

SEVENTH WORKSHOP DYNAMICAL SYSTEMS APPLIED TO BIOLOGY AND NATURAL SCIENCES

Évora (Portugal), February 2-5, 2016

TWO-SEX BRANCHING POPULATIONS

Manuel Molina

Department of Mathematics, University of Extremadura, Spain
mmolina@unex.es

GOBIERNO DE EXTREMADURA



- **Branching processes**
- **Two-sex branching processes**
- **Recent advances on two-sex branching processes**
- **Two-sex branching model with several mating and reproduction strategies**
- **Some questions for research**

BRANCHING PROCESSES

- Motivation
- Bienaymé-Galton-Watson model
- Asexual branching processes investigated

Motivation

The branching process theory was motivated, in the XIX-th century, by the explanation of the extinction phenomenon in certain aristocracy family lines in the European society.

This theory provides mathematical models to describe the probabilistic evolution of dynamical systems whose components (cells, particles, individuals in general) after certain life period reproduce and die, in such a way that transition from one to other state of the system is made according to a probability distribution.

After 1940, the interest in branching processes increased due to their applications in physical and biological sciences. The terminology branching processes was proposed in 1938 by A.N. Kolmogorov when he was interested on applications in the field of Biology.

The subject had major contributions from the Russian school of A.N. Kolmogorov and his many students in the 1940s, 1950s and 1960s. In particular, very important were the contributions by B.A. Sevastyanov.

Nowadays, branching process theory is an active research area of both theoretical interest and applicability to such fields as biology, demography, ecology, epidemiology, genetics, population dynamics, and others.

Bienaymé-Galton-Watson model

Introduced in 1845, the Bienaymé-Galton-Watson model is a discrete-time stochastic sequence $\{Z_n\}_{n=0}^{\infty}$, Z_n representing the total number of individuals in the population at time (generation) n , defined in the recursive form:

$$Z_0 = k_0 \in \mathbb{N}_+, \quad X_{0,1}, \dots, X_{0,k_0} \text{ iid } \{p_k\}_{k=0}^{\infty}, \quad p_k := P(X_{0,1} = k)$$

$$Z_1 := \sum_{i=1}^{k_0} X_{0,i}$$

$$Z_n \in \mathbb{N}_+, \quad X_{n,1}, \dots, X_{n,Z_n} \text{ iid } \{p_k\}_{k=0}^\infty$$

$$Z_{n+1} := \sum_{i=1}^{Z_n} X_{n,i}, \quad n \in \mathbb{N}$$

$\{p_k\}_{k=0}^\infty$, $p_k := P(X_{0,1} = k)$, offspring distribution

- $\{Z_n\}_{n=0}^\infty$ is a homogeneous Markov chain.
- 0 is an absorbent state.
- Each $k \neq 0$ is a transient state.

- If for some positive integer n , $Z_n = 0$, then $Z_{n+k} = 0$, $k \in \mathbb{N}_+$.

$$Q := P(\lim_{n \rightarrow \infty} Z_n = 0).$$

- Q is the smallest root of the equation $f(s) = s$ in $[0, 1]$.

$$\{p_k\}_{k=0}^{\infty}, \quad f(s) := \sum_{k=0}^{\infty} s^k p_k, \quad s \in [0, 1].$$

- Extinction-explosion duality:

$$P(\lim_{n \rightarrow \infty} Z_n = 0) + P(\lim_{n \rightarrow \infty} Z_n = \infty) = 1.$$

- $m := \sum_{k=1}^{\infty} kp_k.$

If $m \leq 1$ then $Q = 1$.

If $m > 1$ then $Q < 1$.

- Classification:

$m < 1$ Subcritical case.

$m = 1$ Critical case.

$m > 1$ Supercritical case.

Several theoretical questions have been investigated, see e.g., Dion (1974), Harris (1948), Heyde (1970, 1971, 1974), Heyde and Seneta (1972), Kesten and Stigum (1966), Kolmogorov (1938), Nagaev (1967), Seneta (1968), Pakes (1970, 1975) or Yaglom (1947).

In 1922, R. Fisher considered the Bienaymé-Galton-Watson model under a genetic context. From then on, it has been applied as mathematical model in different fields, see e.g., Bartoszynsky (1966), Becker (1975, 1976), Bruss and Slatvchova-Bojkova (1999) or Pérez-Abreu (1987).

Asexual branching processes investigated

Also, from the Bienaymé-Galton-Watson model general context, several *asexual reproduction branching processes* have been introduced and investigated, including models with immigration; models in varying environments; models in random environments; controlled models; multitype models; and some special models.

- **T. Harris (1963).**
The theory of branching processes.
Springer-Verlag.
- **K. Athreya and P. Ney (1972).**
Branching processes.
Springer-Verlag.
- **G. Asmussen and H. Hering (1983).**
Branching processes.
Birkhauser.
- **G. Sankaranarayanan (1989).**
Branching processes and its estimation theory.
John Wiley and Sons.
- **P. Guttorp (1991).**
Statistical inference for branching processes.
John Wiley and Sons.

Applications in cell kinetics, cellular biology, chemotherapy, gene amplification, human evolution or molecular biology.

- **P. Jagers (1975).**
Branching processes with biological applications.
John Wiley and Sons.
- **M. Kimmel and D. Axelrod (2002).**
Branching processes in Biology.
Springer-Verlag.
- **A. Pakes (2003).**
Biological applications of branching processes.
Handbook of Statistics, v. 21. Elsevier Sciences B.V.
- **P. Haccou, P. Jagers and V. Vatutin (2005).**
Branching processes: variation, growth, and extinction of populations.
Cambridge University Press.

Branching Processes Research Group University of Extremadura (Spain)

<http://www.unex.es/investigacion/grupos/gipra>

The group was created in 1990 with the main purpose to study some methodological problems, not sufficiently investigated, about certain branching processes.

From then on, we have studied several theoretical and practical questions about some classes of branching processes, especially models with immigration; controlled models; multitype models; and two-sex models. We have also considered applications of such models, mainly in the fields of epidemiology, genetics and population dynamics.

