BAYESIAN INFERENCE FOR BISEXUAL GALTON-WATSON PROCESSES

M. Molina, M. González and M. Mota

Departamento de Matemáticas. Universidad de Extremadura. 06071-Badajoz. Spain.

Key Words: Bivariate series power distribution; Parametric and nonparametric Bayesian inference.

ABSTRACT

In this work, an approach to the Bayesian estimation in a bisexual Galton-Watson process is considered. First we study an important parametric case assuming offspring distribution belonging to the bivariate series power family of distributions and then, we continue to investigate the nonparametric case. In both situations, Bayes estimators under weighted squared error loss function, for means, variances and covariance of the offspring distribution are obtained. For the superadditive case, the Bayes estimation of the asymptotic growth rate is also considered. Illustrative examples are given.

1 INTRODUCTION

The bisexual Galton-Watson process (BGWP) is a modified standard Galton-Watson process with sexual reproduction. It is a two type branching model with $F_n$ females and $M_n$ males in the nth generation, which form $Z_n = L(F_n, M_n)$ mating units according to certain mating function $L$. These mating units reproduce independently through the same offspring distribution for each generation. Introduced by Daley (1968), the BGWP has received attention in the scientific literature of the last years. The extinction problem for this model, has been investigated by Bruss (1984), Daley (1968), Daley et al. (1986) and Hull (1982, 1984) and its limit behaviour by Bagley (1986) and González & Molina (1996, 1997a, 1997b). However inferential questions have not been, until now, sufficiently considered. The only previous paper on inference about BGWP deals with the estimation of its parameters in a parametric context, through the traditional maximum likelihood procedure, see González & Pérez-Abreu (1991).

The purpose of this paper is to consider the estimation problem from a Bayesian outlook. Thus, we regard the parameters as random variables. We begin our study in Section 2 with a brief description on BGWPs where some notation and basic results are provided. In Section

1055

Copyright © 1998 by Marcel Dekker, Inc.
3, we consider an interesting parametric case assuming offspring distribution belonging to the bivariate series power family of distributions. For the main parameters Bayes estimators under weighted squared error loss function are obtained. Section 4 is devoted to look at the nonparametric case. As illustration, some simulated examples are given.

2 THE PROBABILITY MODEL

We consider a BGWP, i.e., a sequence \{(F_n, M_n), n = 0, 1, \ldots\} defined in the recursive manner

\[(F_{n+1}, M_{n+1}) = \sum_{i=1}^{Z_n} (f_{ni}, m_{ni}), Z_n = L(F_n, M_n), n = 0, 1, \ldots \quad (2.1)\]

with the empty sum defined to be (0, 0) and where \(f_{ni}\) (respectively \(m_{ni}\)) denotes the number of females (respectively males) produced by the \(i\)th mating unit in the \(n\)th generation, being \(\{(f_{ni}, m_{ni})\}\) i.i.d. non-negative integer valued bivariate random variables with a non-degenerate probability distribution \(\{p_{k_1 k_2}\}, p_{k_1 k_2} := \Pr[f_{01} = k_1, m_{01} = k_2]\), \((k_1, k_2) \in S\), where \(S := \{(k_1, k_2) \in \mathbb{N} \times \mathbb{N}: p_{k_1 k_2} > 0\}\). Let \(\mu_1 = E[f_{01}], \mu_2 = E[m_{01}], \sigma_1^2 = \text{Var}[f_{01}], \sigma_2^2 = \text{Var}[m_{01}]\) and \(\tau = \text{Cov}[f_{01}, m_{01}]\), the means, variances and covariance of the offspring distribution (supposed finites). The mating function \(L : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^+\) is assumed to be monotonic non-decreasing in each argument, integer-valued for integer-valued arguments and \(L(0, 0) = 0\). For simplicity, we consider that \(\Pr[(F_0, M_0) = (0, 0)] = 1\) with \(f_0, m_0\), given positive integers and suppose offspring distribution and mating function such that the condition \(\Pr[Z_n \rightarrow 0] + \Pr[Z_n \rightarrow \infty] = 1\) holds. In this situation it is easy to verify that the sequences \(\{(F_n, M_n)\}\) and \(\{Z_n\}\) are Markov chains with stationary transition probabilities.

Definition 2.1 (Bruss (1984)) For a BGWP, the concept of average reproduction mean per mating unit is defined as \(r_k = k^{-1}E[Z_{n+1}|Z_n = k], k = 1, 2, \ldots\).

Definition 2.2 (Hull (1982)) A BGWP is said to be superadditive if for all integer \(n \geq 2\), its mating function verify:

\[L\left(\sum_{i=1}^{n}(x_i, y_i)\right) \geq \sum_{i=1}^{n} L(x_i, y_i), x_i, y_i \in \mathbb{R}^+, i = 1, \ldots, n \quad (2.2)\]

Theorem 2.1 (Daley et al. (1986)) Let \(r = \lim_{k \to \infty} r_k\) and let \(q_i = \Pr[Z_n \rightarrow 0|Z_0 = i], i = 1, 2, \ldots, \) then for a superadditive BGWP, it is verified:

(i) \(r = \sup_{k > 0} r_k\)

(ii) \(q_i = 1 \text{ for all } i = 1, 2, \ldots \text{ if and only if } r \leq 1\)

