

# Chapter 1

## Bayesian analysis for controlled branching processes

Miguel González, Cristina Gutiérrez, Rodrigo Martínez, Carmen Minuesa, and Inés M. del Puerto

**Abstract** A controlled branching process is a stochastic model that is well suited to describing the probabilistic evolution of populations in which, for various reasons of an environmental, social, or other nature, there is a mechanism that establishes the number of progenitors who take part in each generation. For this model, a Bayesian analysis is described, considering a non-parametric offspring distribution and control distributions belonging to the power series family that depend on a single parameter termed the control parameter. Inferences on the offspring distribution, on the offspring mean, and on the control parameter (or on its parametrization as the migration parameter) are considered within two sampling schemes: first, the classical branching theory scheme based on the observation of the entire family tree; and second, the more realistic situation in which only the generation-by-generation population size is observed. In this latter case, the Dirichlet process and the Gibbs sampler are used to estimate the posterior density of the main parameters of interest. Inference on posterior predictive distributions for as-yet unobserved generations is also considered. Monte Carlo sampling based and analytical approximations are discussed. The results are illustrated with simulated data.

**Mathematics Subject Classification (2000):** 60J80

**Key words:** controlled branching process, Bayesian inference, Dirichlet process, Gibbs sampler, Monte Carlo method, non-parametric statistics, power series family of distributions, statistical computing.

---

M. González, e-mail: [mvelasco@unex.es](mailto:mvelasco@unex.es) · C. Gutiérrez, e-mail: [cgutierrez@unex.es](mailto:cgutierrez@unex.es) · R. Martínez, e-mail: [rmartinez@unex.es](mailto:rmartinez@unex.es) · C. Minuesa, e-mail: [cminuesaa@unex.es](mailto:cminuesaa@unex.es) · I. del Puerto, e-mail: [idelpuerto@unex.es](mailto:idelpuerto@unex.es), Department of Mathematics, Faculty of Sciences, University of Extremadura, Avda. Elvas, 06006 Badajoz, Spain

This is a post-peer-review, pre-copyedit version of the work published in *Branching Processes and their Applications*, volume 219 of *Lecture Notes in Statistics-Proceedings*, Springer-Verlag, pp. 185-205. The final authenticated version is available online at: [https://doi.org/10.1007/978-3-319-31641-3\\_11](https://doi.org/10.1007/978-3-319-31641-3_11)

This version of the work is subject to Springer Nature's AM terms of use, see: <https://www.springernature.com/gp/open-science/policies/accepted-manuscript-terms>.

## 1.1 Introduction

The branching model considered in the present work is the controlled branching process. This model is a generalization of the standard Bienaymé-Galton-Watson (BGW) branching process, and, in the terminology of population dynamics, is used to describe the evolution of populations in which a control of the population size at each generation is needed. This control consists of determining mathematically the number of individuals with reproductive capacity at each generation through a random process. In practice, this branching model can describe reasonably well the probabilistic evolution of populations in which, for various reasons of an environmental, social, or other nature, there is a mechanism that establishes the number of progenitors which take part in each generation. For example, in an ecological context, one can think of an invasive animal species that is widely recognized as a threat to native ecosystems, but there is disagreement about plans to eradicate it, i.e., while the presence of the species is appreciated by a part of the society, if its numbers are left uncontrolled it is known to be very harmful to native ecosystems. In such a case, it is better to control the population to keep it within admissible limits even though this might mean periods when animals have to be culled. Another practical situation that can be modelled by this kind of process is the evolution of an animal population that is threatened by the existence of predators. In each generation, the survival of each animal (and therefore the possibility of giving new births) will be strongly affected by this factor, making the introduction of a random mechanism necessary to model the evolution of this kind of population.

Mathematically, a controlled branching process with random control function (CBP) is a discrete-time stochastic growth population model  $\{Z_n\}_{n \geq 0}$  defined recursively as

$$Z_0 = N \in \mathbb{N}, \quad Z_{n+1} = \sum_{j=1}^{\phi_n(Z_n)} X_{nj}, \quad n \geq 0, \quad (1.1)$$

where  $\{X_{nj} : n = 0, 1, \dots; j = 1, 2, \dots\}$  and  $\{\phi_n(k) : n, k = 0, 1, \dots\}$  are two independent families of non-negative integer-valued random variables. Moreover,  $X_{nj}$ ,  $n = 0, 1, \dots; j = 1, 2, \dots$ , are independent and identically distributed (i.i.d.) random variables and, for each  $n = 0, 1, \dots$ ,  $\{\phi_n(k)\}_{k \geq 0}$ , are independent stochastic processes with equal one-dimensional probability distributions. The empty sum in (1.1) is defined to be 0. Let  $\{p_k : k \geq 0\}$  denote the common probability distribution of the random variables  $X_{nj}$ , i.e.,  $p_k = P(X_{nj} = k)$ ,  $k \geq 0$ , and  $m = E[X_{nj}]$  (assumed finite).

Intuitively,  $Z_n$  denotes the number of individuals in generation  $n$ , and  $X_{ni}$  the number of offspring of the  $i$ th individual in generation  $n$ . Thus, the probability law  $\{p_k : k \geq 0\}$  is termed the offspring distribution, and  $m$  the offspring mean. The control variables  $\phi_n(\cdot)$  could be seen as a random mechanism determining the individual migration process in each generation depending on its population size.

The probabilistic theory of CBPs, in particular the problem of their extinction and their limiting behaviour, has been extensively investigated (see for example [1], [11] and references therein, and [19]). The presence of the control mechanism com-

plicates the study of this kind of process. Nevertheless, it allows one to model a much greater variety of behaviours than the BGW branching process. One of its features that is important for applications is that it overcomes one of the main deficiencies of the BGW process, which is the possibility of an unbounded population growth with a supercritical offspring law (i.e.,  $m > 1$ ). Thus, in [10], it was proved that a CBP with offspring mean greater than unity (and, of course, other regularity conditions governing the control mechanism) can die out with probability one. The offspring mean continues to play a key role in the probabilistic evolution of a CBP, as the above-cited papers show. Indeed, it was established that the asymptotic mean growth rate of the process denoted by  $\tau$  and defined by  $\tau = \lim_{k \rightarrow \infty} k^{-1} E[Z_{n+1} | Z_n = k] = \lim_{k \rightarrow \infty} k^{-1} \varepsilon(k)m$ , with  $\varepsilon(k) = E[\phi_n(k)]$  (whenever it exists) is the threshold parameter that determines the behaviour of a CBP in relation to its extinction. Hence the importance of making inferences on the offspring mean and on the asymptotic mean growth rate. However, there have as yet been few papers devoted to this topic. Inferential studies from a frequentist standpoint may be found in [7], [8], [9], and [20]. A first approach from a Bayesian standpoint was considered in [14] in a parametric context, and in [5] for the particular case of a deterministic control function.

The present study is a continuation of this line of research by developing the inferential theory in a non-parametric framework for the offspring law and in a parametric setting for the control distributions, depending on a single parameter termed the control parameter. Notice that, in relation to [5], random control rather than deterministic one introduces much more uncertainty in the model becoming its behaviour richer, but adding a considerable difficulty on the estimation of its parameters. Moreover, another important novelty is that we assume an upper bound of the support of the offspring law is unknown. We model this uncertainty considering the support potentially infinite. Although this could seem a strong condition, the gain in flexibility makes the model more attractive and realistic. In this case, to deal with the inference procedure we shall use as prior a Dirichlet process on the space of the nonnegative integers. Consequently, in this paper, we address the inference of the control parameter, of the offspring distribution, and of the offspring mean, as well as the asymptotic mean growth rate and the prediction of future sizes of the population. To this end, Section 1.2 begins by assuming that the entire family tree up to some given generation can be observed. A Dirichlet process is introduced to model the prior distribution of the offspring law, avoiding assumptions on the cardinality of its support. However, actually, in most populations, it is not possible to observe these data, and only the population size at each generation can be recorded. To deal with the Bayesian inference in this case, a Markov chain Monte Carlo (MCMC) method is used, in particular, the Gibbs sampler algorithm, to approximate the posterior distribution of parameters of interest. The present implementation of this algorithm generalizes the results in [5] and [6] and represents the Bayesian analogy of the given in [9].