Branching Processes Research Group University of Extremadura (Spain)

<http://www.unex.es/investigacion/grupos/gipra>



M. Molina



M. González



M. Mota



A. Ramos



I. del Puerto



R. Martínez



C. Gutiérrez



C. Minuesa

TWO-SEX BRANCHING PROCESSES

- Motivation
- Bisexual Galton-Watson model
- Two-sex branching processes investigated

Motivation

In the Symposium held at Wistar Institute of Anatomy and Biology (Philadelphia), in 1966, S.M. Ulam, in his presentation entitled *How to formulate mathematically problems of rate of evolution*, pointed out:

There is a very nice and simple mathematical technique for describing processes starting with a single object, which then duplicates and gives 0, 1, 2 or more descendants. It is called the theory of branching processes. It deals with asexual reproduction and gives methods to calculate the number of existing particles, of various kinds, in future generations, and other questions of this sort. I would like to stress that a corresponding theory for branching with sex, where particles get together, say at random and then produce offspring, i.e., a combination of a binary process of mating and reproduction, is mathematically much more difficult and no exact theory exists as yet.

Bisexual Galton-Watson model

D.J. Daley (1968)

Extinction conditions for certain bisexual Galton-Watson branching processes.
Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 9, 315-322.

- The population of the species under consideration consists of two disjoint types of individuals: females and males.
- Two biological phases are considered: mating and reproduction.
In the mating phase, the couples (mating units) are formed according to a certain function. A couple consists of one female and one male from the same generation who come together for the purpose of reproduction.
In the reproduction phase, each couple produces new descendants (females and males) according to a probability distribution.

$$\{(F_n, M_n)\}_{n=1}^{\infty}, \quad \{Z_n\}_{n=0}^{\infty}$$
$$(F_{n+1}, M_{n+1}) := \sum_{i=1}^{Z_n} (f_{n,i}, m_{n,i}), \quad Z_{n+1} := L(F_{n+1}, M_{n+1}), \quad n \in \mathbb{N}$$

- (F_{n+1}, M_{n+1}) represents the number of females and males at time (generation) $n + 1$.
- These females and males form Z_{n+1} couples through the mating function L .
- Initially, we assume $k_0 \in \mathbb{N}_+$ couples in the population, $Z_0 = k_0$.

$$\{(F_n, M_n)\}_{n=1}^{\infty}$$

$$(F_{n+1}, M_{n+1}) := \sum_{i=1}^{Z_n} (f_{n,i}, m_{n,i}), \quad Z_{n+1} := L(F_{n+1}, M_{n+1}), \quad n \in \mathbb{N}$$

- $(f_{n,i}, m_{n,i})$ is the random vector representing the number of females and males descending from the couple i of the n -th generation.

$(f_{n,i}, m_{n,i})$, $n \in \mathbb{N}$, $i \in \mathbb{N}_+$, are iid random vectors.

Their probability law is called the offspring distribution:

$$\{p_{k,l}\}_{k,l=0}^{\infty}, \quad p_{k,l} = P(f_{0,1} = k, m_{0,1} = l)$$

$$\{(F_n, M_n)\}_{n=1}^{\infty}$$
$$(F_{n+1}, M_{n+1}) := \sum_{i=1}^{Z_n} (f_{n,i}, m_{n,i}), \quad Z_{n+1} := L(F_{n+1}, M_{n+1}), \quad n \in \mathbb{N}$$

- The mating function $L: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$, is assumed to be non-decreasing and such that:

$$L(0, m) = L(f, 0) = 0, \quad f, m \in \mathbb{N}$$

A generation without females or males will not be able to form couples.

$$\{(F_n, M_n)\}_{n=1}^{\infty}, \quad \{Z_n\}_{n=0}^{\infty}$$
$$(F_{n+1}, M_{n+1}) := \sum_{i=1}^{Z_n} (f_{n,i}, m_{n,i}), \quad Z_{n+1} := L(F_{n+1}, M_{n+1}), \quad n \in \mathbb{N}$$

- $\{F_n, M_n\}_{n=1}^{\infty}$ and $\{Z_n\}_{n=0}^{\infty}$ are homogeneous Markov chains.
- If, for some $n \in \mathbb{N}_+$, $Z_n = 0$, then $(F_{n+k}, M_{n+k}) = (0, 0)$, $k \in \mathbb{N}_+$.
Hence, the extinction of the population occurs.

$$Q_{k_0} := P\left(\lim_{n \nearrow \infty} Z_n = 0 \mid Z_0 = k_0\right)$$

Extinction probability when initially there are k_0 couples.

D.J. Daley (1968)

Extinction conditions for certain bisexual Galton-Watson branching processes.
Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 9, 315-322.

$L(f, m) = f \min\{1, m\}$ (completely promiscuous mating)

$$Q_{k_0} = 1 \Leftrightarrow E[f_{0,1}] \leq 1$$

$L(f, m) = \min\{f, Km\}$, $K \in \mathbb{N}_+$ (polygamous mating with perfect fidelity)

$$Q_{k_0} = 1 \Leftrightarrow \min\{E[f_{0,1}], KE[m_{0,1}]\} \leq 1$$

D. J. Daley, D. Hull, J. M. Taylor (1986)

Bisexual Galton-Watson branching processes with superadditive mating function.
Journal of Applied Probability, 23, 585-600.

Superadditive mating function:

$$L(f_1 + f_2, m_1 + m_2) \geq L(f_1, m_1) + L(f_2, m_2), \quad f_i, m_i \in \mathbb{N}, \quad i = 1, 2$$

The superadditivity expresses the intuitive notion that $f_1 + f_2$ females and $m_1 + m_2$ males coexisting together will form a number of couples that is at least as great as the total number of couples formed by f_1 females and m_1 males, and f_2 females and m_2 males, living separately.

$$L(f, m) = fm; \quad L(f, m) = \lfloor (f + m)/2 \rfloor; \quad L(f, m) = \lfloor (fm)^{1/2} \rfloor; \\ L(f, m) = \lfloor 2fm/(f + m) \rfloor; \quad f, m \in \mathbb{N}.$$

- Asymptotic growth rate:

$$R := \lim_{k \rightarrow \infty} \frac{1}{k} E[Z_{n+1} \mid Z_n = k]$$

- For all $k_0 \in \mathbb{N}_+$,

$$Q_{k_0} = 1 \Leftrightarrow R \leq 1$$

- Classification:

$R < 1$ Subcritical case.