Remark 2.1 Really, sufficiency of \(r \leq 1\) for almost certain extinction follows already from Theorem 1 of Bruss (1984), and this for arbitrary mating functions. He showed however that this condition cannot be improved for general mating functions. To obtain a necessary and sufficient condition it must turn to superadditivity. This is not a serious restriction, as was pointed out in Hull (1982), the vast majority of mating functions used in the literature on two-sex population models are superadditive. As noted in the above result, the asymptotic growth rate, \(r\), (for simplicity growth rate) is a fundamental parameter in the extinction problem. Moreover it plays a crucial role in the limit behaviour of the process, see Bagley (1986) and González & Molina (1996, 1997a, 1997b), and for these reasons, we shall consider its Bayesian estimation.
3 BAYES ESTIMATION RESULTS IN A PARAMETRIC CASE

We shall consider a model (2.1) with offspring distribution belonging to the power series family of distributions, i.e. \( p_{k_1,k_2} \) can be written in the form

\[
p_{k_1,k_2} = a_{k_1,k_2} \theta_1^{k_1} \theta_2^{k_2} (A(\theta_1, \theta_2))^{-1}, \quad (\theta_1, \theta_2) \in \Theta \subseteq \mathbb{R}^2
\]  

(3.1)

where \( a_{k_1,k_2} \) is a function of \( k_1, k_2 \) or constant,

\[
A(\theta_1, \theta_2) = \sum_{(k_1,k_2) \in S} a_{k_1,k_2} \theta_1^{k_1} \theta_2^{k_2} \quad \text{with} \quad a_{k_1,k_2} \theta_1^{k_1} \theta_2^{k_2} \geq 0
\]

and \( \Theta = \{(\theta_1, \theta_2) \in \mathbb{R}^2 : A(\theta_1, \theta_2) < \infty \} \).

This is an exponential family that includes many important distributions (e.g. bivariate Poisson, trinomial, negative trinomial \cdots ). It easily follows that for a probability distribution (3.1) the \((i,j)\)-cumulant, denoted \( Q_{ij}, i, j = 0, 1, \ldots, (i,j) \neq (0,0), \) can be obtained as

\[
Q_{ij} := \left. \frac{\partial^{i+j} \log(M(t_1,t_2))}{\partial t_1^i \partial t_2^j} \right|_{t_1=t_2=0} = \left( \theta_2 \frac{\partial}{\partial \theta_2} \right)^j \left( \theta_1 \frac{\partial}{\partial \theta_1} \right)^i \log(A(\theta_1, \theta_2))
\]  

(3.2)

where \( M(t_1,t_2) \) is the moment generating function of the offspring distribution and for \( l = 1, 2, (\theta_i \frac{\partial}{\partial \theta_i})^h \) denotes the operator \( (\theta_i \frac{\partial}{\partial \theta_i}) \) applied \( h \) times.

In particular, from (3.2) it is deduced that

\[
\mu_i = \mu_i(\theta_1, \theta_2) = \theta_i \frac{\partial}{\partial \theta_i} \log(A(\theta_1, \theta_2)) \quad , \quad i = 1, 2
\]

\[
\sigma_i^2 = \sigma_i^2(\theta_1, \theta_2) = \theta_i \frac{\partial}{\partial \theta_i} [\mu_i(\theta_1, \theta_2)] \quad , \quad i = 1, 2
\]

\[
\tau = \tau(\theta_1, \theta_2) = \theta_1 \frac{\partial}{\partial \theta_1} [\mu_2(\theta_1, \theta_2)] = \theta_2 \frac{\partial}{\partial \theta_2} [\mu_1(\theta_1, \theta_2)]
\]  

(3.3)

and it can be shown (Khatri (1959)) that (3.1) is determined uniquely from the moments given in (3.3).

Using that \( \{F_n, M_n\} \) is a Markov chain with stationary transition probabilities, the likelihood based on observing \((f_0,m_0), \ldots, (f_n,m_n)\) namely the numbers of females and males per generation until the \( n \)th generation, verify

\[
\ell[\theta_1, \theta_2] = \prod_{j=0}^{n-1} \Pr[(F_{j+1}, M_{j+1}) = (f_{j+1}, m_{j+1}) | (F_j, M_j) = (f_j, m_j)] = \prod_{j=0}^{n-1} \Pr \left[ \sum_{i=1}^{z_j} (f_{ji}, m_{ji}) = (f_{j+1}, m_{j+1}) \right] \propto \theta_1^{\sum_{i=1}^{z_j} f_{ji}} \theta_2^{\sum_{i=1}^{z_j} m_{ji}} [A(\theta_1, \theta_2)]^{-\sum_{i=0}^{n-1} z_i}
\]

(3.4)

with \( z_j = L(f_j, m_j) \) and where we used that

\[
\Pr \left[ \sum_{i=1}^{z_j} (f_{ji}, m_{ji}) = (f_{j+1}, m_{j+1}) \right] \propto \theta_1^{f_{j+1}} \theta_2^{m_{j+1}} [A(\theta_1, \theta_2)]^{-z_j}
\]

Therefore, a reasonable conjugate class of prior distributions will be

\[
f(\theta_1, \theta_2) = \theta_1^{\alpha_1} \theta_2^{\alpha_2} A(\theta_1, \theta_2)^{-\beta} (\beta(\alpha_1, \alpha_2, \beta))^{-1}
\]
with $\alpha_1, \alpha_2, \beta$ constants such that

$$\varphi(\alpha_1, \alpha_2, \beta) = \int_{\Theta} \theta_1^{\alpha_1} \theta_2^{\alpha_2} (A(\theta_1, \theta_2))^{-\beta} d\theta_1 d\theta_2 < \infty$$  \hspace{1cm} (3.5)$$

This class is flexible enough to describe different prior expectations.

