbb{N}\}.$$

The offspring parameter is $\boldsymbol{\theta}=(K,v)\in\Theta=(0,\infty)\times(1/2,1]$.

We fixed $\boldsymbol{\theta}_0=(K_0,v_0)=(200,0.75)$.

We simulated 1000 non-extinct trajectories, and we are interested in estimating the parameters based on the observation of $Z_0,Z_1,\ldots,Z_{12}$.

Estimation in the growth phase

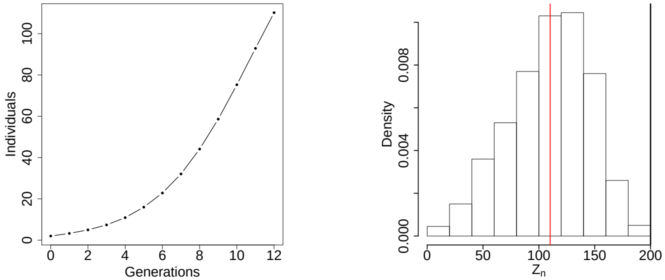


Fig: **Left:** empirical mean of the simulated paths. **Right:** empirical distribution of Z_{12}, together with the empirical mean (red line) and the carrying capacity (grey line).

Simulated example: estimation in the growth phase

$$\hat{\boldsymbol{\theta}}_n = \arg\min_{\boldsymbol{\theta}\in\Theta} \sum_{z=1}^{\infty} \hat{w}_n(z) \left\{ \hat{m}_n(z) - m^{\uparrow}(z, \boldsymbol{\theta}) \right\}^2 \qquad (C\text{-consistent})$$

$$\tilde{\boldsymbol{\theta}}_n^* = \arg\min_{\boldsymbol{\theta}\in\Theta} \sum_{z=1}^{\infty} \hat{w}_n(z) \left\{ \hat{m}_n(z) - m(z, \boldsymbol{\theta}) \right\}^2 \qquad (\text{modified})$$

(1) $\hat{w}_n^{(1)}(z) = \sum_{i=0}^{n-1} \mathbb{1}_{\{Z_i = z\}} / n$

(2) $\hat{w}_n^{(2)}(z) = z \sum_{i=0}^{n-1} \mathbb{1}_{\{Z_i = z\}} / (\sum_{i=0}^{n-1} Z_i)$

Table: Median of the 1000 estimates — $\boldsymbol{\theta}_0 = (K_0, v_0) = (200, 0.75)$

	C-consistent		Modified	
	$\hat{K}_n^*$	$\hat{v}_n^*$	$\tilde{K}_n^*$	$\tilde{v}_n^*$
(1)	177.60010	0.77088	151.31210	0.79670
(2)	191.78720	0.76622	185.79590	0.77025

Carrying capacity of the black robin population

Model	$\hat{K}$	$\hat{v}$
Beverton-Holt	109.61	0.696
Ricker	95.64	0.679

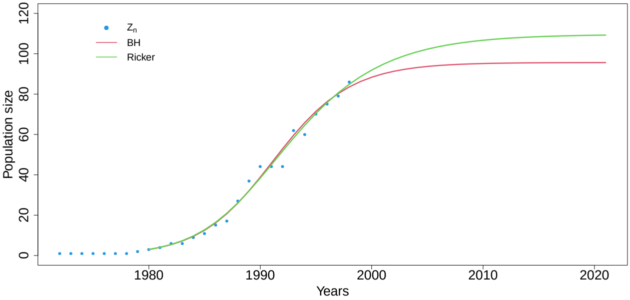


Fig: Number of adult females between 1972 and 1998, together with the mean population size curve of the estimated BH and Ricker models.

Outline

Conclusions

- We obtained the MLEs of the offspring mean function of a PSDBP based on the observation of the population sizes.
- Focussing our attention on the study of PSDBPs whose extinction is certain, we considered the sample of the population sizes and we analysed the asymptotic properties of the MLE $\hat{m}_n(z)$ for a population size z fixed, establishing its Q-consistency and asymptotic normality.
- In a parametric setting, we developed C-consistent estimators for the offspring parameter θ_0 of PSDBPs.
- We applied our results to estimate the carrying capacity of the endangered black robin population in the Chatham Islands.
- The choice of weights and parametric model are still very important.

