Inference in Population-Size-Dependent Branching Processes

Sophie Hautphenne

joint work with Peter Braunsteins and Carmen Minuesa

The University of Melbourne

Seventh Wellington Workshop in Probability and Mathematical Statistics

Galton–Watson Branching processes

- ξ : the offspring distribution
- $m := \mathbb{E}[\xi]$ the mean offspring
- Z_n : the population size at generation n,

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

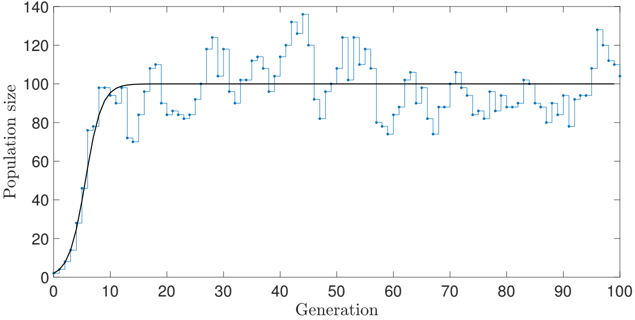
- $\mathbb{E}[Z_n \mid Z_0] = Z_0\, m^n$ (exponential growth)
- The MLE of m based on $Z_0, Z_1, \ldots, Z_n$:

$$\hat{m}_n = \frac{\sum_{i=1}^{n} Z_i}{\sum_{i=1}^{n} Z_{i-1}}$$

- Supercritical case $m > 1$: $\hat{m}_n \to m$ on $\{Z_n \to \infty\}$ (consistency).

In reality...

- Biological populations do not grow exponentially
- Population sizes tend to fluctuate around a carrying capacity
- 99.9% of all species are extinct
- Extinction occurs slowly – often after millions of years



Population-size-dependent branching processes

- $\xi(z)$: the offspring distribution at population size z, $\quad z \geq 1$
- $m(z) := \mathbb{E}[\xi(z)]$ the mean offspring at population size z
- Z_n : the population size at generation n,

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

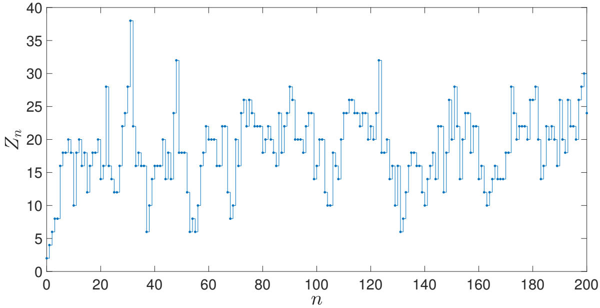
- $\mathbb{E}[Z_n \,|\, Z_{n-1}] = Z_{n-1}\, m(Z_{n-1})$

Example : the Beverton-Holt (B-H) binary splitting model :

- $\xi(z) \sim 2\,\mathrm{Ber}(p(z))$ with $p(z) = \dfrac{K}{K+z}$, where K : carrying capacity
- $m(z) = \dfrac{2K}{K+z}, \quad m(z) \geq 1 \Leftrightarrow z \leq K$

Population-size-dependent branching processes

B-H model with $K = 20$



The MLE of $m(z)$ based on $Z_0, Z_1, \ldots, Z_n$:

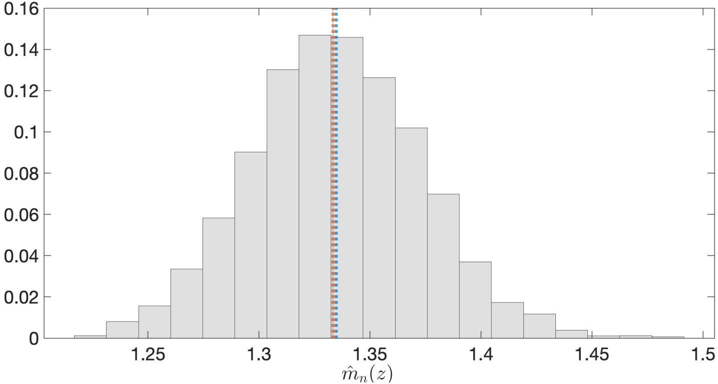
$$\hat{m}_n(z) := \frac{\sum_{i=1}^{n} Z_i \, \mathbb{1}_{\{Z_{i-1}=z\}}}{z \sum_{i=1}^{n} \mathbb{1}_{\{Z_{i-1}=z\}}}$$