From (3.4) and (3.5) it is derived that the posteriori distribution must be

$$f(\theta_1, \theta_2|\{ (f_0, m_0), \ldots, (f_n, m_n) \}) = \theta_1^{\alpha_1} \theta_2^{\alpha_2} (A(\theta_1, \theta_2))^{-\beta} (\varphi(\alpha_1^*, \alpha_2^*, \beta^*))^{-1}$$

where $\alpha_1^* = \alpha_1 + \sum_{i=1}^n f_i$, $\alpha_2^* = \alpha_2 + \sum_{i=1}^n m_i$, and $\beta^* = \beta + \sum_{i=0}^{n-1} z_i$.

In general, if for $n = 1, 2, \ldots$ we denote by $F_n^* = \sum_{i=0}^n F_i$, $M_n^* = \sum_{i=0}^n M_i$, $Z_n^* = \sum_{i=0}^n Z_i$ and $F_n = \sigma(\{ F_0, M_0, \ldots, (F_n, M_n) \})$ (i.e. the $\sigma$-field generated by $\{ F_0, M_0, \ldots, (F_n, M_n) \}$), then we have

$$f(\theta_1, \theta_2|F_n) = \theta_1^{\alpha_1^*} \theta_2^{\alpha_2^*} (A(\theta_1, \theta_2))^{-\beta^*} (\varphi(\alpha_1^*, \alpha_2^*, \beta^*))^{-1}$$ \hspace{1cm} (3.6)$$

with $\alpha_1^* = \alpha_1 + F_n^* - f_0$, $\alpha_2^* = \alpha_2 + M_n^* - m_0$ and $\beta^* = \beta + Z_{n-1}^*$.

Remark 3.1 From (3.6) can be obtained credibility sets, i.e. sets of probable values of $(\theta_1, \theta_2)$. An usual procedure is to determine a set $R(\lambda)$ such that $R(\lambda) = \{ (\theta_1, \theta_2) \in \Theta : f(\theta_1, \theta_2|F_n) \geq \lambda \}$ where $\lambda$ verifies

$$\int_{R(\lambda)} f(\theta_1, \theta_2|F_n) d\theta_1 d\theta_2 = 1 - \alpha$$

for a given credibility coefficient $1 - \alpha$, $0 < \alpha < 1$.

If we consider weighted squared error loss function, the Bayes estimator for a parameter $\Psi(\theta_1, \theta_2)$, denoted $\hat{\Psi}$, will be

$$\hat{\Psi} = \int_{\Theta} \Psi(\theta_1, \theta_2) f(\theta_1, \theta_2|F_n) d\theta_1 d\theta_2$$ \hspace{1cm} (3.7)$$

Therefore, using (3.3) it is readily proved the following result

**Proposition 3.1** For a BGWP (2.1) with offspring distribution (3.1), the Bayes estimators, under weighted squared error loss and considering the conjugate class (3.5), for the means, variances and covariance of the offspring distribution, are obtained through the following expressions:

$$\hat{\mu}_i = k \sum_{(k_1, k_2) \in S} k_i a_{k_1, k_2} \varphi(\alpha_1^* + k_1, \alpha_2^* + k_2, \beta^* + 1), i = 1, 2$$

$$\hat{\sigma}_i^2 = k \left[ \sum_{(k_1, k_2) \in S} k_i^2 a_{k_1, k_2} \varphi(\alpha_1^* + k_1, \alpha_2^* + k_2, \beta^* + 1) \right. - \sum_{(k_1, k_2) \in S} \sum_{(j_1, j_2) \in S} k_i j_i a_{k_1, k_2} a_{j_1, j_2} \varphi(\alpha_1^* + k_1 + j_1, \alpha_2^* + k_2 + j_2, \beta^* + 2) \right]$$

$$\hat{\tau} = k \sum_{(k_1, k_2) \in S} k_1 k_2 a_{k_1, k_2} \varphi(\alpha_1^* + k_1, \alpha_2^* + k_2, \beta^* + 1)$$
\[
- \sum_{(k_1, k_2) \in S} \sum_{(j_1, j_2) \in S} k_1 j_2 a_{k_1} a_{j_1} \varphi(a_1^* + k_1 + j_1, a_2^* + k_2 + j_2, \beta^* + 2) \frac{1}{\varphi(a_1^*, a_2^*, \beta^*)}
\]

where \( k = \frac{1}{\varphi(a_1^*, a_2^*, \beta^*)} \).

We now consider, for superadditive BGWPs, the Bayes estimation for the growth rate \( r \). It can be shown, see Daley et al. (1986), that \( r = \lim_{j \to \infty} Y_j \), where \( Y_j = \frac{L}{j} j^{-1} L(j\mu_1, j\mu_2) \), \( j = 1, 2, \ldots \).

**Proposition 3.2** For a superadditive BGWP (2.1) with offspring distribution (3.1), the Bayes estimator, under weighted squared error loss and considering the conjugate class (3.5), for the growth rate, verifies

(i) \( \bar{L} \leq \hat{r} \leq \liminf_{j \to \infty} \bar{Y}_j \), where \( \bar{L} = \mathbb{E}(L(\mu_1, \mu_2)|\mathcal{F}_n) \).

(ii) If the mating function is such that for \( j = 1, 2, \ldots \), \( L(jx, jy) = b_j L(x, y) \) holds, where \( b_j \) is a function of \( j \) such that \( \lim_{j \to \infty} b_j/j = b \) exists, then \( \hat{r} = b \bar{L} \).

(iii) If \( \{Y_j\} \) is non-decreasing or there exists a random variable \( Z \) such that for \( j = 1, 2, \ldots \), \( Y_j \leq Z \) almost surely, then \( \hat{r} = \lim_{j \to \infty} \bar{Y}_j \).