Conclusions

- We obtained the MLEs of the offspring mean function of a PSDBP based on the observation of the population sizes.
- Focussing our attention on the study of PSDBPs whose extinction is certain, we considered the sample of the population sizes and we analysed the asymptotic properties of the MLE $\hat{m}_n(z)$ for a population size z fixed, establishing its Q-consistency and asymptotic normality.
- In a parametric setting, we developed C-consistent estimators for the offspring parameter θ_0 of PSDBPs.
- We applied our results to estimate the carrying capacity of the endangered black robin population in the Chatham Islands.
- The choice of weights and parametric model are still very important.

Conclusions

- We obtained the MLEs of the offspring mean function of a PSDBP based on the observation of the population sizes.
- Focussing our attention on the study of PSDBPs whose extinction is certain, we considered the sample of the population sizes and we analysed the asymptotic properties of the MLE $\hat{m}_n(z)$ for a population size z fixed, establishing its Q-consistency and asymptotic normality.
- In a parametric setting, we developed C-consistent estimators for the offspring parameter $\boldsymbol{\theta}_0$ of PSDBPs.
- We applied our results to estimate the carrying capacity of the endangered black robin population in the Chatham Islands.
- The choice of weights and parametric model are still very important.

Conclusions

- We obtained the MLEs of the offspring mean function of a PSDBP based on the observation of the population sizes.
- Focussing our attention on the study of PSDBPs whose extinction is certain, we considered the sample of the population sizes and we analysed the asymptotic properties of the MLE $\hat{m}_n(z)$ for a population size z fixed, establishing its Q-consistency and asymptotic normality.
- In a parametric setting, we developed C-consistent estimators for the offspring parameter $\boldsymbol{\theta}_0$ of PSDBPs.
- We applied our results to estimate the carrying capacity of the endangered black robin population in the Chatham Islands.
- The choice of weights and parametric model are still very important.

Conclusions

- We obtained the MLEs of the offspring mean function of a PSDBP based on the observation of the population sizes.
- Focussing our attention on the study of PSDBPs whose extinction is certain, we considered the sample of the population sizes and we analysed the asymptotic properties of the MLE $\hat{m}_n(z)$ for a population size z fixed, establishing its Q-consistency and asymptotic normality.
- In a parametric setting, we developed C-consistent estimators for the offspring parameter $\boldsymbol{\theta}_0$ of PSDBPs.
- We applied our results to estimate the carrying capacity of the endangered black robin population in the Chatham Islands.
- The choice of weights and parametric model are still very important.

References

P. Braunsteins, S. Hautphenne, and C. Minuesa.
Parameter estimation in branching processes with almost sure extinction.
Bernoulli, vol. 28, no. 1, pp. 33-63, 2022a.

P. Braunsteins, S. Hautphenne, and C. Minuesa.
Consistent least squares estimation in population-size-dependent branching processes.
`arXiv:2211.10898`, 2022b.

D. R. Klein.
The introduction, increase, and crash of reindeer on St. Matthew Island.
Journal of Wildlife Management, vol. 32, no. 2, pp. 350-367, 1968.

M. Massaro, A. Chick, E.S. Kennedy, and R. Whitsed.
Post-reintroduction distribution and habitat preferences of a spatially limited island bird species.
Animal Conservation, vol. 21, no. 1, pp. 54-64, 2018.

Thank you for your attention!

Acknowledgements: This work was started while Carmen Minuesa was a visiting postdoctoral researcher in the School of Mathematics and Statistics at The University of Melbourne, and she is grateful for the hospitality and collaboration. She also acknowledges the Australian Research Council Centre of Excellence for Mathematical and Statistical Frontiers for partially supporting her research visit at The University of Melbourne.

Funding: This research is part of the R&D&I project PID2019-108211GB-I00, funded by MCIN/AEI/10.13039/501100011033/.