Example :

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

Asymptotic properties of the estimator

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



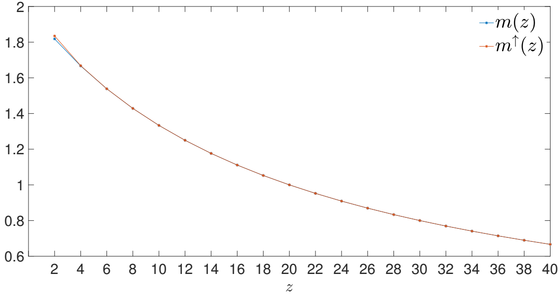
Real value of $m(z) = 1.3333$; Empirical mean of $\hat{m}_n(z) = 1.3349$

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

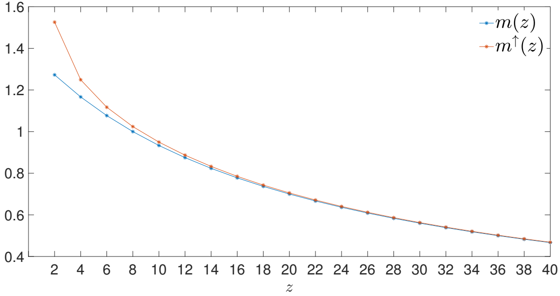
How do $m(z)$ and $m^{\uparrow}(z)$ differ ?

$K = 20, \ v = 1$

(v : efficiency)



$K = 8, \ v = 0.7$



Conditioning on $Z_n > 0$

- Q : the sub-stochastic transition probability matrix of $\{Z_n\}$ restricted to the transient states $\{1, 2, \ldots\}$
- We set the following conditions :

(A1) There exists $z \in \mathbb{N}$ and $n \geq 1$ such that $(Q^n)_{zz} > 0$
(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 $\boldsymbol{u}, \boldsymbol{v} > \boldsymbol{0}$ such that

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

Conditioning on $Z_n > 0$

- For n fixed : the process $\{Z_\ell\}_{0\le\ell\le n}$ 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 $n \to \infty$:

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

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

The Q-process and 'Q-consistency'

- 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, H., Minuesa (2019))

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

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

Asymptotic normality

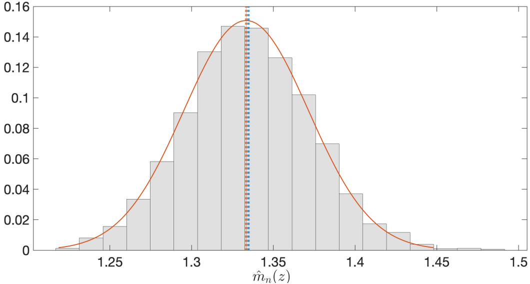
$(u_i v_i)_{i\geq 1}$: stationary distribution of $\{Z_n^{\uparrow}\}$, $\quad \sigma^{2\uparrow}(z) = \frac{\sum_{k=1}^{\infty} k^2 Q_{zk}^{\uparrow}}{z^2} - (m^{\uparrow}(z))^2$

Theorem (Braunsteins, H., Minuesa (2019))

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

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

where $\Phi(x)$ is the standard normal distribution.