$R = 1$ Critical case.

$R > 1$ Supercritical case.

Several theoretical and applied contributions to the bisexual Galton-Watson process theory have been derived.

D. Hull (2003).

A survey of the literature associated with the bisexual Galton-Watson process.
Extracta Mathematicae, 18, 321-343.

P. Haccou, P. Jagers, V. Vatutin (2005).

Branching Processes: Variation, growth and extinction of populations.
Cambridge University Press.
G. Alsmeyer, pages: 43-46, 135-145, 177-179.

Two-sex branching processes investigated

From Daley's bisexual model, different classes of two-sex branching processes have been developed, including models with immigration; models depending on the number of couples in the population; models in varying environments; models in random environments; etc.

M. Molina (2010)

Two-sex branching process literature.

Lectures Notes in Statistics, vol. 197, 279-293.

Springer-Verlag.

RECENT ADVANCES ON TWO-SEX BRANCHING PROCESSES

- **Motivation**
- **New research lines**

Motivation

- We are interested in to develop stochastic models to describe the demographic dynamics of biological populations with sexual reproduction.
- The range of two-sex processes investigated is not large enough in order to describe the probabilistic behavior of several biological populations. In particular, significant efforts have been made to develop two-sex processes based on the assumptions that, in each generation:

The number of couple which take part in the reproduction phase (progenitor couples) is determined in a predictable environment.

Mating and reproduction (or both phases) are affected by the number of couples in the population.

However, in many biological populations, due to several random factors (weather conditions, food supply, fertility parameters, predators, human activities, and others):

- The number of progenitor couples is determined in a non-predictable environment.
- Mating and reproduction are usually influenced by the numbers of females and males in the population.

In an attempt to contribute some solution to the mathematical modeling in such populations, in the last years, in collaboration with some researchers from China and France, we have considered two new research lines on two-sex branching processes:

Research line 1

To investigate two-sex branching models where, in each generation, the number of progenitor couples is determined in a non-predictable environment.



M. Molina
UEX (Spain)



M. Mota
UEX (Spain)



A. Ramos
UEX (Spain)



S. Ma
HUT (China)



Y. Xing
SIT (China)

M. Molina, M. Mota, A. Ramos (2012)

Two-sex branching models with random control on the number of progenitor couples. Methodology and Computing in Applied Probability, 14, 35-48.

S. Ma, M. Molina, Y. Xing (2012)

Two-sex branching populations with progenitor couples in a random environment. Mathematical Population Studies, 19, 177-187.

M. Molina, M. Mota, A. Ramos (2015)

Mathematical modeling in biological populations through branching processes. Application to salmonid populations. Journal of Mathematical Biology, 70, 197-212.

S. Ma, M. Molina, Y. Xing (2016)

Two-sex branching processes with reproduction phase in a random environment. Stochastics (accepted)

Research line 2

To investigate two-sex branching models where, in each generation, mating and reproduction are influenced by the numbers of females and males in the population.



M. Molina
UEX (Spain)



M. Mota
UEX (Spain)



A. Ramos
UEX (Spain)



C. Jacob
INRA (France)

M. Molina, M. Mota, A. Ramos (2013)

Modeling in random environments through two-sex branching processes. Applied Methods of Statistical Analysis, vol. 1, 206-213. S.B. Lemesko, M. Nikukin, N. Balakrishnan, (eds), NSTU Publisher, ISBN 978-5-7782-2311-0.

M. Molina, M. Mota, A. Ramos (2014)

Stochastic modeling in biological populations with sexual reproduction through branching models. Application to Coho salmon populations. Mathematical Biosciences, 258, 182-188.

C. Jacob, M. Molina, M. Mota (2016)

A general class of population-dependent two-sex processes with random mating. Bernoulli (accepted)

M. Molina, M. Mota, A. Ramos (2016)

Two-sex branching processes with several mating and reproduction strategies: extinction versus survival. Lectures Notes in Statistics (accepted)

TWO-SEX BRANCHING MODEL WITH SEVERAL MATING AND REPRODUCTION STRATEGIES

- Definition of the model
- Probabilistic results
- Inferential results
- Application

Definition of the model

Let us consider biological populations with sexual reproduction with the following basic characteristics:

- Several strategies for the mating between females and males are feasible, depending on the current numbers of females and males in the population. One of them is implemented in each generation.
- Once formed the couples, there are several reproduction strategies to produce new descendants. The choice of the reproduction strategy depends also of the current numbers of females and males in the population.

Let N_m and N_r be positive integers representing the numbers of mating and reproduction strategies, respectively. We model the dynamics of such populations as follows:

Mating phase

- Let $L_l, l = 1, \dots, N_m$, be deterministic functions, on \mathbb{N}^2 and taking values in \mathbb{N} . The role of L_l is to determine the number of couples formed in the population when the l -th mating strategy is considered. Each function L_l is assumed to be non-decreasing and such that, $L_l(f, 0) = L_l(0, m) = 0, f, m \in \mathbb{N}$. These functions will be referred as to *mating functions*.

Reproduction phase

- Let $\{p_{k,j}^h\}_{(k,j) \in S_h}$, $h = 1, \dots, N_r$, be probability distributions.

$p_{k,j}^h$ is the probability that a couple produces k females and j males when the h -th reproduction strategy is considered. In this reproduction phase, it is assumed that the couples act independently of the others and, inside a generation, all of them adopt the same reproductive strategy. These probability distributions will be referred as to *offspring distributions*.

Determination of the mating function and the offspring distribution

- Let ψ and φ be deterministic functions on \mathbb{N}^2 and taking values in the sets $\{1, \dots, N_m\}$ and $\{1, \dots, N_r\}$ respectively.

ψ and φ determine, according to the number of females and males in the population, the mating function and the offspring distribution to be considered in the mating and reproduction phases, respectively.