Section 1.2 also deals with the problem of approximating the predictive posterior distributions. As illustration, in Section 1.3 a simulated example is presented. Finally, some concluding remarks are given in Section 1.4.

## 1.2 Bayesian Analysis

For the purpose of this paper, we consider a CBP with an offspring distribution  $p = \{p_k : k \geq 0\}$ , without assuming any knowledge about the cardinality of its support. With respect to the random control mechanism, note that one has different probability distributions for each population size  $k \geq 0$ , that corresponding to  $\phi_n(k)$ . Consequently, from a finite sample, it is not possible to deal with the inference problems arising from this model (at least for the control distributions) without assuming that there exists some stable structure. We therefore consider a parametric scheme for the control process. In particular, we take a CBP with control distributions belonging to the power series family. Formally, for each  $k \geq 0$ ,

$$P(\phi_n(k) = j) = a_k(j)\theta^j/A_k(\theta), \quad j = 0, 1, \dots; \theta \in \Theta_k, \quad (1.2)$$

with  $a_k(j)$  known non-negative values,  $A_k(\theta) = \sum_{j=0}^{\infty} a_k(j)\theta^j$ , and  $\Theta_k = \{\theta > 0 : 0 < A_k(\theta) < \infty\}$  being an open subset of  $\mathbb{R}$ . We also assume that the sets  $\Theta_k$  are independent of  $k$ . Hence, we shall henceforth drop the index  $k$  from  $\Theta_k$ , the control parameter space. This implies a certain stability in the probability laws of the control variables  $\phi_n(k)$ , for  $k \geq 0$ , depending on a single parameter  $\theta$ , the control parameter, and on the size of the population,  $k$ . Let us write  $\epsilon(k, \theta) = E[\phi_n(k)]$ , and assume the following regularity condition:

$$\prod_{k \in B} A_k(\theta) = A_{\sum_{k \in B} k}(\theta), \quad \text{for every } B \subseteq \mathbb{N}, \theta \in \Theta. \quad (1.3)$$

Condition (1.3), satisfied by a wide family of probability distributions, is a technical hypothesis that allows the theory of conjugate families to be made use of in the Bayesian analysis to be developed below. Moreover (1.3) implies that  $A_k(\theta) = (A_1(\theta))^k$ , so that

$$\epsilon(k, \theta) = \theta A'_k(\theta)(A_k(\theta))^{-1} = k\theta(A_1(\theta))^{k-1}A'_1(\theta)(A_1(\theta))^{-k} = k\epsilon(1, \theta).$$

This allows the family of distributions verifying (1.3) to be reparametrized by  $\mu = \mu(\theta) = \epsilon(1, \theta)$ . It also guarantees the existence of the limit  $\tau$  and its parametrization in terms of the two mean-value parameters, the offspring mean  $m$  and  $\mu$ , with  $\tau = m\mu$ .

Unlike the parameter  $\theta$ ,  $\mu$  has the same interpretation for all the power series families of distributions satisfying (1.3). We refer to  $\mu$  as the migration parameter since, whenever  $\mu < 1$ , one can use such control distributions to model processes with expected emigration, whereas  $\mu > 1$  can model processes with expected immigration. For  $\mu = 1$  neither emigration nor immigration is expected.

*Remark 1.* Some interesting particular cases of control distributions verifying (1.2) and (1.3) are the following:

- a) Consider for each  $k$  that  $\phi_n(k)$  follows a binomial distribution with parameters  $k$  and  $0 < q < 1$ . In such a case  $\theta = q(1 - q)^{-1}$ . It is easy to see that conditions (1.2) and (1.3) hold, and in particular that  $\mu = \theta(1 + \theta)^{-1} (= q)$ . From a practical

standpoint, this control mechanism could be reasonable to model situations in which, in each generation, each individual can be removed from the population with probability  $1 - q$ , not participating in its subsequent evolution, and can survive and give birth to offspring in the next generation with probability  $q$ . The value of  $q$  is considered to be independent of the total population size in each generation. These control functions always take into account the possibility of an emigration phenomenon in each generation ( $\mu < 1$ ). A CBP with this family of control distributions can be useful to model predator effects in a population.

- b) Consider for each  $k$  that  $\phi_n(k)$  has a Poisson distribution with parameter  $k\theta$ . Again conditions (1.2) and (1.3) hold, and one deduces that  $\mu = \theta$ . Depending on the value of  $\mu$ , a CBP with these control functions can model either expected immigration or expected emigration processes.
- c) Consider for each  $k$  that  $\phi_n(k)$  has a negative binomial distribution with parameters  $k$  and  $q$ , with  $0 < q < 1$ . Now  $\theta = 1 - q$ . It is easy to see again that conditions (1.2) and (1.3) hold. In this case  $\mu = \theta(1 - \theta)^{-1}$ . The associated CBP can also be used to model different migratory movements.

To infer the posterior distribution of the main parameters of interest, we consider initially that the entire family tree up to the current  $n$ th generation can be observed. After studying this case, we consider a more realistic requirement that only the total population size at each generation can be sampled. To deal with this second situation, we shall need the Gibbs sampler.

### 1.2.1 Analysis based on the entire family tree

We consider that the entire family tree up to the current  $n$ th generation can be observed, i.e.,  $\{X_{lj} : j = 1, \dots, \phi_l(Z_l); l = 0, 1, \dots, n - 1\}$ , or at least the variables

$$Z_{n,k}^* = \sum_{l=0}^{n-1} Z_l(k), \text{ where } Z_l(k) = \sum_{j=1}^{\phi_l(Z_l)} I_{\{X_{lj}=k\}}, k \geq 0,$$

with  $I_A$  standing for the indicator function of the set  $A$ . Intuitively,  $Z_l(k)$  represents the number of progenitors at the  $l$ th generation with exactly  $k$  offspring, and therefore  $Z_{n,k}^*$  is the accumulated number up to generation  $n$  of progenitors that give rise to exactly  $k$  offspring. Let us write  $\mathcal{Z}_n^* = \{Z_l(k), k \geq 0, l = 0, 1, \dots, n - 1\}$ , and introduce the following variables:

$$Y_n = \sum_{l=0}^{n-1} Z_l \quad \text{and} \quad Y_n^* = \sum_{l=0}^{n-1} \phi_l(Z_l),$$

i.e.,  $Y_n$  and  $Y_n^*$  represent, respectively, the total number of individuals and progenitors in the population up to the  $(n - 1)$ th generation. Using (1.2) and (1.3), one can deduce that the likelihood based on the sample  $\mathcal{Z}_n^*$  verifies

$$f(\mathcal{Z}_n^* | p, \theta) \propto \prod_{k \geq 0}^{Z_{n,k}^*} p_k^{Z_{n,k}^*} \theta^{Y_n^*} / A_{Y_n}(\theta). \quad (1.4)$$

Hence, since no restriction has been imposed on the cardinality of support of the reproduction law, and the offspring and control distributions are independent, an appropriate conjugate class of prior distributions for  $(p, \theta)$  is  $\pi(p, \theta) = \pi(p)\pi(\theta)$ , with  $\pi(p)$  being the distribution corresponding to

$$p \sim \text{DP}(p(0), \alpha), \quad (1.5)$$

where DP denotes the Dirichlet process, with  $p(0) = \{p_k(0), k \geq 0\}$  being the base measure and  $\alpha$  the concentration parameter,  $\alpha > 0$ , and  $\pi(\theta)$  the distribution given by the density

$$\varphi(a, b)^{-1} \theta^a / A_b(\theta), \quad (1.6)$$

with

$$\varphi(a, b) = \int_{\Theta} \theta^a / A_b(\theta) d\theta,$$

where  $a, b \geq 0$ .