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA E INNOVACIÓN

Unión Europea

JUNTA DE EXTREMADURA
Consejería de Economía e Infraestructuras

Fondo Europeo de Desarrollo Regional
Una manera de hacer Europa

Our assumptions

(A1) If $m^{\uparrow}(z, \boldsymbol{\theta}_1) = m^{\uparrow}(z, \boldsymbol{\theta}_2)$, for each $z \in \text{supp}(\boldsymbol{w}(\boldsymbol{\theta}_0))$, then $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$

(A2) $M = \sup_{\boldsymbol{\theta}\in\Theta} \sup_{z\in\mathbb{N}} m^{\uparrow}(z, \boldsymbol{\theta}) < \infty$.

(A3) For each $\epsilon > 0$,
$\lim_{n\to\infty} P_i[\sup_{z\in\mathbb{N}} |\hat{m}_n(z) - m^{\uparrow}(z, \boldsymbol{\theta}_0)| > \epsilon | Z_n > 0] = 0$, for any initial state $i \in \mathbb{N}$.

(A4) The function $m^{\uparrow}(z, \boldsymbol{\theta})$ is twice continuously differentiable with respect to $\boldsymbol{\theta}$ for each $z \in \mathbb{N}$. Moreover, for each $\boldsymbol{\theta}' \in \Theta$, there exists a compact set C such that $\boldsymbol{\theta}' \in int(C)$:

1. $M_1^*(\boldsymbol{\theta}') = \sup_{\boldsymbol{\theta}\in C} \sup_{z\in\mathbb{N}} \left\| \nabla m^{\uparrow}(z, \boldsymbol{\theta}) \right\|_{\infty} < \infty$.
2. $M_2^*(\boldsymbol{\theta}') = \sup_{\boldsymbol{\theta}\in C} \sup_{z\in\mathbb{N}} \max_{i,j=1,\dots,d} \left| \frac{\partial^2 m^{\uparrow}(z,\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j} \right| < \infty$.

Ingredient for the proof of C-consistency

We use the following lemma:

Lemma

Let $f : \boldsymbol{\theta} \in \Theta \mapsto [0, \infty)$ be a continuous function with a unique minimum at $\boldsymbol{\theta}_0$ that satisfies $f(\boldsymbol{\theta}_0) = 0$.

Let $\hat{f}_n(\boldsymbol{\theta})$ be a (random) measurable function of the random variables $Z_0, \ldots, Z_n$ for each $\boldsymbol{\theta} \in \Theta$, s.t. for each $\omega \in \Omega$, $\hat{f}_n(\omega, \cdot)$ is a continuous function on Θ.

$$\text{If} \quad \forall \epsilon > 0, \quad \lim_{n\to\infty} P\left[\sup_{\boldsymbol{\theta}\in\Theta} |\hat{f}_n(\boldsymbol{\theta}) - f(\boldsymbol{\theta})| > \epsilon | Z_n > 0\right] = 0, \tag{1}$$

then

1. $\lim_{n\to\infty} P\left[\exists \arg\min_{\boldsymbol{\theta}\in\Theta} \hat{f}_n(\boldsymbol{\theta}) | Z_n > 0\right] = 1.$
2. $\forall \epsilon > 0$, $\lim_{n\to\infty} P\left[\left\| \arg\min_{\boldsymbol{\theta}\in\Theta} \hat{f}_n(\boldsymbol{\theta}) - \boldsymbol{\theta}_0 \right\| > \epsilon | Z_n > 0\right] = 0.$

Idea of the proof of C-consistency

We apply the previous lemma with

$$\hat{f}_n(\boldsymbol{\theta}) = \sum_{z=1}^{\infty} \hat{w}_n(z) \left\{ \hat{m}_n(z) - m^{\uparrow}(z, \boldsymbol{\theta}) \right\}^2$$

and

$$f(\boldsymbol{\theta}) = \sum_{z=1}^{\infty} w_z(\boldsymbol{\theta}_0) \left\{ m^{\uparrow}(z, \boldsymbol{\theta}_0) - m^{\uparrow}(z, \boldsymbol{\theta}) \right\}^2$$

and prove that (1) holds. To that end we introduce

$$\tilde{f}_n(\boldsymbol{\theta}) = \sum_{z=1}^{\infty} w_z(\boldsymbol{\theta}_0) \mathbb{1}_{\{j_n(z)>0\}} \left\{ \hat{m}_n(z) - m^{\uparrow}(z, \boldsymbol{\theta}) \right\}^2, \quad n \in \mathbb{N},\ \boldsymbol{\theta} \in \Theta,$$