For instance, assume that in some generation there are f females and m males in the population. If $\psi(f, m) = l$ and $\varphi(f, m) = h$ then L_l and $\{p_{k,j}^h\}_{(k,j) \in S_h}$ are the mating function and the offspring distribution, respectively, to be considered in such generation.

Let us define the process $\{\mathbf{X}_n\}_{n=0}^{\infty}$, where $\mathbf{X}_n = (F_n, M_n)$ is the random vector representing the numbers of females and males at time (generation) n , in the form:

- $P(\mathbf{X}_0 = \mathbf{x}_0) = 1$, $\mathbf{x}_0 = (f_0, m_0)$, $f_0 \geq 1$, $m_0 \geq 1$.
- If, in the n -th generation, $\mathbf{X}_n = \mathbf{x} = (f, m) \in \mathbb{N}_+^2$ then $Z_n = L_{l_n}(\mathbf{x})$ couples are formed where $l_n = \psi(\mathbf{x})$. Each of them produces a random number of females and males, according to $\{p_{k,j}^{h_n}\}_{(k,j) \in S_{h_n}}$, $h_n = \varphi(\mathbf{x})$, and then disappears. Hence,

$$\mathbf{X}_{n+1} = \sum_{i=1}^{Z_n} (f_{n,i}^{h_n}, m_{n,i}^{h_n})$$

where $(f_{n,i}^{h_n}, m_{n,i}^{h_n})$, $i = 1, \dots, Z_n$ are iid random vectors such that:

$$P(f_{n,1}^{h_n} = k, m_{n,1}^{h_n} = j) = p_{k,j}^{h_n}, (k, j) \in S_{h_n}.$$

$\{\mathbf{X}_n\}_{n=0}^{\infty}$ is a homogeneous Markov chain with state space on \mathbb{N}^2 developing in an environment which changes, stochastically in time, influenced by the current number of females and males in the population.

In addition to its theoretical interest, this two-sex process has also practical implications. Really, it is appropriate to model the probabilistic evolution of semelparous species (biological species characterized by a single reproductive episode before death).

Probabilistic results

Next we provide some probabilistic results. To this end, we will consider the sequence $\{g(\mathbf{X}_n)\}_{n=0}^{\infty}$ where g has the functional form:

$$g(f, m) = a_1 f + a_2 m, \quad a_1 \geq 0, \quad a_2 \geq 0, \quad a_1 + a_2 > 0, \quad (f, m) \in \mathbb{N}^2 \quad (1)$$

Some particular cases of special interest are:

- $g(f, m) = f$
- $g(f, m) = m$
- $g(f, m) = f + m$

corresponding to the numbers of females, males, or total individuals, respectively.

Result 1

Let us consider the process $\{\mathbf{X}_n\}_{n=0}^{\infty}$. Let $\mathbf{x} \in \mathbb{N}^2$, $l = \psi(\mathbf{x})$, $h = \varphi(\mathbf{x})$, and g as in (1). Then:

- $E[\mathbf{X}_{n+1} \mid \mathbf{X}_n = \mathbf{x}] = L_l(\mathbf{x})\mu^h.$
- $Var[\mathbf{X}_{n+1} \mid \mathbf{X}_n = \mathbf{x}] = L_l(\mathbf{x})\Sigma^h.$
- $E[g(\mathbf{X}_{n+1}) \mid \mathbf{X}_n = \mathbf{x}] = L_l(\mathbf{x})g(\mu^h).$
- $Var[g(\mathbf{X}_{n+1}) \mid \mathbf{X}_n = \mathbf{x}] = L_l(\mathbf{x})\mathbf{a}\Sigma^h\mathbf{a}', \mathbf{a} = (a_1, a_2)$

where μ^h and Σ^h denote, respectively, the mean vector and the covariance matrix of the h -th offspring distribution.

Result 2

Let us consider the process $\{\mathbf{X}_n\}_{n=0}^\infty$. Then:

- $\mathbf{0} = (0, 0)$ is an absorbing state.
- $(f, 0), (0, m), f, m \in \mathbb{N}_+$ are transient states.
- Let $\mathbf{x} \in \mathbb{N}_+^2$ and $h = \varphi(\mathbf{x})$. If

$$\max \left\{ \sum_{k=0}^{\infty} p_{k,0}^h, \sum_{j=0}^{\infty} p_{0,j}^h \right\} > 0$$

then \mathbf{x} is a transient state.

Result 3

Let us consider the process $\{\mathbf{X}_n\}_{n=0}^{\infty}$ and g as in (1) with $a_1, a_2 > 0$. Assume that, for all $h \in \{1, \dots, N_r\}$,

$$\max \left\{ \sum_{k=0}^{\infty} p_{k,0}^h, \sum_{j=0}^{\infty} p_{0,j}^h \right\} > 0$$

Then:

$$P \left(\lim_{n \rightarrow \infty} \mathbf{X}_n = \mathbf{0} \right) + P \left(\lim_{n \rightarrow \infty} g(\mathbf{X}_n) = \infty \right) = 1 \quad (2)$$

Definition

Let $\mathbf{x} = (f, m) \in \mathbb{N}_+^2$ and g as in (1). In order to study results concerning the extinction/survival of the process, the following rate will play an important role:

$$m_g(\mathbf{x}) := \frac{1}{g(\mathbf{x})} E[g(\mathbf{X}_{n+1}) \mid \mathbf{X}_n = \mathbf{x}]$$

From Result 1,

$$m_g(\mathbf{x}) = \frac{1}{g(\mathbf{x})} L_l(\mathbf{x}) g(\mu^h), \quad l = \psi(\mathbf{x}), \quad h = \varphi(\mathbf{x})$$

Definition

$$q(\mathbf{x}_0) := P\left(\lim_{n \rightarrow \infty} \mathbf{X}_n = \mathbf{0} \mid \mathbf{X}_0 = \mathbf{x}_0\right), \quad \mathbf{x}_0 = (f_0, m_0) \in \mathbb{N}_+^2$$

Result 4 (Extinction)

Let us consider the process $\{\mathbf{X}_n\}_{n=0}^{\infty}$ and g as in (1) such that (2) holds and, for some $g^* \geq 0$,

$$\sup_{\{\mathbf{x}: g(\mathbf{x}) > g^*\}} m_g(\mathbf{x}) \leq 1$$

Then: $q(\mathbf{x}_0) = 1$ for all $\mathbf{x}_0 \in \mathbb{N}_+^2$.