Then, using (1.4)–(1.6), one has that the posterior distribution

$$\pi(p, \theta | \mathcal{Z}_n^*) \propto \pi(p | \mathcal{Z}_n^*) \pi(\theta | \mathcal{Z}_n^*), \quad (1.7)$$

with  $\pi(p | \mathcal{Z}_n^*)$  being the distribution corresponding to

$$p | \mathcal{Z}_n^* \sim \text{DP} \left( \frac{\alpha}{\alpha + Y_n^*} p(0) + \frac{1}{\alpha + Y_n^*} \sum_{k \geq 0}^{Z_{n,k}^*} \delta_k, \alpha + Y_n^* \right),$$

with  $\delta_k$  a Dirac delta at  $k$ ,  $k \geq 0$ , and

$$\pi(\theta | \mathcal{Z}_n^*) = \varphi(a + Y_n^*, b + Y_n)^{-1} \theta^{a+Y_n^*} / A_{b+Y_n}(\theta).$$

From (1.7), using Dirichlet process properties and considering the squared error loss function, it follows straightforwardly that the Bayes estimator for the offspring distribution and  $\theta$  are, respectively:

$$\widehat{p}_k = (\alpha p_k(0) + Z_{n,k}^*) / (\alpha + Y_n^*), \quad k \geq 0,$$

and

$$\widehat{\theta} = \varphi(a + Y_n^* + 1, b + Y_n) / \varphi(a + Y_n^*, b + Y_n).$$

As a consequence, one obtains that the Bayes estimator for the offspring mean based on the sample  $\mathcal{Z}_n^*$ , under squared error loss, is given by

$$\widetilde{m} = (\alpha m^{(0)} + Y_n + Z_n - Z_0) / (\alpha + Y_n^*), \quad (1.8)$$

with  $m^{(0)}$  being the mean of  $p(0)$ , and for  $\mu$  and  $\tau$  one has  $\widetilde{\mu} = \int_{\Theta} \mu(\theta) \pi(\theta | \mathcal{Z}_n^*) d\theta$  and  $\widetilde{\tau} = \widetilde{m} \widetilde{\mu}$ , respectively.

*Remark 2.* In particular, for the examples considered in Remark 1, using (1.6), one has that the beta distribution family is an appropriate conjugate class of prior distributions for  $q$  in the binomial control case and for  $\theta$  when the control distributions are negative binomial. For the Poisson control distribution case, an appropriate conjugate class of priors for  $\theta$  is the gamma distribution family.

*Remark 3.* The theoretical approach to dealing with the inference issues related to a CBP is to assume that the control law belongs to a power series distribution family. This is an exponential family that includes many important distributions. It is worth noting that, from a practical standpoint, in most situations the choice of the control process, whether it is governed by a Poisson, binomial, negative binomial, or some other scheme, should be a prior specification based on knowledge of the development of the population.

### 1.2.2 Analysis based on population size in each generation: Gibbs sampler

In real situations, it is difficult to observe the whole family tree up to the current generation or even the random variables  $Z_l(k)$ ,  $k \geq 0$ ,  $l = 0, \dots, n-1$ . Hence, in this subsection we shall assume the more realistic requirement that these are unobservable, with the observable data being  $\mathcal{Z}_n = \{Z_0, \dots, Z_n\}$ . Given the definition of the model, an expression of the posterior distribution for  $(p, \theta)$  after observing  $\mathcal{Z}_n$  can not be displayed in a closed form. Consequently, we shall describe an algorithm based on the Gibbs sampler (see e.g., [2]) to approximate it only by observing  $\mathcal{Z}_n$ . To this end, it is necessary to take the unobservable variables  $Z_l(k)$ ,  $k \geq 0$ ,  $l = 0, 1, \dots, n-1$  as being latent variables, and consider the augmented parameter vector  $(p, \theta, \mathcal{Z}_n^*)$ . Let  $\pi(p, \theta | \mathcal{Z}_n)$  denote the posterior distribution of  $(p, \theta)$  after observing  $\mathcal{Z}_n$ . We shall approximate the posterior distribution of  $(p, \theta, \mathcal{Z}_n^*)$  after observing  $\mathcal{Z}_n$ , denoted by  $\pi(p, \theta, \mathcal{Z}_n^* | \mathcal{Z}_n)$ , and from this obtain an approximation for its marginal distribution  $\pi(p, \theta | \mathcal{Z}_n)$ . To use the Gibbs sampler, first, it is necessary to obtain the conditional posterior distribution of  $(p, \theta)$  after observing  $\mathcal{Z}_n$  and  $\mathcal{Z}_n^*$ , which is denoted by  $\pi(p, \theta | \mathcal{Z}_n, \mathcal{Z}_n^*)$ , and the conditional posterior distribution of  $\mathcal{Z}_n^*$  after observing  $(p, \theta, \mathcal{Z}_n)$ , denoted by  $f(\mathcal{Z}_n^* | p, \theta, \mathcal{Z}_n)$ .

Taking into account that, for  $l = 0, \dots, n-1$ ,

$$Z_{l+1} = \sum_{k \geq 0} k Z_l(k), \quad (1.9)$$

$\pi(p, \theta | \mathcal{Z}_n, \mathcal{Z}_n^*)$  is the same as  $\pi(p, \theta | \mathcal{Z}_n^*)$  given in (1.7). Let us now consider  $f(\mathcal{Z}_n^* | p, \theta, \mathcal{Z}_n)$ . Denoting by  $P(\cdot)$  the conditional probability given an offspring distribution  $p$  and control distributions governed by  $\theta$  (the explicit indication of the conditioning on  $p$  and  $\theta$  is dropped for notational clarity), since the individuals reproduce independently and the control distributions are independent of the offspring distribution, one has that, for  $z_l(k) \in \mathbb{N} \cup \{0\}$ ,  $k \geq 0$ ,  $l = 0, 1, \dots, n-1$ ,

$z_l \in \mathbb{N}$ ,  $l = 0, \dots, n$ , satisfying the constraints  $z_l = \sum_{k \geq 0} k z_{l-1}(k)$ ,  $l = 1, \dots, n$ ,

$$\begin{aligned} P(Z_l(k) = z_l(k), k \geq 0, l = 0, 1, \dots, n-1 \mid Z_0 = z_0, \dots, Z_n = z_n) \\ = \prod_{l=0}^{n-1} P(Z_l(k) = z_l(k), k \geq 0 \mid Z_l = z_l, Z_{l+1} = z_{l+1}). \end{aligned}$$

Hence,

$$f(\mathcal{Z}_n^* \mid p, \theta, \mathcal{Z}_n) = \prod_{l=0}^{n-1} f(Z_l(k), k \geq 0 \mid p, \theta, Z_l, Z_{l+1}),$$

where  $f(Z_l(k), k \geq 0 \mid p, \theta, Z_l, Z_{l+1})$  denotes the conditional distribution of the random sequence  $\{Z_l(k), k \geq 0\}$  given  $p, \theta, Z_l$ , and  $Z_{l+1}$ . Now, writing  $\phi_l^* = \sum_{k \geq 0} z_l(k)$ ,

$$\begin{aligned} P(Z_l(k) = z_l(k), k \geq 0 \mid Z_l = z_l, Z_{l+1} = z_{l+1}) \\ = \frac{1}{P(Z_{l+1} = z_{l+1} \mid Z_l = z_l)} \frac{\phi_l^*!}{\prod_{k \geq 0} z_l(k)!} \prod_{k \geq 0} p_k^{z_l(k)} a_{z_l}(\phi_l^*) \theta^{\phi_l^*} / A_{z_l}(\theta). \end{aligned}$$

Thus, computationally, an appropriate way to obtain a sample from  $f(\mathcal{Z}_n^* \mid p, \theta, \mathcal{Z}_n)$  is as follows. Given the known sample  $\{z_0, \dots, z_n\}$  and known values of  $\theta$  and  $p$ , one samples, for each  $l = 0, 1, \dots, n-1$ , a value  $\phi_l^*(z_l)$  from the distribution of the variable  $\phi_l(z_l)$  given by (1.2). Then, for each  $l = 0, 1, \dots, n-1$ , one samples a sequence  $\{z_l(k), k \geq 0\}$  from the multinomial probabilities  $\frac{\phi_l^*(z_l)!}{\prod_{k \geq 0} z_l(k)!} \prod_{k \geq 0} p_k^{z_l(k)}$ ,  $k \geq 0$ , normalized by considering the constraint  $z_{l+1} = \sum_{k \geq 0} k z_l(k)$ . Notice that, although the cardinality of the support of the reproduction law may be infinite, once  $z_{l+1}$  is known, only a finite number of coordinates of sequence  $\{z_l(k), k \geq 0\}$  are non-null. Indeed,  $z_l(k) = 0$  for all  $k \geq z_{l+1}$ .