**Proof**

(i) From (2.2) it is deduced that \( Y_j \geq L(\mu_1, \mu_2) \), then

\[
\hat{r} = \mathbb{E}\left[ \lim_{j \to \infty} Y_j | \mathcal{F}_n \right] \geq \mathbb{E}[L(\mu_1, \mu_2)|\mathcal{F}_n]
\]

The other inequality follows from conditional Fatou's Lemma.

(ii) Obviously, if \( b = 0 \) or \( \infty \), we have that \( r = 0 \) or \( \infty \) respectively. Thus, we consider \( 0 < b < \infty \). Then

\[
\hat{r} = \mathbb{E}\left[ \lim_{j \to \infty} Y_j | \mathcal{F}_n \right] = \mathbb{E}\left[ \lim_{j \to \infty} j^{-1} b_j L(\mu_1, \mu_2)|\mathcal{F}_n \right] = b \bar{L}
\]

(iii) \( \hat{r} = \mathbb{E}\left[ \lim_{j \to \infty} Y_j | \mathcal{F}_n \right] = \lim_{j \to \infty} \mathbb{E}[Y_j | \mathcal{F}_n] = \lim_{j \to \infty} \bar{Y}_j \)

where, we have used the conditional monotone convergence theorem (if \( \{Y_j\} \) is non-decreasing) or the conditional dominated convergence theorem (if \( Y_j \leq Z \), \( j = 1, 2, \ldots \) a.s.).

**Remark 3.2** The hypothesis in proposition 3.2 (ii), is verified by many important mating functions (e.g. \( L(x, y) = x + y \), \( L(x, y) = \min\{x, y\} \), \( L(x, y) = x, \ldots \)). Another interesting mating functions fulfill some of the hypotheses in proposition 3.2 (iii), (e.g. \( L(x, y) = [(x + y)/2] \), \( L(x, y) = \lfloor 2xy/(x + y) \rfloor \) where we use \( \lfloor z \rfloor \) to denote the largest integer \( \leq z \).

**Remark 3.3** From (3.7) it is obtained that

\[
\hat{\psi} = \frac{\int_{\Theta} \Psi(\theta_1, \theta_2) \theta_1^{\alpha_1} \theta_2^{\alpha_2} (A(\theta_1, \theta_2))^{-\beta} d\theta_1 d\theta_2}{\int_{\Theta} \theta_1^{\alpha_1} \theta_2^{\alpha_2} (A(\theta_1, \theta_2))^{-\beta} d\theta_1 d\theta_2}
\]
When there exists computational problems, both integrals can be approximated through the Laplace’s method. A good description of this procedure can be seen, for example, in Bernardo & Smith (1994), Section 5.5.1.

**Example 3.1 (Negative trinomial)** We consider a BGWP (2.1) with offspring distribution

\[
p_{k_1k_2} = \frac{(N + k_1 + k_2 - 1)!}{k_1!k_2!(N - 1)!}\theta_1^{k_1}\theta_2^{k_2}(1 - \theta_1 - \theta_2)^N, \quad k_1, k_2 = 0, 1, \ldots \tag{3.8}
\]

where for \(i = 1, 2\), \(\theta_i > 0\) and \(\theta_1 + \theta_2 < 1\), with \(N\) a positive integer. This is a negative trinomial distribution, being the particular case of (3.1) where

\[
ak_{k_1k_2} = (N + k_1 + k_2 - 1)!/k_1!k_2!(N - 1)!
\]

and

\[
A(\theta_1, \theta_2) = (1 - \theta_1 - \theta_2)^{-N}
\]

From (3.3) it is derived that

\[
\mu_i = \mu_i(\theta_1, \theta_2) = N\theta_i(1 - \theta_1 - \theta_2)^{-1}, \quad i = 1, 2
\]

\[
\sigma_i^2 = \sigma_i^2(\theta_1, \theta_2) = N(\theta_i - \theta_1\theta_2)(1 - \theta_1 - \theta_2)^{-2}, \quad i = 1, 2
\]

\[
\tau = \tau(\theta_1, \theta_2) = -N\theta_1\theta_2(1 - \theta_1 - \theta_2)^{-2}
\]

(3.9)

and according to (3.5)

\[
\varphi(\alpha_1, \alpha_2, \beta) = \int_0^1 \theta_1^{\alpha_1} \left( \int_0^{1 - \theta_1} \theta_2^{\alpha_2}(1 - \theta_1 - \theta_2)^{N\beta} \, d\theta_2 \right) \, d\theta_1
\]

Substituting \(z = \theta_2/(1 - \theta_1)\) in the second integral we obtain

\[
\varphi(\alpha_1, \alpha_2, \beta) = \left( \int_0^1 \theta_1^{\alpha_1}(1 - \theta_1)^{\alpha_2+N\beta+1} \, d\theta_1 \right) \left( \int_0^1 z^{\alpha_2}(1 - z)^{N\beta} \, dz \right)
\]

\[
= \Gamma(\alpha_1 + 1)\Gamma(\alpha_2 + 1)\Gamma(N\beta + 1)/\Gamma(\alpha_1 + \alpha_2 + N\beta + 3)
\]

(3.10)

provided that \(\alpha_1 + 1 > 0\), \(i = 1, 2\), \(N\beta + 1 > 0\).