Result 5 (Survival)

Let us consider the process $\{\mathbf{X}_n\}_{n=0}^\infty$, g as in (1), and $g^* \geq 0$, such that:

- $\inf_{\{\mathbf{x}: g(\mathbf{x}) > g^*\}} m_g(\mathbf{x}) > 1$
- There exists $\gamma > 1$ verifying that

$$\sum_{k=0}^{\infty} \sup_{\{\mathbf{x}: g(\mathbf{x}) > \nu^k g(\mathbf{x}_0)\}} \frac{M_g^\gamma(\mathbf{x})}{g(\mathbf{x})^\gamma} < \infty$$

$$M_g^\gamma(\mathbf{x}) := E[|g(\mathbf{X}_{k+1}) - E[g(\mathbf{X}_{k+1} | \mathbf{X}_k = \mathbf{x})]|^\gamma | \mathbf{X}_k = \mathbf{x}]$$

Then, for all $\mathbf{x}_0 \in \mathbb{N}_+^2$ with $g(\mathbf{x}_0) > g^*$,

$$P\left(\lim_{n \rightarrow \infty} g(\mathbf{X}_n) = \infty \mid \mathbf{X}_0 = \mathbf{x}_0\right) > 0$$

Inferential results

$$\mu^h = (\mu_1^h, \mu_2^h), \quad \Sigma^h = \left(\sigma_{ij}^h \right)_{i,j=1,2}, \quad h \in \{1, \dots, N_r\}$$

$$\mu_1^h = E[f_{0,1}^h], \quad \mu_2^h = E[m_{0,1}^h]$$

$$\sigma_{11}^h = \text{Var}[f_{0,1}^h], \quad \sigma_{22}^h = \text{Var}[m_{0,1}^h], \quad \sigma_{12}^h = \sigma_{21}^h = \text{Cov}[f_{0,1}^h, m_{0,1}^h]$$

Expected number of females and males per couple, and the corresponding variances and covariance, for the h -th reproductive strategy.

Bayesian estimation

We shall consider the estimation, from a Bayesian point of view, of the reproductive parameters:

$$\mu_i^h, \sigma_{ij}^h, i, j = 1, 2, h = 1, \dots, N_r.$$

To this end, we assume the observation of the number of females and males up to a certain pre-set generation $n > 1$ is reached:

$$\mathcal{A}_n := \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n\}, \mathbf{x}_i = (f_i, m_i) \in \mathbb{N}^2$$

Let $\{\mathbf{X}_n\}_{n=0}^\infty$ be a two-sex process with offspring distributions $\{p_{k,j}^h\}_{(k,j) \in S_h}$, $h = 1, \dots, N_r$, belonging to the bivariate power series distributions family. Hence, $p_{k,j}^h$ can be rewritten in the functional form:

$$p_{k,j}^h = p_{k,j}^h(\theta_1^h, \theta_2^h) = \left(A_h(\theta_1^h, \theta_2^h) \right)^{-1} (\theta_1^h)^k (\theta_2^h)^j a_{k,j}^h, \quad (3)$$

$$A_h(\theta_1^h, \theta_2^h) = \sum_{k,j \in S_h} (\theta_1^h)^k (\theta_2^h)^j a_{k,j}^h, \quad (\theta_1^h, \theta_2^h) \in \Theta_h \subseteq \mathbb{R}^2,$$

$$\Theta_h = \{(\theta_1^h, \theta_2^h) : (\theta_1^h)^k (\theta_2^h)^j a_{k,j}^h \geq 0, A_h(\theta_1^h, \theta_2^h) < \infty\},$$

(θ_1^h, θ_2^h) are vectors of parameters and $a_{k,j}^h$ are real constants.

Because $\{\mathbf{X}_n\}_{n=0}^\infty$ is a Markov chain and $P(\mathbf{X}_0 = \mathbf{x}_0) = 1$,

$$P(\mathbf{X}_0 = \mathbf{x}_0, \dots, \mathbf{X}_n = \mathbf{x}_n) = \prod_{j=1}^n P\left(\sum_{i=1}^{z_{j-1}} \left(f_{j-1,i}^{\varphi(\mathbf{x}_{j-1})}, m_{j-1,i}^{\varphi(\mathbf{x}_{j-1})}\right) = \mathbf{x}_j\right) \quad (4)$$

$$z_{j-1} = L_{\psi(\mathbf{x}_{j-1})}(\mathbf{x}_{j-1}), \quad j = 1, \dots, n$$

From (3), for $j = 1, \dots, n$,

$$P\left(\sum_{i=1}^{z_{j-1}} \left(f_{j-1,i}^{\varphi(\mathbf{x}_{j-1})}, m_{j-1,i}^{\varphi(\mathbf{x}_{j-1})}\right) = \mathbf{x}_j\right) \propto$$

$$\left(A_{\varphi(\mathbf{x}_{j-1})}(\theta_1^{\varphi(\mathbf{x}_{j-1})}, \theta_2^{\varphi(\mathbf{x}_{j-1})})\right)^{-z_{j-1}} (\theta_1^{\varphi(\mathbf{x}_{j-1})})^{f_j} (\theta_2^{\varphi(\mathbf{x}_{j-1})})^{m_j}. \quad (5)$$