Once it is known how to obtain samples from the distributions  $\pi(p, \theta \mid \mathcal{Z}_n, \mathcal{Z}_n^*)$  and  $f(\mathcal{Z}_n^* \mid p, \theta, \mathcal{Z}_n)$ , the Gibbs sampler algorithm works as follows:

```

Initialize  $l = 0$ 
Generate  $p^{(0)} \sim \text{DP}(p(0), \alpha)$ 
Generate  $\theta^{(0)}$  from (1.6)
Iterate
   $l = l + 1$ 
  Generate  $\mathcal{Z}_n^{*(l)} \sim f(\mathcal{Z}_n^* \mid p^{(l-1)}, \theta^{(l-1)}, \mathcal{Z}_n)$ 
  Generate  $(p^{(l)}, \theta^{(l)}) \sim \pi(p, \theta \mid \mathcal{Z}_n^{*(l)})$ 

```

Notice that, given the sample  $\mathcal{Z}_n$ , the maximum number of coordinates of  $p^{(l)}$ , for all  $l \geq 0$ , involved in the algorithm is  $1 + \max_{1 \leq k \leq n} \{Z_k\}$ . Hence, in the last step of the algorithm, bearing in mind Equation (1.7) and the properties of the Dirichlet process, one obtains these probabilities from the Dirichlet distribution.

The sequence  $\{(p^{(l)}, \theta^{(l)}, \mathcal{Z}_n^{*(l)})\}_{l \geq 0}$  is an ergodic Markov chain, and the stationary distribution of that Markov chain is just the sought-after joint distribution,  $\pi(p, \theta, \mathcal{Z}_n^* \mid \mathcal{Z}_n)$ . Several practical implementation issues must be taken into account



for success with the sample obtained by the method described above. Common approaches to reaching the equilibrium distribution as well as to reducing the autocorrelation in the sample are to choose a sufficient burn-in period,  $N$ , and to thin the output by storing only every  $G$ th value after the burn-in period ( $G$  is known as the batch size). Thus, for a run of the sequence  $\{(p^{(l)}, \theta^{(l)}, \mathcal{Z}_n^{*(l)})\}_{l \geq 0}$ , one chooses  $Q + 1$  vectors  $\{(p^{(N)}, \theta^{(N)}), (p^{(N+G)}, \theta^{(N+G)}), \dots, (p^{(N+QG)}, \theta^{(N+QG)})\}$ . These vectors are approximately independent sampled values of the distribution  $\pi(p, \theta | \mathcal{Z}_n)$  if  $G$  and  $N$  are large enough (see [21]). Since they could be affected by the initial state  $(p^{(0)}, \theta^{(0)})$ , the algorithm is applied  $T$  times, obtaining a final sample of length  $T(Q + 1)$ . To determine  $N$ ,  $G$ , and  $T$  in practice, we shall make use of the Gelman-Rubin-Brooks and autocorrelation diagnostics (see [3] and [4]). From this sample one can estimate  $\pi(p, \theta | \mathcal{Z}_n)$  and its marginal distributions,  $\pi(p | \mathcal{Z}_n)$  and  $\pi(\theta | \mathcal{Z}_n)$ , by making use of kernel density estimators. These posterior densities can be used to calculate numerically highest-probability-density (HPD) credible sets for the respective parameters, yielding sets in which there is a high probability of finding those parameters. In general, if  $\Psi(p, \theta)$  denotes a function of the offspring law and the control parameter (we shall be interested below in  $m$ ,  $\mu$ , and  $\tau$ ) then

$$\pi(\Psi | \mathcal{Z}_n) = \int \pi(\Psi | \mathcal{Z}_n, p, \theta) \pi(p, \theta | \mathcal{Z}_n) dp d\theta.$$

Again using kernel density estimators, one can also approximate  $\pi(\Psi | \mathcal{Z}_n)$  and calculate its HPD sets.

### 1.2.3 Approaches to prediction

A very important problem from a practical standpoint is to infer the size of future generations from currently available information. Thus, from a sample  $\{Z_0, \dots, Z_n\}$ , one desires inferential statements about unobserved  $Z_{n+l}$ ,  $l \geq 1$ . Few results related to this topic can be found in the branching process theory literature (see [13] and [15] for BGW processes). From a Bayesian standpoint, any inferential statement about  $Z_{n+l}$ ,  $l \geq 1$ , given known population sizes until generation  $n$ , is contained in the posterior predictive distribution  $f(Z_{n+l} | \mathcal{Z}_n)$ . Of course, the inferential content of the predictive distribution may be appropriately summarized to provide an estimator of  $Z_{n+l}$  as the mean of  $f(Z_{n+l} | \mathcal{Z}_n)$ , and interval estimates of  $Z_{n+l}$  such as the class of HPD sets which may be derived from  $f(Z_{n+l} | \mathcal{Z}_n)$ . For simplicity, we shall focus on the set  $\{Z_{n+l} > 0\}$ , avoiding approximating the density of the mass point  $\{Z_{n+l} = 0\}$  whose estimation is obvious. The difficulty lies in finding a closed form for this distribution.

We shall present two ways of approximating the predictive distribution, both applying a Monte Carlo procedure. The first is a sampling-based method, and the

second is based on approximating  $E[Z_{n+l} | \mathcal{Z}_n]$  and  $Var[Z_{n+l} | \mathcal{Z}_n]$  and then seeking a parametric model.

**Method A.** We consider a random sample  $\{(p^{(1)}, \theta^{(1)}), \dots, (p^{(r)}, \theta^{(r)})\}$  from  $\pi(p, \theta | \mathcal{Z}_n)$ . In particular, we shall use the one obtained with the Gibbs sampler in Subsection 1.2.2, i.e.,  $r = T(Q + 1)$ . It is clear that

$$f(Z_{n+l} | \mathcal{Z}_n) = \int f(Z_{n+l} | \mathcal{Z}_n, p, \theta) \pi(p, \theta | \mathcal{Z}_n) dp d\theta.$$

On the basis of this formula, for each  $(p^{(i)}, \theta^{(i)})$ ,  $i = 1, \dots, r$ , one can simulate  $s$  processes until the  $l$ th generation, which started with  $Z_n$  individuals, obtaining the values  $z_{n+l,1}^{(i)}, z_{n+l,2}^{(i)}, \dots, z_{n+l,s}^{(i)}$  from  $Z_{n+l} > 0$ , and use them to approximate  $f(Z_{n+l} | \mathcal{Z}_n, p^{(i)}, \theta^{(i)})$  by a Gaussian kernel estimator

$$f^{(i)}(x) = \frac{1}{s} \sum_{j=1}^s \frac{1}{b^{(i)}} K\left(\frac{x - z_{n+l,j}^{(i)}}{b^{(i)}}\right), \quad x \in \mathbb{R},$$

with  $b^{(i)}$  an appropriate bandwidth and  $K(x)$  the density of the standard normal distribution. Thus,  $f(Z_{n+l} | \mathcal{Z}_n)$  is estimated by

$$\hat{f}(x) = \frac{1}{r} \sum_{i=1}^r f^{(i)}(x), \quad x \in \mathbb{R}. \quad (1.10)$$