Proof approach : coupling and martingale CLT

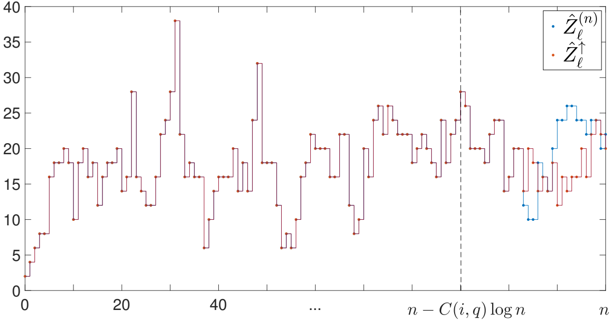
We use a MEXIT coupling $(\{\widehat{Z}_\ell^{(n)}\}, \{\widehat{Z}_\ell^{\uparrow}\})$ of

- $\{Z_\ell^{(n)}\}_{0\le\ell\le n}$: the time-inhomogeneous Markov chain $\{Z_\ell\}_{0\le\ell\le n}$ conditioned on $Z_n > 0$
- $\{Z_\ell^{\uparrow}\}_{0\le\ell\le n}$: the time-homogeneous Markov chain $\{Z_\ell\}_{0\le\ell\le n}$ conditioned on $Z_\infty > 0$

Let $\tau := \min\{\ell \le n : \widehat{Z}_\ell^{(n)} \ne \widehat{Z}_\ell^{\uparrow}\}$ (un-coupling time)

$$\hat{\mathbb{P}}(\tau \le k) = d_{TV}\left(\{Z_\ell^{(n)}\}_{0\le\ell\le k}, \{Z_\ell^{\uparrow}\}_{0\le\ell\le k}\right), \qquad k \le n$$

Proof approach : coupling and martingale CLT



Theorem (Braunsteins, H., Minuesa (2019))

Under Assumptions (A1)–(A3), for any initial state i and any $q > 0$, there exist constants $C(i,q)$ and $N(i,q)$ such that

$$\hat{\mathbb{P}}(\tau \leq n - C(i,q)\log n) \leq \frac{1}{n^q} \qquad \textit{for all } n \geq N(i,q).$$

'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

Thankfully, in our PSDBP, $m^{\uparrow}(z) \approx m(z)$ because $\{Z_\ell^{\uparrow}\} \approx \{Z_\ell\}$

Are there situations where we may prefer C-consistency ?

'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

Thankfully, in our PSDBP, $m^{\uparrow}(z) \approx m(z)$ because $\{Z_\ell^{\uparrow}\} \approx \{Z_\ell\}$

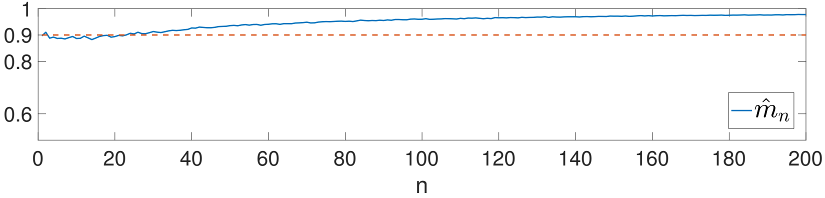
Are there situations where we may prefer C-consistency ?

We often study endangered populations because they are still alive !

Subcritical Galton–Watson branching processes

- Mean offspring $m < 1$ (almost sure extinction – rapid extinction)
- Conditional on $Z_n > 0$, regardless the value of $m < 1$

$$\hat{m}_n = \frac{\sum_{i=1}^{n} Z_i}{\sum_{i=1}^{n} Z_{i-1}} \longrightarrow 1 \qquad \text{as } n \to \infty \qquad \text{(Q-consistency)}$$



Geometric offspring distribution with mean $m = 0.9$, and $Z_0 = 100$

- For small values of n and large values of Z_0, $\hat{m}_n$ is a decent (but not efficient) estimator of n
- For large values of n, the sample $Z_0, Z_1, \ldots, Z_n > 0$ is too biased

Subcritical Galton–Watson branching processes

Assume that

$$\frac{Var[\xi]}{m} = am + b$$

This holds if ξ is geometric, Poisson, binomial (and more ?)