and we prove (using our assumptions) that for each $\epsilon > 0$,

1. $\lim_{n\to\infty} P\left[\sup_{\boldsymbol{\theta}\in\Theta} \left|\hat{f}_n(\boldsymbol{\theta}) - \tilde{f}_n(\boldsymbol{\theta})\right| > \epsilon | Z_n > 0\right] = 0.$
2. $\lim_{n\to\infty} P\left[\sup_{\boldsymbol{\theta}\in\Theta} \left|\tilde{f}_n(\boldsymbol{\theta}) - f(\boldsymbol{\theta})\right| > \epsilon | Z_n > 0\right] = 0.$

Idea of the proof of asymptotic normality

We use a Taylor expansion for the function $\nabla \hat{f}_n(\cdot)$ around $\boldsymbol{\theta}_0$ on the set $\{Z_n > 0\}$,

$$\mathbf{0} = \nabla \hat{f}_n(\hat{\boldsymbol{\theta}}_n) = \nabla \hat{f}_n(\boldsymbol{\theta}_0) + \nabla^2 \hat{f}_n(\boldsymbol{\theta}_n)^\top (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0),$$

where $\boldsymbol{\theta}_n$ is a point between $\hat{\boldsymbol{\theta}}_n$ and $\boldsymbol{\theta}_0$, and $\nabla^2 \hat{f}_n(\boldsymbol{\theta}_n)$ is the Jacobian matrix of $\hat{f}_n(\cdot)$ at $\boldsymbol{\theta}_n$. Then,

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) = -\left(\nabla^2 \hat{f}_n(\boldsymbol{\theta}_n)\right)^{-1} \sqrt{n}\, \nabla \hat{f}_n(\boldsymbol{\theta}_0).$$

Idea of the proof of asymptotic normality

We then prove the result in two steps:

1. If $\Psi_{\zeta(\boldsymbol{\theta}_0)}(\cdot)$ is the distribution function of a b-dimensional normal distribution with mean vector $\mathbf{0}$ and covariance matrix $\zeta(\boldsymbol{\theta}_0)$, and $x_1, \ldots, x_b \in \mathbb{R}$, then

$$\lim_{n\to\infty} P\left[-\sqrt{n}\,\frac{\partial \hat{f}_n(\boldsymbol{\theta}_0)}{\partial\theta_j} \leq x_j,\ j=1,\ldots,b \,\middle|\, Z_n > 0\right] = \Psi_{\zeta(\boldsymbol{\theta}_0)}(x_1,\ldots,x_b).$$

2. For each $\epsilon > 0$,

$$\lim_{n\to\infty} P\left[\left|\frac{\partial^2 \hat{f}_n(\boldsymbol{\theta}_n)}{\partial\theta_j\partial\theta_l} - \eta_{jl}(\boldsymbol{\theta}_0)\right| > \epsilon | Z_n > 0\right] = 0, \quad j,l=1,\ldots,b.$$

Estimation in the quasi-stationary phase

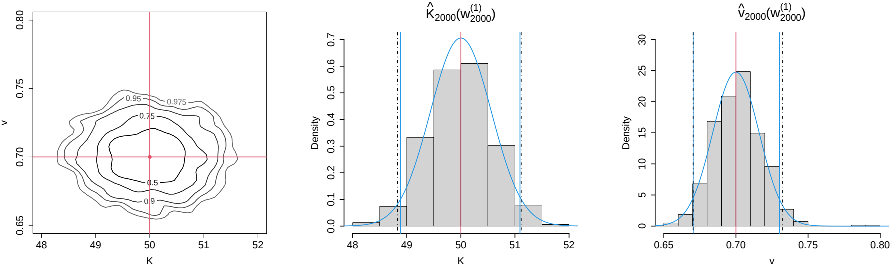


Fig: Left: confidence regions for levels 50%, 75%, 90%, 95%, 97.5%. Centre: marginal distribution of K. Right: marginal distribution of the estimator of v.