**Method B.** We consider a random sample  $\{(p^{(1)}, \theta^{(1)}), \dots, (p^{(r)}, \theta^{(r)})\}$  (again the sample obtained in Subsection 1.2.2) from  $\pi(p, \theta | \mathcal{Z}_n)$ , and for each  $i = 1, \dots, r$  one simulates  $s$  processes until the  $l$ th generation, which started with  $Z_n$ , reproduction law  $p^{(i)}$ , and control distribution governed by  $\theta^{(i)}$ . One calculates the mean and the variance of the  $s$ -values  $z_{n+l,j}^{(i)} > 0$ ,  $j = 1, \dots, s$ , obtaining an approximation to  $E[Z_{n+l} | \mathcal{Z}_n, p^{(i)}, \theta^{(i)}]$  and to  $Var[Z_{n+l} | \mathcal{Z}_n, p^{(i)}, \theta^{(i)}]$ . Finally,

$$E[Z_{n+l} | \mathcal{Z}_n] \approx \frac{1}{r} \sum_{i=1}^r E[Z_{n+l} | \mathcal{Z}_n, p^{(i)}, \theta^{(i)}],$$

and, considering that

$$Var[Z_{n+l} | \mathcal{Z}_n] = E[Var[Z_{n+l} | \mathcal{Z}_n, p, \theta]] + Var[E[Z_{n+l} | \mathcal{Z}_n, p, \theta]],$$

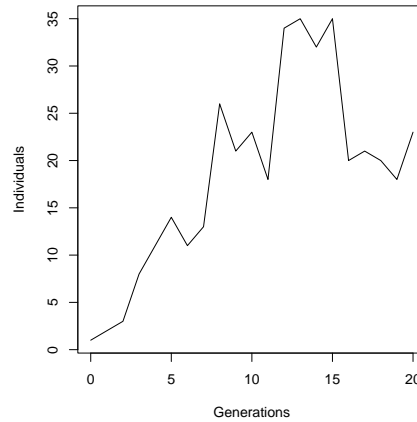
with the mean and the variance in the right term of the previous equality considered with respect to the distribution  $\pi(p, \theta | \mathcal{Z}_n)$ , then

$$\begin{aligned} \text{Var}[Z_{n+1} | \mathcal{Z}_n] &\approx \frac{1}{r} \sum_{i=1}^r \text{Var}[Z_{n+1} | Z_n, p^{(i)}, \theta^{(i)}] + \\ &+ \frac{1}{r-1} \sum_{i=1}^r \left( E[Z_{n+1} | Z_n, p^{(i)}, \theta^{(i)}] - \frac{1}{r} \sum_{i=1}^r E[Z_{n+1} | Z_n, p^{(i)}, \theta^{(i)}] \right)^2. \end{aligned}$$

As was already proposed in Mendoza and Gutiérrez-Peña (2000) for BGW processes, we also use a gamma distribution with mean and variance  $E[Z_{n+1} | \mathcal{Z}_n]$  and  $\text{Var}[Z_{n+1} | \mathcal{Z}_n]$ , respectively, (justified by the minimum logarithmic divergence criterion) to approximate  $f(Z_{n+1} | \mathcal{Z}_n)$ .

### 1.3 Simulated Example

In this section, we shall illustrate the methods described above by analysing an example with simulated data.



**Fig. 1.1** Evolution of the simulated population sizes.

We simulated 20 generations of a CBP with  $Z_0 = 1$ , offspring law  $p_0 = 0.0778$ ,  $p_1 = 0.2592$ ,  $p_2 = 0.3456$ ,  $p_3 = 0.2304$ ,  $p_4 = 0.0768$ ,  $p_5 = 0.0102$ , and  $\phi_n(k)$  having a Poisson distribution with parameter  $0.51k$ . Thus the offspring mean is  $m = 2$ , the control and the immigration parameters coincide, being  $\mu = \theta = 0.51$ , and in this case the asymptotic mean growth rate is  $\tau = m\mu = 1.02$ . As  $\mu < 1$ , we are considering a CBP with expected emigration. Figure 1.1 shows the evolution of the simulated population sizes. In an emulation of the classification of standard BGW processes, it was established in [11] that one refers to a subcritical, critical, or supercritical CBP depending on whether  $\tau$  is less than, equal to, or greater than unity. Despite the

expected emigration, using the results in [10], one can deduce that this supercritical CBP process has a positive probability of non-extinction.

We now focus on the estimation of  $p$ ,  $\mu$ ,  $m$ , and  $\tau$  based on the population size in each generation, by using the Gibbs sampler. We specify a Dirichlet process on the non-negative integers to model the prior distribution of the offspring law, avoiding any assumption about the cardinality of its support. Initially we choose  $\alpha = 1$  as concentration parameter, and a Poisson distribution as base measure. It is well known that the Poisson distribution models the number of events occurring within a given time interval when those events occur at a known average rate and independently of the time since the last event. It is thus appropriate for modeling a generic offspring process. We propose that the average rate of the Poisson distribution will be initially estimated by considering that no control is imposed on the population. Therefore, the maximum likelihood estimator of the offspring mean corresponding to a BGW process can be used, i.e., we propose the Poisson base distribution with mean  $(Z_1 + \dots + Z_n)/(Z_0 + \dots + Z_{n-1})$ , in this example, 1.06. Recall that for a Poisson control distribution –see Remark 1, b)–  $\mu = \theta$ , so that we shall henceforth refer in the analysis to  $\mu$ . With respect to the prior for the parameter  $\mu$ , as one does not know a priori what kind of expected migration is taking place, one can take the value of  $\mu$  to be 1, and set as a prior distribution for  $\mu$  a gamma distribution with mean 1 (the shape parameter is chosen to be unity). How the choices of prior elicitation affect the inferences will be evaluated below.

In the simulation, we set  $T = 50$  and ran the algorithm described in Subsection 1.2.2 20 000 times for each chain. Using the Gelman-Rubin-Brooks diagnostic plots for  $(p, \mu)$ , we took  $N = 5 000$ . Table 1.1 lists the values of the estimated potential scale reduction factor together with 97.5% confidence upper bounds and the autocorrelation values for  $\mu$  and the first values of  $p$ . That the values of the estimated scale reduction factor are close to unity suggests that further simulations will not improve the values of the listed scalar estimators (see [3] and [4]). Finally, for the autocorrelation study, we chose  $G = 600$ , and consequently  $Q = 25$ . The final sample size was therefore 1 300.