Let \mathcal{L} be the likelihood function. From (4) and (5), we derive that:

$$\mathcal{L}(\theta_1^1, \theta_2^1, \dots, \theta_1^{N_r}, \theta_2^{N_r} \mid \mathcal{A}_n) \propto \prod_{h=1}^{N_r} (A_h(\theta_1^h, \theta_2^h))^{-z_h^*} (\theta_1^h)^{f_h^*} (\theta_2^h)^{m_h^*} \quad (6)$$

$$z_h^* = \sum_{j=1}^n z_{j-1} I_{j,h}, \quad f_h^* = \sum_{j=1}^n f_j I_{j,h}, \quad m_h^* = \sum_{j=1}^n m_j I_{j,h},$$

$$I_{j,h} = 1 \text{ if } \varphi(\mathbf{x}_{j-1}) = h; \text{ or } 0 \text{ if } \varphi(\mathbf{x}_{j-1}) \neq h.$$

Prior distribution

Taking into account expression (6), we deduce as appropriate conjugate class of prior distributions on $\Theta = \prod_{h=1}^{N_r} \Theta_h$:

$$\pi(\theta_1^1, \theta_2^1, \dots, \theta_1^{N_r}, \theta_2^{N_r}) = \prod_{h=1}^{N_r} \delta_{\alpha_{0h}, \alpha_{1h}, \alpha_{2h}}^{-1} (A_h(\theta_1^h, \theta_2^h))^{-\alpha_{0h}} (\theta_1^h)^{\alpha_{1h}} (\theta_2^h)^{\alpha_{2h}} \quad (7)$$

α_{0h} , α_{1h} and α_{2h} are constants such that:

$$\delta_{\alpha_{0h}, \alpha_{1h}, \alpha_{2h}} = \int_{\Theta_h} (A_h(\theta_1^h, \theta_2^h))^{-\alpha_{0h}} (\theta_1^h)^{\alpha_{1h}} (\theta_2^h)^{\alpha_{2h}} d\theta_1^h d\theta_2^h < \infty \quad (8)$$

This general class of prior distributions is flexible enough to describe different prior beliefs, see e.g. Guttorp (1991) (*Statistical Inference for Branching Processes*, John-Wiley and Sons.)

Posterior distribution

From (6) and (7), we derive as posterior distribution on Θ :

$$\pi(\theta_1^1, \theta_2^1, \dots, \theta_1^{N_r}, \theta_2^{N_r} \mid \mathcal{A}_n) = \prod_{h=1}^{N_r} \delta_{\alpha_{0h}^*, \alpha_{1h}^*, \alpha_{2h}^*}^{-1} (A_h(\theta_1^h, \theta_2^h))^{-\alpha_{0h}^*} (\theta_1^h)^{\alpha_{1h}^*} (\theta_2^h)^{\alpha_{2h}^*}$$

$$\alpha_{0h}^* = \alpha_{0h} + z_h^*, \quad \alpha_{1h}^* = \alpha_{1h} + f_h^*, \quad \alpha_{2h}^* = \alpha_{2h} + m_h^*$$

Result 6 (Bayes estimators)

- $\widehat{\mu}_i^h = \sum_{(k_1, k_2) \in S_h} k_i \Delta_{k_1, k_2}^h, \quad i = 1, 2$
- $\widehat{\sigma}_{ij}^h = \sum_{(k_1, k_2) \in S_h} k_i k_j \Delta_{k_1, k_2}^h - \sum_{(k_1, k_2) \in S_h} \sum_{(l_1, l_2) \in S_h} k_i l_j \Delta_{k_1, k_2; l_1, l_2}^h, \quad i, j = 1, 2$

where, for $(k_1, k_2), (l_1, l_2) \in S_h$,

$$\Delta_{k_1, k_2}^h = a_{k_1, k_2}^h \delta_{\alpha_{0h}^*, \alpha_{1h}^*, \alpha_{2h}^*}^{-1} \delta_{\alpha_{0h}^* + 1, \alpha_{1h}^* + k_1, \alpha_{2h}^* + k_2},$$

$$\Delta_{k_1, k_2; l_1, l_2}^h = a_{k_1, k_2}^h a_{l_1, l_2}^h \delta_{\alpha_{0h}^*, \alpha_{1h}^*, \alpha_{2h}^*}^{-1} \delta_{\alpha_{0h}^* + 2, \alpha_{1h}^* + k_1 + l_1, \alpha_{2h}^* + k_2 + l_2}.$$

High posterior density credibility sets

From the posterior distributions $\pi(\theta_1^h, \theta_2^h \mid \mathcal{A}_n)$, $h = 1, \dots, N_r$ we can determine high posterior density credibility sets:

$$\mathcal{I}_h(c) = \{(\theta_1^h, \theta_2^h) : \pi(\theta_1^h, \theta_2^h \mid \mathcal{A}_n) \geq c\}$$

where, given a credibility coefficient $1 - \alpha$, the constant c is chosen such that

$$\int_{\mathcal{I}_h(c)} \pi(\theta_1^h, \theta_2^h \mid \mathcal{A}_n) d\theta_1^h d\theta_2^h = 1 - \alpha$$

Hence, from the posterior distributions of μ_i^h and σ_{ij}^h , $i, j = 1, 2$, we can derive the corresponding highest posterior density credibility sets. It is not easy to obtain closed expressions for such posterior distributions. In the following result we propose a method to obtain accurate approximations.

Result 7

$\pi(\mu_i^h \mid \mathcal{A}_n), \pi(\sigma_{ij}^h \mid \mathcal{A}_n), i, j = 1, 2, h = 1, \dots, N_r$ can be approximated by considering the following computational algorithm:

- To generate a sufficiently large number of values for (θ_1^h, θ_2^h) from the posterior density $\pi(\theta_1^h, \theta_2^h \mid \mathcal{A}_n)$.
- To compute, from the generated data in the previous step, the corresponding values of μ_i^h and $\sigma_{ij}^h, i, j = 1, 2$.
- To estimate, the posterior densities of μ_i^h and $\sigma_{ij}^h, i, j = 1, 2$, with the data simulated in the previous step. To this end, Gaussian kernel estimation can be used, see e.g. Silverman(1986) *Density Estimation, Chapman and Hall.*)

Software developed

In order to simulate the two-sex process and calculate the Bayes estimations and the corresponding HPD credibility sets, we have developed some specific software. To this end, we have considered the language and environment for statistical computing and graphics R.

R Development Core Team: A language and Environment for Statistical Computing. R Foundation for Statistical Computing, 2009.