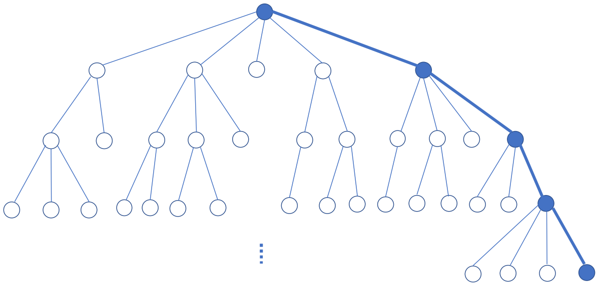
Theorem (Braunsteins, H., Minuesa (2019))

A C-consistent estimator for m is given by

$$\tilde{m}_n := \frac{\sum_{i=1}^{n}(Z_i - b)}{\sum_{i=1}^{n}(Z_{i-1} + a)}.$$

The Q-process $\{Z_\ell^{\uparrow}\}$ of a GW process

The size-biased tree



White nodes have offspring distribution ξ

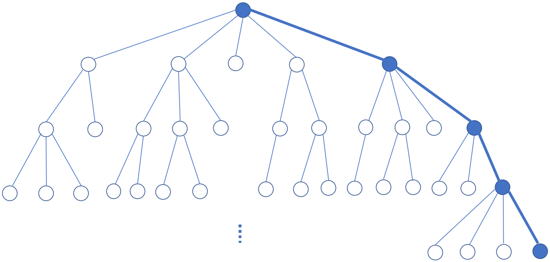
Blue nodes have offspring distribution SB[ξ] (size-biased distribution of ξ)

$$E[Z_\ell^{\uparrow} \mid Z_{\ell-1}^{\uparrow}] \;=\; (Z_{\ell-1}^{\uparrow} - 1)\, m + (E[\xi^2]/m) \;=\; Z_{\ell-1}^{\uparrow}\, m + (Var[\xi]/m)$$

$$\frac{Var[\xi]}{m} = am + b \quad \Rightarrow \quad m = \frac{E[Z_\ell^{\uparrow} \mid Z_{\ell-1}^{\uparrow}] - b}{Z_{\ell-1}^{\uparrow} + a}$$

Unbiasing

$$\frac{Var[\xi]}{m} = am + b, \qquad \tilde{m}_n := \frac{\sum_{i=1}^{n}(Z_i - b)}{\sum_{i=1}^{n}(Z_{i-1} + a)} = \frac{\text{children}}{\text{parents}}$$



$\xi \sim \text{Poi}(m)$: $a = 0,\ b = 1,$ $\text{SB}[\xi] \overset{d}{=} 1 + \xi$

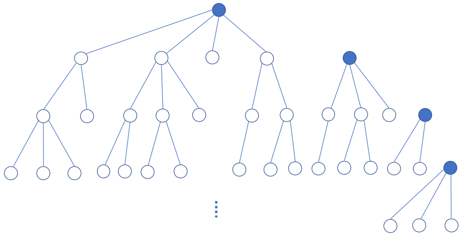
$\xi \sim \text{Geom}(1/(m+1))$: $a = 1,\ b = 1,$ $\text{SB}[\xi] \overset{d}{=} 1 + \xi + \xi'$

$\xi \sim 2\,\text{Ber}(m/2)$: $a = -1,\ b = 2,$ $\text{SB}[\xi] = 2$

Another useful interpretation of $\{Z_\ell^\uparrow\}$

$\{Z_\ell^\uparrow - 1\}$: GW process with immigration and offspring law ξ

Law of the number of immigrants : $\mathrm{SB}[\xi] - 1$



Theorem (Braunsteins, H., Minuesa (2019))

A C-consistent estimator for m is given by

$$\bar{m}_n := 1 - \frac{1}{2}\frac{\sum_{i=1}^n (Z_i - Z_{i-1})^2}{\sum_{i=1}^n (Z_{i-1} - \bar{Z}_n)^2}, \qquad \text{where } \bar{Z}_n := \frac{\sum_{i=1}^n Z_{i-1}}{n}$$

+ a C-consistent estimator for $Var[\xi]$.