	Potential Scale Reduction		Autocorrelation		
	Est.	97.5%	lag1	lag100	lag600
$\mu$	1.04	1.06	0.9894	0.6669	0.0577
$p_0$	1.03	1.04	0.9937	0.6558	0.0776
$p_1$	1.02	1.03	0.9925	0.5538	0.0446
$p_2$	1.02	1.02	0.9895	0.4922	0.0326
$p_3$	1.02	1.04	0.9893	0.5055	0.0302
$p_4$	1.02	1.02	0.9684	0.1478	0.0020
$p_5$	1.03	1.04	0.9337	0.0500	0.0000
$p_6$	1.06	1.06	0.7524	0.0059	-0.0016

**Table 1.1** Potential scale reduction factor and autocorrelation for  $\mu$  and the first values of  $p$ .

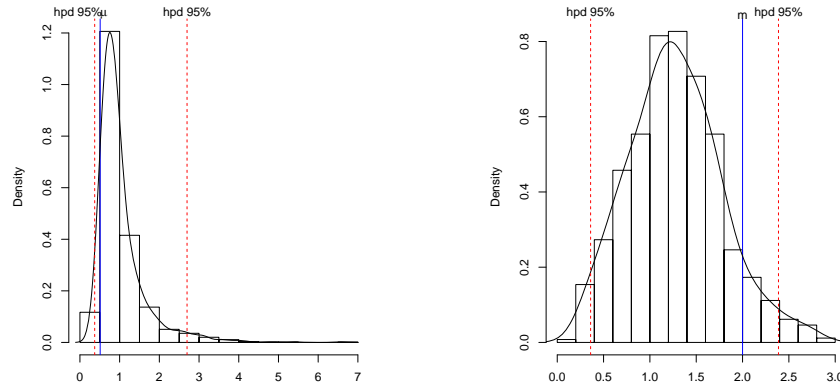
To evaluate the algorithm’s efficiency, Table 1.2 presents some summary statistics for the posterior distributions of  $\mu$ ,  $m$ , and  $\tau$ . Note that, due to the batch procedure,

the time-series standard errors (TSSE) are very close to the Monte Carlo standard errors (MCSE). Also, for the three parameters, the standard errors (MCSE and TSSE) are less than 5% of the posterior standard deviation (SD), indicating that the number of observations considered seems to be a reasonable choice.

	MEAN	SD	MCSE	TSSE
$\mu$	1.0187	0.6015	0.0167	0.0155
$m$	1.2751	0.4987	0.0138	0.0127
$\tau$	1.0597	0.0728	0.0020	0.0018

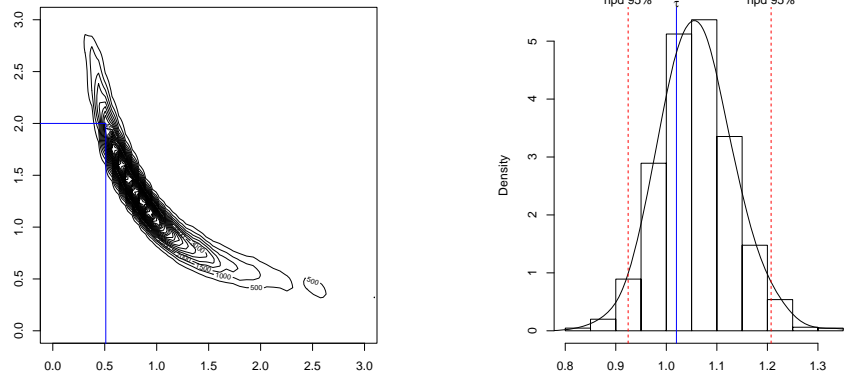
**Table 1.2** Summary statistics for the posterior distributions of  $\mu$ ,  $m$ , and  $\tau$ .

Figures 1.2 and 1.3 show the estimated posterior density for  $\mu$ ,  $m$ , and  $\tau$  together with their Bayes estimates under squared error loss, and the 95% HPD sets. The contour plot of the estimated posterior density for  $(\mu, m)$  is also shown. One observes that the 95% HPD sets contain the true values of the parameters. It is particularly noteworthy that the method is not very accurate at identifying the parameters  $m$  and  $\mu$ , although  $P(\mu < 1 \mid \mathcal{Z}_{20})$  is estimated at 0.662, identifying the process as having expected emigration, and  $P(m > 1 \mid \mathcal{Z}_{20})$  is estimated at 0.711, identifying a mean reproduction capacity of greater than unity. The method also provides a good estimate of the process's asymptotic mean growth rate,  $\tau$ , which is the parameter that determines the limiting behaviour of the process. In this case,  $P(\tau > 1 \mid \mathcal{Z}_{20})$  is approximated by 0.80, identifying a supercritical CBP. The contour plot shows clearly the interdependence of the parameters  $m$ ,  $\mu$ , and  $\tau$ .

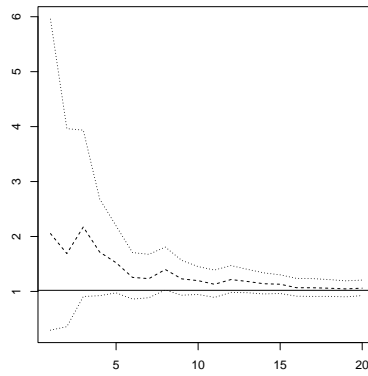


**Fig. 1.2** Estimated posterior density for  $\mu$  (left) and  $m$  (right).

Figure 1.4 illustrates the *long-term* behaviour of the estimates of  $\tau$ , showing for each generation their Bayes estimates under squared error loss and their respective 95% HPD sets. Note that one has estimates closer to the real value and narrower HPD intervals as the generations advance.



**Fig. 1.3** Contour plot of estimated posterior density for  $(\mu, m)$  (left) and estimated posterior density for  $\tau$  (right).

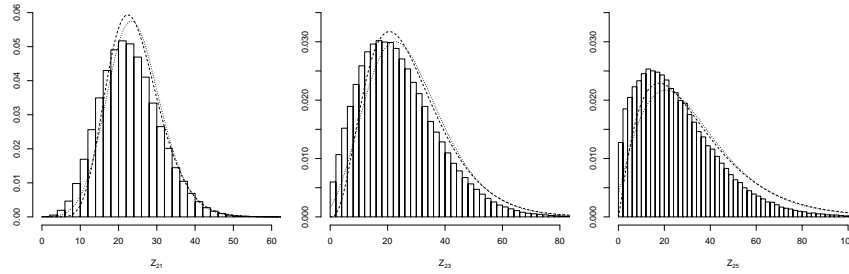


**Fig. 1.4** Evolution of the squared error loss estimates of  $\tau$  with 95% HPD bands. The horizontal line represents the true parameter value.

Using the information on the population sizes until generation 20, we obtained via the methods described in Subsection 1.2.3 the predictive distributions of  $Z_{21}$ ,  $Z_{23}$ , and  $Z_{25}$  on their respective non-extinction sets. Methods A and B were applied by simulating  $s = 1\,000$  processes, started with  $Z_{20} = 23$  until the 5th generation, and reproduction law and control parameter  $(p^{(i)}, \theta^{(i)})$ ,  $i = 1, \dots, 1\,300$ .

	MEAN	95% prediction HPD
	Method A/Method B	Method A/Method B
$Z_{21}$	24.3844/24.3840	11.2112–38.2382 / 11.6116–38.7387
$Z_{23}$	27.8852/27.8475	2.8028–57.8579 / 4.6046–58.9590
$Z_{25}$	31.7907/32.6389	3.9039–86.9870 / 1.1011–84.4845

**Table 1.3** Mean and 95% prediction HPD sets for  $Z_{21}$ ,  $Z_{23}$ , and  $Z_{25}$ .



**Fig. 1.5** Histogram of the data from the true distribution together with estimated posterior densities for  $Z_{21}$ ,  $Z_{23}$ , and  $Z_{25}$ . Method A: dotted line; Method B: discontinuous line.

Figure 1.5 shows the predictive posterior distribution of  $Z_{21}$ ,  $Z_{23}$ , and  $Z_{25}$  estimated by the two methods described above, together with the histogram of the data from the true distribution. The sampling-based approach and the gamma model lead to similar estimates for  $Z_{21}$ ,  $Z_{23}$ , and  $Z_{25}$ . Table 1.3 presents estimates for  $Z_{21}$ ,  $Z_{23}$ , and  $Z_{25}$  together with the 95% HPD sets. The accuracy of these approximations, assessed by comparison with the true distribution, is reasonable. In spite of the fact that the estimation of  $m$  and  $\mu$  are not so precise as desirable, the combination of both to estimate  $\tau$  is quite good, as shown Figure 1.3 (left) and this leads to the estimation of the forecast values of the process are accurate. Nevertheless, according to Figure 1.3 (right), the estimated values of  $\tau$  tend to be slightly greater than the true value, so that the predictive values also show this tendency.