From (3.9) and taking into account (3.10), it is a matter of straightforward computations to prove that the corresponding Bayes estimators, under weighted error loss, are given by

\[
\hat{\mu}_i = (\alpha_i^* + 1)/\beta^*, \quad i = 1, 2
\]

\[
\hat{\sigma}_i^2 = (\alpha_i^* + 1)(\alpha_i^* + N\beta^* + 1)/\beta^*(N\beta^* - 1), \quad i = 1, 2
\]

\[
\hat{\tau} = -(\alpha_1^* + 1)(\alpha_2^* + 1)/\beta^*(N\beta^* - 1)
\]

(3.11)

As illustration, we have considered a BGWP with offspring distribution (3.8), (where \(\theta_1 = 0.05\), \(\theta_2 = 0.15\), \(N = 10\)) and governed by the mating function \(L(x, y) = x + y\). Under these conditions, we obtain that

\[
\mu_1 = 0.625 \quad \mu_2 = 1.875 \quad \sigma_1^2 = 0.66401 \quad \sigma_2^2 = 2.22656 \quad \tau = -0.11719 \quad r = 2.5
\]

Then, considering the prior distribution with \(\alpha_1 = 1.5\), \(\alpha_2 = -0.7\), \(\beta = 0.5\), and \((f_0, m_0) = (2, 3)\), 17 generations were simulated and from (3.11) the Bayes estimates were calculated. The table I exhibits the results obtained.
TABLE I. Simulated data and Bayes estimates

<table>
<thead>
<tr>
<th>n</th>
<th>F_n</th>
<th>M_n</th>
<th>Z_n</th>
<th>(\hat{\mu}_1)</th>
<th>(\hat{\mu}_2)</th>
<th>(\hat{\sigma}^2_1)</th>
<th>(\hat{\sigma}^2_2)</th>
<th>(\hat{\tau})</th>
<th>(\hat{\tilde{\tau}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>125</td>
<td>386</td>
<td>511</td>
<td>0.62293</td>
<td>1.91041</td>
<td>0.66193</td>
<td>2.27606</td>
<td>-1.1904</td>
<td>2.53333</td>
</tr>
<tr>
<td>6</td>
<td>311</td>
<td>965</td>
<td>1276</td>
<td>0.61424</td>
<td>1.89709</td>
<td>0.65205</td>
<td>2.25726</td>
<td>-1.1654</td>
<td>2.51134</td>
</tr>
<tr>
<td>7</td>
<td>777</td>
<td>2385</td>
<td>3162</td>
<td>0.61050</td>
<td>1.88024</td>
<td>0.64841</td>
<td>2.23388</td>
<td>-1.1490</td>
<td>2.49129</td>
</tr>
<tr>
<td>8</td>
<td>1950</td>
<td>5910</td>
<td>7860</td>
<td>0.61443</td>
<td>1.87355</td>
<td>0.65220</td>
<td>2.22462</td>
<td>-1.1512</td>
<td>2.48798</td>
</tr>
<tr>
<td>9</td>
<td>4969</td>
<td>14087</td>
<td>19656</td>
<td>0.60900</td>
<td>1.87058</td>
<td>0.66413</td>
<td>2.22050</td>
<td>-1.1692</td>
<td>2.49563</td>
</tr>
<tr>
<td>10</td>
<td>12360</td>
<td>36800</td>
<td>49160</td>
<td>0.62731</td>
<td>1.87155</td>
<td>0.66666</td>
<td>2.22183</td>
<td>-1.1740</td>
<td>2.49886</td>
</tr>
<tr>
<td>11</td>
<td>30941</td>
<td>92281</td>
<td>123222</td>
<td>0.62856</td>
<td>1.87491</td>
<td>0.66807</td>
<td>2.22645</td>
<td>-1.1785</td>
<td>2.50347</td>
</tr>
<tr>
<td>12</td>
<td>77433</td>
<td>230173</td>
<td>307606</td>
<td>0.62847</td>
<td>1.87073</td>
<td>0.66796</td>
<td>2.22070</td>
<td>-1.1757</td>
<td>2.49920</td>
</tr>
<tr>
<td>13</td>
<td>191812</td>
<td>572889</td>
<td>767701</td>
<td>0.62553</td>
<td>1.87159</td>
<td>0.66465</td>
<td>2.22188</td>
<td>-1.1707</td>
<td>2.49712</td>
</tr>
<tr>
<td>14</td>
<td>480207</td>
<td>1439562</td>
<td>1919769</td>
<td>0.62552</td>
<td>1.87373</td>
<td>0.66465</td>
<td>2.22482</td>
<td>-1.1721</td>
<td>2.49925</td>
</tr>
<tr>
<td>15</td>
<td>1201154</td>
<td>3598276</td>
<td>4799430</td>
<td>0.62561</td>
<td>1.87409</td>
<td>0.66475</td>
<td>2.22531</td>
<td>-1.1725</td>
<td>2.49970</td>
</tr>
<tr>
<td>16</td>
<td>2997164</td>
<td>8999008</td>
<td>11990072</td>
<td>0.62494</td>
<td>1.87463</td>
<td>0.66399</td>
<td>2.22266</td>
<td>-1.1715</td>
<td>2.49957</td>
</tr>
<tr>
<td>17</td>
<td>7497179</td>
<td>22494894</td>
<td>29992073</td>
<td>0.62496</td>
<td>1.87497</td>
<td>0.66401</td>
<td>2.22265</td>
<td>-1.1718</td>
<td>2.49992</td>
</tr>
</tbody>
</table>

The figures 1, 2 and 3 show the evolution of the estimates.

4 BAYES ESTIMATION RESULTS IN A NONPARAMETRIC CASE

We now assume a nonparametric situation considering a BGWP (2.1) with offspring distribution \(\{p_{k_1,k_2}\}\) such that its support \(S\) is a finite set. In this case, it is easy to prove that the likelihood based on the random sample \(\{(f_{j1},m_{j1}), i = 1, \ldots, Z_j, j = 0, \ldots, n\}\), i.e. the entire family tree up to the current nth generation, verifies

\[
\ell(p|\{(f_{j1},m_{j1}), i = 1, \ldots, Z_j, j = 0, \ldots, n\}) \propto \prod_{(k_1,k_2) \in S} p_{k_1,k_2}^{Y_n(k_1,k_2)}
\]

(4.1)

with \(p = (p_{k_1,k_2}, (k_1,k_2) \in S)^1\) and \(Y_n(k_1,k_2) = \sum_{j=0}^{n} Z_j(k_1,k_2)\), where

\(Z_j(k_1,k_2) = \#\{i \in \{1, \ldots, Z_j\}: (f_{j1}, m_{j1}) = (k_1,k_2)\}\).