<http://www.r-project.org>

Application

We shall consider an application to Coho salmon populations. This type of salmon, found along the west coast of North America between California and Alaska, constitute a highly valued species.

They born in fresh water, spend most of their life in the sea and, at time of reproduction, return to fresh water to spawn. The spawning process involves mature male and female salmons swimming upstream, overcoming strong rivers currents, waterfalls, and other obstacles to reach their home spawning ground. The female releases her eggs and the male fertilizes them. They spawn only once and then both parents die.

The number of offspring significantly reduces due to various causes: they are the prey of a variety of animals (squawfish, mergansers, kingfishers, bears, etc.); the human activities (mainly fishing); some environmental factors (elevated water temperatures, stream damming, decomposition of organics, etc.), and others.

The salmon is called smolt when it is ready to migrate out to the sea. From various studies, see e.g. Laufle et al. (*Serv. Biol. Rep.* 82 (1986)), the Washington Department of Fisheries has estimated an average production of 75 smolts per couple.

The demographic dynamics of smolt Coho salmon is developed in a non-predictable environment influenced by the number of female and male smolts in the population. See for more details Fleming (1996) or Hautekeete et al. (2001).

These characteristics make them interesting species for describing their demographic dynamics (in a first mathematical attempt) through the class of two-sex branching models considered in this work.

As illustration, let us assume a Coho salmon population such that if, in certain generation, there are f female and m male smolts in the population then:

Mating strategies

The female and male smolts form couples taking into account three mating strategies ($N_m = 3$):

$$L_1(f, m) = \lfloor (1 \wedge f)mc_1 \rfloor, L_2(f, m) = \lfloor (f \wedge m)c_1 \rfloor, L_3(f, m) = \lfloor (1 \wedge m)fc_1 \rfloor$$

$c_1 \in (0, 1)$ represents a coefficient which takes into account the rate of decline in the smolt population due to predators, human activity and other environmental factors.

In a given generation, the mating function is determined through the following function ψ :

$$\psi(f, m) = 1 \text{ if } m \leq c_2^{-1}f; \quad 2 \text{ if } c_2^{-1}f < m < c_2f; \quad \text{or } 3 \text{ if } m \geq c_2f$$

where $c_2 > 1$ can be seen as a threshold for the proportion between female and male smolts in the population.

Reproduction strategies

The couples produce new female and male smolts according to two reproduction strategies ($N_r = 2$):

$$\{p_{k,j}^1\}_{(k,j) \in \mathbb{N}^2}, \quad p_{k,j}^1 = e^{-75} (40)^k (35)^j (k!j!)^{-1}$$

$$\{p_{k,j}^2\}_{(k,j) \in \mathbb{N}^2}, \quad p_{k,j}^2 = e^{-75} (35)^k (40)^j (k!j!)^{-1}$$

For $h = 1, 2$, $p_{k,j}^h$ is the probability that a couple gives birth to k female and j male smolts, when it is considered the h -th reproductive strategy.

$$\mu_1^1 = \sigma_{11}^1 = 40, \quad \mu_2^1 = \sigma_{22}^1 = 35, \quad \sigma_{12}^1 = 0.$$

$$\mu_1^2 = \sigma_{11}^2 = 35, \quad \mu_2^2 = \sigma_{22}^2 = 40, \quad \sigma_{12}^2 = 0.$$

Thus, according to the estimation provided by the Washington Department of Fisheries, in both distributions we obtain an average number of 75 smolts per couple.

In order to determine the corresponding offspring distribution, we consider the following function φ :

$$\varphi(f, m) = 1 \text{ if } f \leq m; \text{ or } 2 \text{ if } f > m.$$

When the number of females is less than or equal to the number of males then it is considered the offspring distribution 1, which promotes the production of females.

On the contrary, if the number of females is greater than the number of males, the offspring distribution 2, which promotes the production of males, is then considered.

Taking $c_1 = 0.04$ and $c_2 = 1.15$, we have simulated such a process up to the twentieth generation was reached.

Generation	f	m	$\psi(f, m)$	z	$\varphi(f, m)$
0	300	80	1	3	2
1	115	123	2	4	1
2	145	127	2	5	2
3	178	192	2	7	1
4	268	245	2	9	2
5	299	332	2	11	1
6	431	397	2	15	2
7	492	619	3	19	1
8	759	653	1	26	2
...
19	28187	31816	2	1127	1
20	45100	39532	2	1581	2

By applying Result 6 we have derived the Bayes estimators:

$$\widehat{\mu}_i^h = \widehat{\sigma}_{ii}^h = (\alpha_{ih}^* + 1)(\alpha_{0h}^*)^{-1}, \quad i, h = 1, 2.$$

$$\alpha_{0h}^* = \alpha_{0h} + \sum_{j=1}^n z_{j-1} I_{j,h}$$

$$\alpha_{1h}^* = \alpha_{1h} + \sum_{j=1}^n f_j I_{j,h}, \quad \alpha_{2h}^* = \alpha_{2h} + \sum_{j=1}^n m_j I_{j,h}$$

$$I_{j,h} = 1 \text{ if } \varphi(\mathbf{x}_{j-1}) = h; \text{ or } 0 \text{ if } \varphi(\mathbf{x}_{j-1}) \neq h$$

From the number of female and male smolts obtained in the simulated data and using the non-informative prior distribution with $\alpha_{0h} = 1$, $\alpha_{1h} = 100$ and $\alpha_{2h} = 100$, $h = 1, 2$, we have obtained that:

Parameter	μ_1^1	μ_2^1	μ_1^2	μ_2^2
True value	40	35	35	40
Estimation	39.987	34.992	35.083	39.993