Concentration parameter	Base measure $m^{(0)} = 1.06$					
	Poisson			Geometric		
	MEAN	HPD 95%		MEAN	HPD 95%	
0.25	1.0629	0.9244	1.2060	1.0584	0.9164	1.2107
0.50	1.0646	0.9258	1.2136	1.0626	0.9251	1.2122
0.75	1.0616	0.9195	1.2118	1.0618	0.9187	1.2180
1	1.0597	0.9244	1.2073	1.0585	0.9053	1.2129
5	1.0583	0.9249	1.2104	1.0618	0.9062	1.2271
10	1.0613	0.9204	1.2081	1.0610	0.8980	1.2194
20	1.0554	0.9099	1.2041	1.0642	0.9065	1.2391

**Table 1.4** Sensitivity analysis for  $\tau|Z_{20}$ .

Finally, we examine the sensitivity of inferences of the main parameters of interest to the choice of the priors (in particular, focusing on the concentration parameter, the base distribution, and the gamma parameters). For simplicity, we present the analysis focussing on the asymptotic mean growth rate,  $\tau$ , a parameter that determines the future evolution of the process. First, we analyse the influence on the choices of the concentration parameter and of the kind of base distribution (setting its mean to 1.06). With the prior on  $\mu$  assumed to be a gamma distribution with mean

1 (its shape parameter taken as unity), the results in Table 1.4 show the estimation of  $\tau$  not to be very sensitive to such changes. Second, with the same prior on  $\mu$  and the concentration parameter equal to unity, the results in Table 1.5 show the estimation of  $\tau$  neither to be very sensitive to the choice of the base distribution or of its mean. Finally, we analyse the influence of the choice of the shape and mean parameters of the gamma distribution. Taking the previous study into account, we took the concentration parameter to be equal to unity, and the base distribution to be a Poisson distribution with mean 1.06. The results in Table 1.6 again allow one to conclude that the estimation of  $\tau$  is not very sensitive to the prior parameters.

$m^{(0)}$	Base measure $\alpha = 1$					
	Poisson			Geometric		
	MEAN	HPD 95%		MEAN	HPD 95%	
0.25	1.05592	0.93007	1.19095	1.05441	0.92359	1.19157
0.50	1.05687	0.92847	1.19287	1.05850	0.92709	1.19397
0.75	1.05983	0.93648	1.19943	1.06178	0.91882	1.21939
1	1.05664	0.91961	1.21131	1.05850	0.91213	1.20456
1.5	1.06283	0.91825	1.22327	1.06127	0.90646	1.22152
2	1.06270	0.90100	1.21972	1.06129	0.91375	1.21722
2.5	1.06631	0.91151	1.22647	1.06287	0.90328	1.24079
3	1.06887	0.90780	1.23867	1.06043	0.90816	1.23099
3.5	1.06978	0.89095	1.24436	1.06787	0.90444	1.25410

**Table 1.5** Sensitivity analysis for  $\tau | \mathcal{Z}_{20}$ .

In most of the situations, as was noted in Remark 3, one has prior knowledge of the kind of the control distribution, and this was indeed our approach in the simulated example where we assumed the control to be applied through a Poisson control distribution. We next examined how the method works when assuming complete ignorance of the control law. To this end, we implemented the method by considering prior binomial, negative binomial, and Poisson control laws. Thus, besides the data already simulated and presented in Figure 1.1 which corresponded to an expected emigration, we also considered a new sample from a CBP with Poisson control distributions with the same asymptotic mean growth rate as the previous example but now with an expected immigration (with parameters  $m = 0.51$  and  $\mu = 2$ ). The results in Table 1.7 illustrate how well (or now poorly) the fitted models, in both the expected emigration and the expected immigration examples, identify the asymptotic mean growth rate of the process as well as a supercritical or a subcritical reproduction mean (i.e.,  $m > 1$  or  $m < 1$ , respectively). To assess whether the fitted models detect an expected emigration or immigration, we calculated  $P(\mu < 1 | \mathcal{Z}_{20})$  or  $P(\mu > 1 | \mathcal{Z}_{20})$ , respectively. The approximations of these probabilities are also given in the table. The results in the table show the asymptotic mean growth rate of the process to usually be well identified whichever situation is considered. This is because the generation-by-generation population sizes provide enough information to estimate this value. However, the estimates of the offspring mean and of the migration process are not generally appropriate in the absence of prior knowledge



about the type of control. Analysing the fitted model considering a prior Poisson control (the genuine control distribution), we obtained acceptable results for the estimates of the offspring mean and of the expected migration process in both the emigration and the immigration examples. Considering a prior binomial control, we found a tendency to overestimate the offspring mean for the two simulated data sets. This was especially so in case of the expected immigration situation, precisely to compensate for the effect of the immigration (recall that the binomial control only allows for emigration). Finally, with a negative binomial control, one deduces from the results in Table 1.7 that, despite both the simulated and the fitted models allowing for any kind of migration, the fitted model identifies well neither the offspring mean nor the emigration process when the latter is expected. The conclusion to be drawn from these simulated examples is thus that, to obtain reliable estimates of the parameters of interest, one should have some prior knowledge of the kind of control law.

*Remark 4.* For the computation of the examples, we used the statistical software **R**, a language and environment for statistical computing and graphics (see [17]), performing the simulations by parallel computing using the `Rmpi` ([12]) and `snow` (see [22]) packages, and the convergence diagnostics using the `coda` package (see [16]).

## 1.4 Concluding Remarks

As was noted in the Introduction, nowadays controlled branching processes are one of the most relevant branching models for study. Apart from their inherent interest, this current relevance is due to the fact that they generalize many important branching processes – migration models, for instance. Moreover, beyond the framework of population dynamics, these models cover and generalize popular models for the analysis of count-data time series, such as the INAR(1) models (see [23] and [24] for reviews). Consequently, the development of the inferential theory of CBPs is an important task to provide a guarantee for their application to many real-world problems. Some contributions to this theory have already been made from a frequentist standpoint, either for the general model itself ([7] and [20]) or for some of its particular cases, such as the models with immigration (see, e.g., [18] and the references therein) or INAR(1) (see [25]).

Respect to a Bayesian perspective, this paper extends in several senses the preliminary contributions established in [5] and [6]. In particular, we have focused on a CBP with non-parametric offspring law and with random control variables instead of deterministic ones, assuming these belonging to a parametric family of distributions that depend on a single parameter, termed the control parameter (or its equivalent, the migration parameters). To avoid any assumption about the cardinality of the support of the offspring law, a Dirichlet process was introduced. The classical application of this methodological approach to a branching context in a non-parametric framework for the offspring law requires the observation of the entire family tree

Mean parameter		Gamma distribution					
		Shape parameter					
		0.25	0.50	0.75	1	2.5	7.5
0.25	MEAN	1.0607	1.0546	1.0563	1.0506	1.0408	1.0293
	95% HPD lower bound	0.9234	0.9021	0.9138	0.9023	0.8980	0.8841
	95% HPD upper bound	1.2075	1.2219	1.1988	1.2153	1.1848	1.1860
0.5	MEAN	1.0597	1.0524	1.0570	1.0523	1.0518	1.0453
	95% HPD lower bound	0.9168	0.9189	0.9052	0.9152	0.9047	0.9101
	95% HPD upper bound	1.2128	1.1997	1.2160	1.2063	1.2066	1.1950
0.75	MEAN	1.0638	1.0660	1.0632	1.0598	1.0571	1.0575
	95% HPD lower bound	0.9242	0.9242	0.9198	0.9233	0.9177	0.9055
	95% HPD upper bound	1.2065	1.2315	1.2136	1.2043	1.2042	1.2133
1	MEAN	1.0617	1.0613	1.0606	1.0597	1.0620	1.0629
	95% HPD lower bound	0.9193	0.9146	0.9164	0.9244	0.9191	0.9219
	95% HPD upper bound	1.2132	1.2144	1.2119	1.2073	1.2121	1.2125
5	MEAN	1.0599	1.0609	1.0652	1.0662	1.0628	1.0596
	95% HPD lower bound	0.9042	0.9247	0.9302	0.9289	0.9241	0.9185
	95% HPD upper bound	1.2102	1.2045	1.2124	1.2083	1.2107	1.2102
10	MEAN	1.0617	1.0660	1.0620	1.0641	1.0607	1.0659
	95% HPD lower bound	0.9179	0.9225	0.9188	0.9273	0.9212	0.9319
	95% HPD upper bound	1.2101	1.2128	1.2089	1.2037	1.1952	1.2101
20	MEAN	1.0590	1.0630	1.0644	1.0603	1.0632	1.0633
	95% HPD lower bound	0.9232	0.9323	0.9195	0.9131	0.9259	0.9355
	95% HPD upper bound	1.2109	1.2084	1.2042	1.2082	1.2099	1.1960