Thus, from this multinomial form, the appropriate conjugate class of prior densities to consider will be

\[
f(p|\mathcal{F}_n) = d(\alpha) \prod_{(k_1,k_2) \in S} p_{k_1,k_2}^{\alpha_{k_1,k_2} - 1}
\]

(4.2)

where \(\alpha = (\alpha_{k_1,k_2}, (k_1,k_2) \in S), \alpha_{k_1,k_2} > 0\) and

\[
d(\alpha) = \Gamma(\alpha_*) \left( \prod_{(k_1,k_2) \in S} \Gamma(\alpha_{k_1,k_2}) \right)^{-1}
\]

with \(\alpha_* = \sum_{(k_1,k_2) \in S} \alpha_{k_1,k_2}\), i.e. \(p\) has a Dirichlet distribution with parameter vector \(\alpha\).

Then, if \(\mathcal{F}_n = \sigma(\{f_{j1}, m_{j1}), i = 1, \ldots, Z_j, j = 0, \ldots, n\}, n = 1, 2, \ldots\), taking into account (4.1) and (4.2) it is derived that

\[
f(p|\mathcal{F}_n) = d(\beta) \prod_{(k_1,k_2) \in S} p_{k_1,k_2}^{\beta_{k_1,k_2} - 1}
\]

(4.3)

\(^1\)In order to place the elements in the vector, given \((k_1,k_2),(k'_1,k'_2) \in S\), we consider \((k_1,k_2) < (k'_1,k'_2)\) if \(k_1 < k'_1\) or \(k_2 < k'_2\) (if \(k_1 = k'_1\)).
FIG. 1. Bayes estimates for $\mu_1$ and $\mu_2$

FIG. 2. Bayes estimates for $\sigma_1^2$ and $\sigma_3^2$
where $\beta = (\beta_{k_1,k_2}, (k_1, k_2) \in S)$, being $\beta_{k_1,k_2} = \alpha_{k_1,k_2} + Y_{n(k_1,k_2)}$.

From Dirichlet distributions theory it is deduced, according to (4.3), that for $I_{k_1,k_2} = 0, 1, \ldots$

$$E \left[ \prod_{(k_1, k_2) \in S} p_{k_1,k_2}^{l_{k_1,k_2}} \mid F'_n \right] = \frac{\Gamma(\beta_*)}{\Gamma(\beta_* + I_*)} \prod_{(k_1, k_2) \in S} \frac{\Gamma(\beta_{k_1,k_2} + I_{k_1,k_2})}{\Gamma(\beta_{k_1,k_2})}$$

with $I_* = \sum_{(k_1, k_2) \in S} I_{k_1,k_2}$

Hence that

$$E[p_{k_1,k_2} \mid F'_n] = \beta_{k_1,k_2} / \beta_*$$
$$E[p_{k_1,k_2}^2 \mid F'_n] = \beta_{k_1,k_2}(1 + \beta_{k_1,k_2}) / \beta_*(\beta_* + 1)$$
$$E[p_{k_1,k_2}p_{j_1,j_2} \mid F'_n] = \beta_{k_1,k_2}\beta_{j_1,j_2} / \beta_*(\beta_* + 1), \quad (k_1, k_2) \neq (j_1, j_2)$$

Moreover, it can be verified that $p_{k_1,k_2} \mid F'_n$ has a Beta distribution with parameters $\beta_{k_1,k_2}$ and $\beta_* - \beta_{k_1,k_2}$.

Consequently, if we consider weighted squared error loss function, the Bayes estimator for $p_{k_1,k_2}$ will be

$$\hat{p}_{k_1,k_2} = \beta_{k_1,k_2} / \beta_*, \quad (k_1, k_2) \in S$$

and the following result can be stated
Proposition 4.1

(i) The nonparametric Bayes estimators, under weighted squared error loss, for the means, variances and covariances of the offspring distribution, are given by:

\[
\hat{\mu}_i = \frac{1}{\beta_*} \sum_{k_i \in S_i} k_i \beta_k, \quad i = 1, 2
\]

\[
\hat{\sigma}_i^2 = \frac{1}{\beta_* + 1} \left( \sum_{k_i \in S_i} k_i^2 \beta_k - \beta_* \hat{\mu}_i^2 \right), \quad i = 1, 2
\]

\[
\hat{r} = \frac{1}{\beta_*(\beta_* + 1)} \left( \beta_* \sum_{(k_1, k_2) \in S_i} k_1 k_2 \beta_{k_1, k_2} - \sum_{(k_1, k_2) \in S_i} k_1 k_2 \beta_{k_1, k_2}^2 \right.
\]

\[
- \sum_{(k_1, k_2) \neq (j_1, j_2)} k_1 j_2 \beta_{k_1, k_2, j_1, j_2})
\]

(4.6)

where \( S_1 \) (respectively \( S_2 \)) denotes the support of \( f_{01} \) (respectively \( m_{01} \)) and \( \beta_{k, i} = \sum_{k_1(i)} \beta_{k_1, k_2} \), being \( \Phi(i) = 2 \) if \( i = 1 \) or 1 if \( i = 2 \).