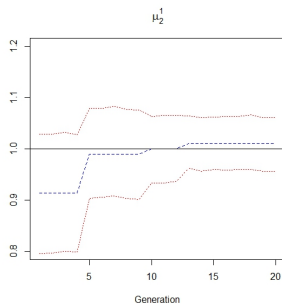
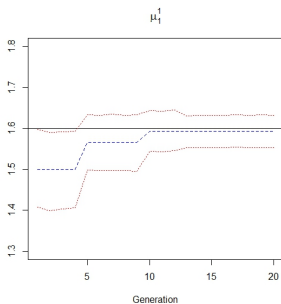
Therefore,

$$\max_{h=1,2} \{ \max_{i=1,2} \{ |\hat{\mu}_i^h - \mu_i^h| \} \} = 0.083$$

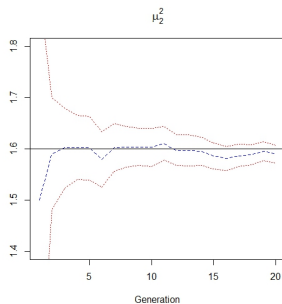
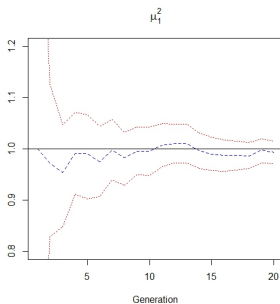
which provides information about the accuracy of the estimates.

The evolution of the estimations and the corresponding 95% highest posterior density credibility sets, through the successive generations, are showed in following Figures:

Offspring distribution 1: $\mu_1^1 = \sigma_{11}^1 = 40$, $\mu_2^1 = \sigma_{22}^1 = 35$, $\sigma_{12}^1 = 0$



Offspring distribution 2: $\mu_1^2 = \sigma_{11}^2 = 35$, $\mu_2^2 = \sigma_{22}^2 = 40$, $\sigma_{12}^2 = 0$



SOME QUESTIONS FOR RESEARCH

In the model we have considered in the last section, the mating function and the offspring distribution are determined, in each generation, through the deterministic functions ψ and φ respectively. A question for research is to consider random variables ψ and φ .

To study some probabilistic questions concerning the new two-sex models introduced. In particular, much of the research associated with two-sex branching processes has been concentrated on extinction probabilities. However, research on the time to extinction has not been sufficiently investigated. This topic needs to be addressed with greater intensity. Also, assuming the non-extinction of the population, it is necessary to investigate results about the limiting random variables.

To study some inferential questions concerning the new two-sex models introduced. In fact, it is necessary to propose for both cases, parametric and non-parametric, estimators for the main reproductive parameters involved in the probability model.

From an applied point of view, it is important to explore fields where the new two-sex processes investigated can be applied as appropriate mathematical model.

It can be stated that significant effort have been made regarding discrete-time two-sex branching processes. Similar effort should be made to develop a continuous-time two-sex branching process theory.

References

- R. Bartoszynski (1966).
Branching processes and the theory of epidemic.
Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability.
- N. Becker (1975).
The use of mathematical models in determining vaccination policies.
Bulletin of the International Statistical Institute, 46, 478-490.
- N. Becker (1976).
Estimation for discrete time branching processes with application to epidemics.
Biometrics, 33, 515-522.
- F.T. Bruss and M. Slatvchova-Bojkova (1999).
On waiting times to populate an environment and a question of statistical inference.
Journal of Applied Probability, 36, 261-267.

References

- J.P. Dion (1974).
Estimation of the mean and the initial probabilities of a branching process.
Journal of Applied Probability, 11, 176-192.
- I.A. Fleming (1996).
Reproductive strategies of Atlantic salmon: ecology and evolution.
Reviews in Fish Biology and Fisheries, 6, 379-416.
- T. Harris (1948).
Branching processes.
Annals of Mathematical Statistics, 19, 474-494.
- N.C. Hautekeete, Y., Piquot and H. Van Dijk (2001).
Investment in survival and reproduction along a semelparity-iteroparity gradient in the Beta species complex.
Journal of Evolutionary Biology, 14, 795-804.

References

- C.C. Heyde (1970).
A rate of convergence result for the supercritical Galton-Watson process.
Journal of Applied Probability, 7, 451-454.
- C.C. Heyde (1971).
Some central limit analogues for supercritical Galton-Watson processes.
Journal of Applied Probability, 8, 52-59.
- C.C. Heyde (1974).
On estimating the variance of the offspring distribution in a simple branching process.
Advances in Applied Probability, 3, 421-433.
- C.C. Heyde and E. Seneta (1972).
The simple branching process, a turning point test and a fundamental inequality. A historical note on I.J. Bienayme.
Biometrika, 59, 680-683.

References

- K. Kesten and B.P. Stigum (1966).
Additional limit theorems for indecomposable multidimensional Galton-Watson processes.
Annals of Mathematical Statistics, 37, 1463-1481.
- A.N. Kolmogorov (1938).
Zur Losung einer biologischen Aufgabe. K reseniyu odnoy biologiceskoy zadaci (To the solution of one biological problem).
IZv. NII Matem. i Meck. Tomskogo Un., 2, 1-6 (German), 7-12 (Russian).
- A.V. Nagaev (1967).
On estimating the expected number of direct descendent of a particle in a branching process.
Theory Probability and Applications, 12, 314-320.
- A.G. Pakes (1970).
An asymptotic result for a subcritical branching process with immigration.
Bulletin of the Australian Mathematics Society, 2, 223-228.

References

- A. G. Pakes (1975).
Non-parametric estimation in the Galton-Watson process.
Mathematical Biosciences, 26, 1-18.
- V. Pérez-Abreu (1987).
Los procesos ramificados como modelos para detectar brotes de epidemia de una enfermedad contagiosa: Aspectos estadísticos.
Memorias del II Foro de Estadística Aplicada, UNAM, Mexico.
- E. Seneta (1968).
On recent theorems concerning the supercritical Galton-Watson process.
Annals of Mathematical Statistics, 39, 1998-2012.
- A.M. Yaglom (1947).
Certain limit theorem of the theory of branching processes.
Doklady Akademii Nauk SSSR, 56, 793-795.

Acknowledgements

Ministerio Economía y Competitividad (Spain) and the FEDER,
grant: MTM2015-70522-P

Gobierno de Extremadura (Spain), grant: GR15105

National Fund for Scientific Research at the Ministry of Education
and Science (Bulgaria), grant: DFNI-I02/17

University of Extremadura (Spain)