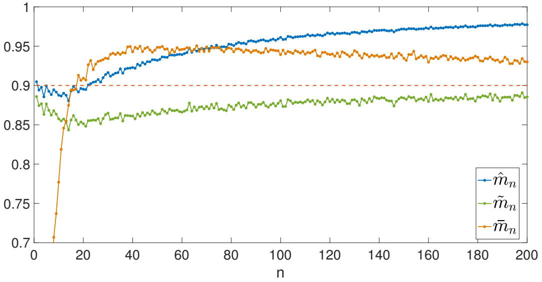
Comparison of the estimators

Given $Z_n > 0$,

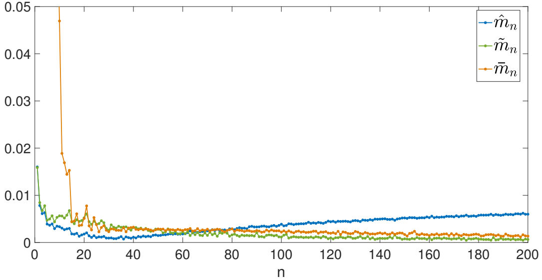
$$\hat{m}_n = \frac{\sum_{i=1}^n Z_i}{\sum_{i=1}^n Z_{i-1}} \to 1,$$

$$\tilde{m}_n := \frac{\sum_{i=1}^n (Z_i - b)}{\sum_{i=1}^n (Z_{i-1} + a)} \to m\ ,$$

$$\bar{m}_n := 1 - \frac{1}{2} \frac{\sum_{i=1}^n (Z_i - Z_{i-1})^2}{\sum_{i=1}^n (Z_{i-1} - \bar{Z}_n)^2} \to m$$



Mean squared error :



Geometric offspring distribution with mean $m = 0.9$, and $Z_0 = 100$

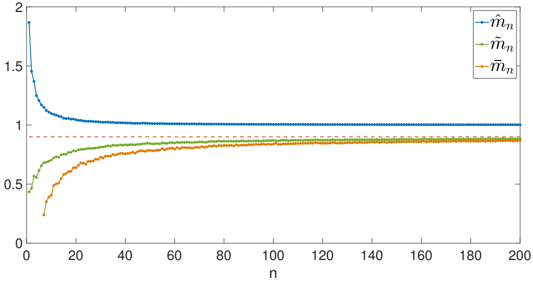
Comparison of the estimators

Given $Z_n > 0$,

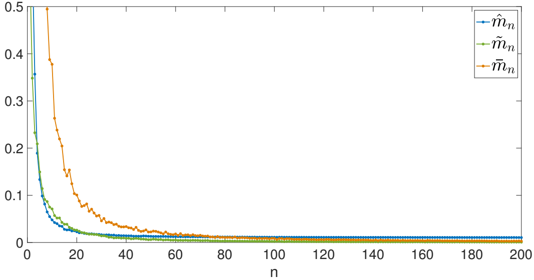
$$\hat{m}_n = \frac{\sum_{i=1}^n Z_i}{\sum_{i=1}^n Z_{i-1}} \to 1,$$

$$\tilde{m}_n := \frac{\sum_{i=1}^n (Z_i - b)}{\sum_{i=1}^n (Z_{i-1} + a)} \to m\ ,$$

$$\bar{m}_n := 1 - \frac{1}{2} \frac{\sum_{i=1}^n (Z_i - Z_{i-1})^2}{\sum_{i=1}^n (Z_{i-1} - \bar{Z}_n)^2} \to m$$



Mean squared error :