(ii) If the BGWP is superadditive, then the Bayes estimator, under weighted squared error loss, for the growth rate, verifies:

\[
\hat{r} \geq \sum_{(k_1, k_2) \in S} \hat{L}_{k_1, k_2} \text{ where } \hat{L}_{k_1, k_2} = E[L(k_1 p_{k_1, k_2}, k_2 p_{k_1, k_2}) | \mathcal{F}_n']
\]

Proof

(i) From (4.4) and (4.5), we have

\[
\hat{\mu}_i = \sum_{k_i \in S_i} \left( \sum_{k_{i(i)} \in S_{i(i)}} \frac{\beta_{k_1, k_2}}{\beta_*} \right) = \frac{1}{\beta_*} \sum_{k_i \in S_i} k_i \beta_k, \quad i = 1, 2
\]

Similarly

\[
\hat{\sigma}_i^2 = \sum_{(k_1, k_2) \in S} k_i^2 E[p_{k_1, k_2} | \mathcal{F}_n'] - \sum_{(k_1, k_2) \in S} k_i^2 E[p_{k_1, k_2}^2 | \mathcal{F}_n']
\]

\[
- \sum_{(k_1, k_2) \neq (j_1, j_2)} k_1 j_2 E[p_{k_1, k_2} p_{j_1, j_2} | \mathcal{F}_n']
\]

\[
= \frac{1}{\beta_*(\beta_* + 1)} \left( \beta_* \sum_{k_i \in S_i} k_i^2 \beta_k - \left( \sum_{k_i \in S_i} k_i \beta_k \right)^2 \right)
\]

\[
= \frac{1}{\beta_* + 1} \left( \sum_{k_i \in S_i} k_i^2 \beta_k - \beta_* \hat{\mu}_i^2 \right) \quad i = 1, 2.
\]

and

\[
\hat{r} = \sum_{(k_1, k_2) \in S} k_1 k_2 E[p_{k_1, k_2} | \mathcal{F}_n'] - \sum_{(k_1, k_2) \in S} k_1 k_2 E[p_{k_1, k_2}^2 | \mathcal{F}_n']
\]

\[
- \sum_{(k_1, k_2) \neq (j_1, j_2)} k_1 j_2 E[p_{k_1, k_2} p_{j_1, j_2} | \mathcal{F}_n']
\]
TABLE II. Bayes estimates for $p_{k_1 k_2}$

<table>
<thead>
<tr>
<th>$(k_1, k_2)$</th>
<th>$n = 10$</th>
<th>$n = 15$</th>
<th>$n = 20$</th>
<th>$n = 27$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0,0)</td>
<td>0.003112</td>
<td>0.001125</td>
<td>0.001029</td>
<td>0.001037</td>
</tr>
<tr>
<td>(0,1)</td>
<td>0.013486</td>
<td>0.013875</td>
<td>0.011387</td>
<td>0.011862</td>
</tr>
<tr>
<td>(0,2)</td>
<td>0.044606</td>
<td>0.049625</td>
<td>0.045469</td>
<td>0.048202</td>
</tr>
<tr>
<td>(0,3)</td>
<td>0.054280</td>
<td>0.068375</td>
<td>0.064410</td>
<td>0.063894</td>
</tr>
<tr>
<td>(1,0)</td>
<td>0.015560</td>
<td>0.012875</td>
<td>0.015003</td>
<td>0.014772</td>
</tr>
<tr>
<td>(1,1)</td>
<td>0.121369</td>
<td>0.112625</td>
<td>0.119683</td>
<td>0.118806</td>
</tr>
<tr>
<td>(1,2)</td>
<td>0.239627</td>
<td>0.230125</td>
<td>0.237729</td>
<td>0.239246</td>
</tr>
<tr>
<td>(2,0)</td>
<td>0.086000</td>
<td>0.084125</td>
<td>0.077777</td>
<td>0.075882</td>
</tr>
<tr>
<td>(2,1)</td>
<td>0.295643</td>
<td>0.299375</td>
<td>0.306056</td>
<td>0.301534</td>
</tr>
<tr>
<td>(3,0)</td>
<td>0.125519</td>
<td>0.127875</td>
<td>0.126366</td>
<td>0.124775</td>
</tr>
</tbody>
</table>

TABLE III. Simulated data and Bayes estimates

<table>
<thead>
<tr>
<th>n</th>
<th>$F_n$</th>
<th>$M_n$</th>
<th>$Z_n$</th>
<th>$\hat{\mu}_1$</th>
<th>$\hat{\mu}_2$</th>
<th>$\hat{\sigma}_1^2$</th>
<th>$\hat{\sigma}_2^2$</th>
<th>$\hat{r}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>23</td>
<td>24</td>
<td>23</td>
<td>1.585</td>
<td>1.244</td>
<td>1.587</td>
<td>1.191</td>
<td>-3.103</td>
</tr>
<tr>
<td>10</td>
<td>157</td>
<td>132</td>
<td>157</td>
<td>1.486</td>
<td>1.179</td>
<td>0.789</td>
<td>0.744</td>
<td>-0.904</td>
</tr>
<tr>
<td>15</td>
<td>1369</td>
<td>1058</td>
<td>1369</td>
<td>1.521</td>
<td>1.179</td>
<td>0.786</td>
<td>0.746</td>
<td>-0.669</td>
</tr>
<tr>
<td>20</td>
<td>10718</td>
<td>8339</td>
<td>10718</td>
<td>1.514</td>
<td>1.185</td>
<td>0.750</td>
<td>0.723</td>
<td>-0.607</td>
</tr>
<tr>
<td>21</td>
<td>16238</td>
<td>12760</td>
<td>16238</td>
<td>1.514</td>
<td>1.187</td>
<td>0.748</td>
<td>0.722</td>
<td>-0.604</td>
</tr>
<tr>
<td>22</td>
<td>24345</td>
<td>19526</td>
<td>24345</td>
<td>1.509</td>
<td>1.192</td>
<td>0.747</td>
<td>0.721</td>
<td>-0.602</td>
</tr>
<tr>
<td>23</td>
<td>36536</td>
<td>29224</td>
<td>36536</td>
<td>1.506</td>
<td>1.195</td>
<td>0.749</td>
<td>0.720</td>
<td>-0.601</td>
</tr>
<tr>
<td>24</td>
<td>54966</td>
<td>43773</td>
<td>54966</td>
<td>1.506</td>
<td>1.196</td>
<td>0.751</td>
<td>0.722</td>
<td>-0.603</td>
</tr>
<tr>
<td>25</td>
<td>82422</td>
<td>66165</td>
<td>82422</td>
<td>1.504</td>
<td>1.199</td>
<td>0.750</td>
<td>0.721</td>
<td>-0.601</td>
</tr>
<tr>
<td>26</td>
<td>123791</td>
<td>98508</td>
<td>123791</td>
<td>1.503</td>
<td>1.198</td>
<td>0.750</td>
<td>0.721</td>
<td>-0.600</td>
</tr>
<tr>
<td>27</td>
<td>185662</td>
<td>148725</td>
<td>185662</td>
<td>1.502</td>
<td>1.199</td>
<td>0.750</td>
<td>0.720</td>
<td>-0.601</td>
</tr>
</tbody>
</table>

\[
\hat{r} = \frac{1}{\beta_n(\beta_n + 1)} \left( \beta_n \sum_{(k_1, k_2) \in S} k_1 k_2 \beta_{k_1 k_2} - \sum_{(k_1, k_2) \in S} k_1 k_2 \beta_{k_1 k_2}^2 \right)
\]

(ii) From (2.2), we obtain

\[
\hat{r} = E \left[ \lim_{j \to \infty} j^{-1} L \left( j \sum_{(k_1, k_2) \in S} (k_1 p_{k_1 k_2}, k_2 p_{k_1 k_2}) \right) \right]
\]

\[
\geq \sum_{(k_1, k_2) \in S} E \left[ L(k_1 p_{k_1 k_2}, k_2 p_{k_1 k_2}) \right] \hat{F}_n = \sum_{(k_1, k_2) \in S} \hat{L}_{k_1 k_2}
\]

and the result is proved.
FIG. 4. Bayes estimates for \( \mu_1 \) and \( \mu_2 \).

FIG. 5. Bayes estimates for \( \sigma_1^2 \) and \( \sigma_2^2 \).
FIG. 6. Bayes estimates for $\tau$.

FIG. 7. Empirical and approximate distributions of $\mu_1$ (left) and $\mu_3$ (right). (Shapiro-Wilk's test for normality: $p=0.7328$, $p=0.9811$, respectively).
Example 4.1 As illustration, suppose a BGWP with the following offspring distribution

\[
\begin{array}{ccccccccccc}
(k_1, k_2) & (0.0) & (0.1) & (0.2) & (0.3) & (1.0) & (1.1) & (1.2) & (2.0) & (2.1) & (3.0) \\
 p_{k_1+k_2} & 0.001 & 0.012 & 0.048 & 0.064 & 0.015 & 0.12 & 0.24 & 0.075 & 0.3 & 0.125 \\
\end{array}
\]

and governed by the historical mating function \(L(x, y) = x \min\{1, y\}\). In this situation, it is readily obtained that

\[
\mu_1 = 1.5 \quad \mu_2 = 1.2 \quad \sigma_1^2 = 0.75 \quad \sigma_2^2 = 0.72 \quad \tau = -0.6
\]

Moreover, for the mating function considered

\[
\tau = \lim_{j \to \infty} j^{-1} L(j \mu_1, j \mu_2) = \lim_{j \to \infty} \mu_1 \min\{1, j \mu_2\} = \mu_1
\]

so \(\tau = 1.5\).
Taking into account the paper of Berger & Bernardo (1992), we shall consider as prior density of reference the Dirichlet distribution with parameters $\alpha_{1,2} = 1/2$, $(k_1, k_2) \in S$. Under those conditions, 27 generations have been simulated and from (4.5) and (4.6) the Bayes estimates were calculated. The tables II and III and the figures 4, 5 and 6 exhibit the results obtained.

From (4.3) the posteriori distribution, when $n = 27$, will be a Dirichlet distribution with parameters

$$
\beta_{00} = 383.5, \beta_{01} = 4387.5, \beta_{02} = 17828.5, \beta_{03} = 23632.5, \beta_{10} = 5463.5, \beta_{11} = 43942.5, \beta_{12} = 88469.5, \beta_{20} = 28066.5, \beta_{21} = 111524.5, \beta_{22} = 46150.5.
$$

The figures 7, 8 and 9 show the empirical and approximate distributions of $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ and $\tau$, obtained by the simulation of 2000 vectors from such Dirichlet distribution.

**Remark 4.1** For simplicity, in this section, we have considered finite support. Making use of the Dirichlet processes theory, see Ferguson (1973), the Bayes estimators obtained could be adapted to the infinite case.

**ACKNOWLEDGEMENTS**

The authors would like to give special thanks to Horacio M. González Velasco (Departamento de Electrónica e Ingeniería Electromecánica. Universidad de Extremadura) for developing the computer programs, which have been used in the simulation process. We also wish to thank the referees for their corrections and helpful suggestions for improvement.

**BIBLIOGRAPHY**


Received February, 1997; Revised November, 1997.