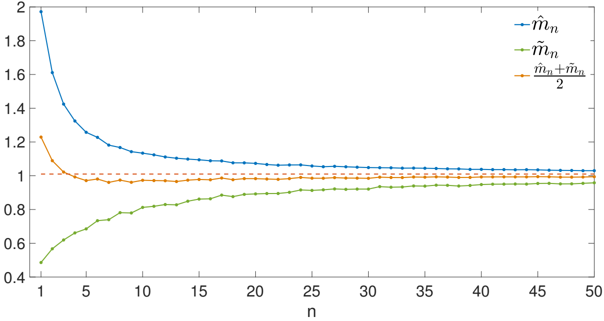
Geometric offspring distribution with mean $m = 0.9$, and $Z_0 = 1$

Supercritical Galton–Watson branching processes

- Mean offspring $m > 1$ (positive chance of survival)
- Conditional on $Z_n > 0$,

$$\hat{m}_n = \frac{\sum_{i=1}^{n} Z_i}{\sum_{i=1}^{n} Z_{i-1}} \longrightarrow m \qquad \text{as } n \to \infty \quad \text{(C-consistency)}$$

$$\tilde{m}_n = \frac{\sum_{i=1}^{n} (Z_i - b)}{\sum_{i=1}^{n} (Z_{i-1} + a)} \longrightarrow m \qquad \text{as } n \to \infty \quad \text{(C-consistency).}$$

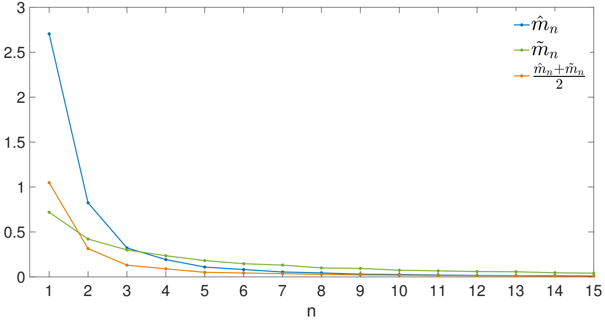


Geometric offspring distr. with mean $m = 1.01$, and $Z_0 = 1$

Supercritical Galton–Watson branching processes

Mean squared errors

$$\hat{m}_n = \frac{\sum_{i=1}^n Z_i}{\sum_{i=1}^n Z_{i-1}}, \qquad \tilde{m}_n = \frac{\sum_{i=1}^n (Z_i - b)}{\sum_{i=1}^n (Z_{i-1} + a)}$$



Geometric offspring distr. with mean $m = 1.01$, and $Z_0 = 1$

Conclusion

We studied the asymptotic properties of the estimator $\hat{m}_n(z)$ of the mean offspring at population size z in PSDBP

This lead us to

- the concept of 'Q-consistency' for estimators in branching processes which become extinct almost surely
- the construction of C-consistent estimators for subcritical GW processes

Ongoing work :

- How to use the estimators $\hat{m}_n(z)$ for $z \geq 1$ to estimate a single parameter θ ?
- How to construct C-consistent estimators for PSDBP ?

Main references

P. Braunsteins, S. Hautphenne, and C. Minuesa
Parameter estimation in branching processes with almost sure extinction.
In preparation, 2019.

P. A. Ernst, W. S. Kendall, G. O. Roberts, and J. S. Rosenthal.
MEXIT : Maximal un-coupling times for stochastic processes.
Stochastic Processes and their Applications, 129(2) :355–380, 2019.

F. Gosselin.
Asymtotic behaviour of absorbing Markov chains conditional on nonabsorption for applications in conservation biology.
The Annals of Applied Probability, 11(1) :261–284, 2001.

C. C. Heyde, and E. Seneta.
Estimation theory for growth and immigration rates in a multiplicative process.
Journal of Applied Probability, 9(2) :235–256, 1972.

A. G. Pakes.
Non-parametric estimation in the Galton-Watson process.
Mathematical Biosciences, 26(1) :1–18, 1975.

Thank you for your attention

ACEMS