**Table 1.6** Sensitivity analysis for  $\tau | \mathcal{Z}_{20}$ .

Simulated model	Fitted models					
	P control		B control		NB control	
Expected emigration	$\tau   \mathcal{Z}_{20}$					
$m = 2; \tau = 1.02$	Mean	Variance	Mean	Variance	Mean	Variance
	1.060	0.005	1.054	0.005	1.062	0.007
	$P(m > 1   \mathcal{Z}_{20})$					
	0.711		0.995		0.400	
	$P(\mu < 1   \mathcal{Z}_{20})$					
	0.662		0.317			
Expected immigration	$\tau   \mathcal{Z}_{20}$					
$m = 0.51; \tau = 1.02$	Mean	Variance	Mean	Variance	Mean	Variance
	1.061	0.006	1.056	0.007	1.064	0.012
	$P(m < 1   \mathcal{Z}_{20})$					
	0.626		0.013		0.735	
	$P(\mu > 1   \mathcal{Z}_{20})$					
	0.767		0.801			

**Table 1.7** Sensitivity analysis: control prior distribution. P control = Poisson control; B control = binomial control; NB control = negative binomial control.

(up to some generation). Although we developed this approach -with the novelty of using the Dirichlet process-, we reckon that from a practical standpoint it is more realistic and relevant to avoid the need for observation of the complete family tree and we only consider the record of the total generation sizes. On the basis of this kind of sample, and making use of the Gibbs sampler and kernel density estimators, we proposed a method to approximate the posterior density of the control parameter (or of its parametrization as the migration parameter), of the offspring law, and consequently the posterior densities of the offspring mean and asymptotic mean growth rate, regardless of the nature of the offspring law (whether subcritical, critical, or supercritical) and the kind of migration. It is worth noting that the frequentist methods given until now depend strongly on this nature, making a Bayesian method preferable in this sense. In fact, one can make inferences on the nature of the offspring law or the kind of migration (immigration or emigration) from the posterior density of the offspring mean or of the migration parameter, respectively.

Furthermore, the above method also allows one to approximate the predictive posterior densities. Two methods were proposed to make inferences on the size of future generations.

We considered a simulated example in some depth to illustrate our findings, and included a detailed sensitivity analysis regarding the choices of the priors. This showed the methodological approach to not be unduly influenced by the choice of priors of the control parameter or the priors of the concentration parameter or the base distribution. However, it revealed the need for prior knowledge of the kind of control being applied. With respect to the comparison of the two methods for predictions, the simulated example showed the two approaches to lead essentially to the same results.

As an overall conclusion, we would state that the proposed procedure allowing inference based only on total generation sizes constitutes the main contribution of the present work.

**Acknowledgements** The authors would like to thank Professors Horacio González-Velasco and Carlos García-Orellana for providing them with computational support. This research has been supported by the Ministerio de Educación, Cultura y Deporte (grant FPU13/03213), Ministerio de Economía y Competitividad of Spain (grants MTM2012-31235 and MTM2015-70522-P), the Junta of Extremadura (grant GR15105), the FEDER, and the National Fund for Scientific Research at the Ministry of Education and Science of Bulgaria (grant DFNI-I02/17).

## References

1. Bagley, J.H.: On the almost sure convergence of controlled branching processes. *J. Appl. Probab.* **23**, 827–831 (1986)
2. Brooks, S.: Markov Chain Monte Carlo method and its application. *J. R. Stat. Soc. Ser. D (The Statistician)* **47**, 69–100 (1998)
3. Brooks, S., Gelman, A.: General methods for monitoring convergence of iterative simulations. *J. Comput. Graph. Statist.* **7**, 434–455 (1998)

4. Gelman, A., Rubin, D.: Inference from iterative simulation using multiple sequences. *Statist. Sci.* **7**, 457–511 (1992)
5. González, M., Gutiérrez, C., Martínez, R., del Puerto, I.: Bayesian inference for controlled branching processes through MCMC and ABC methodologies. *Rev. R. Acad. Cien. Serie A. Mat.* **107**, 459–473 (2013)
6. González, M., Martín, J., Martínez, R., Mota, M.: Non-parametric Bayesian estimation for multitype branching processes through simulation-based methods. *Comput. Statist. Data Anal.* **52**, 1281–1291 (2008)
7. González, M., Martínez, R., del Puerto, I.: Nonparametric estimation of the offspring distribution and mean for a controlled branching process. *Test* **13**, 465–479 (2004)
8. González, M., Martínez, R., del Puerto, I.: Estimation of the variance for a controlled branching process. *Test* **14**, 199–213 (2005)
9. González, M., Minuesa, C., del Puerto, I.: Maximum likelihood estimation and Expectation-Maximization algorithm for controlled branching process. *Comput. Statist. Data Anal.* **93**, 209–227 (2016)
10. González, M., Molina, M., del Puerto, I.: On the class of controlled branching process with random control functions. *J. Appl. Probab.* **39**, 804–815 (2002)
11. González, M., Molina, M., del Puerto, I.: Asymptotic behaviour of critical controlled branching process with random control function. *J. Appl. Probab.* **42**:2, 463–477 (2005)
12. Yu, H.: Rmpi: Parallel Statistical Computing in R. *R News* **2**, 10–14 (2002)
13. Jagers, P.: *Branching processes with biological applications*. John Wiley & Sons (1975)
14. Martínez, R., Mota, M. and del Puerto, I.: On asymptotic posterior normality for controlled branching processes with random control function. *Statistics* **43**, 367–378 (2009)
15. Mendoza, M., Gutiérrez-Peña, E.: Bayesian conjugate analysis of the Galton-Watson process. *Test* **9**, 149–171 (2000)
16. Plummer, M., Best, N., Cowles, K., Vines, K.: coda: Output analysis and diagnostics for MCMC. R package version 0.13-5 (2010)
17. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2014)
18. Rahimov, I.: Limit distributions for weighted estimators of the offspring mean in a branching process. *Test* **18**, 568–583 (2009)
19. Sevastyanov, B.A., Zubkov, A.: Controlled branching processes. *Theor. Prob. Appl.* **19**, 14–24 (1974)
20. Sriram, T., Bhattacharya, A., González, M., Martínez, R., del Puerto, I.: Estimation of the offspring mean in a controlled branching process with a random control function. *Stochastic Processes. Appl.* **117**, 928–946 (2007)
21. Tierney, L.: Markov chains for exploring posterior distributions. *Ann. Statist.* **22**, 1701–1762 (1994)
22. Tierney, L., Rossini, A.J., Li, N., Sevcikova, H.: snow: Simple Network of Workstations. R package version 0.3-3 (2013)
23. Tjøstheim, D.: Some recent theory for autoregressive count time series. *Test* **21**, 413–438 (2012)
24. Weiß, C.H.: Thinning operations for modelling time series of counts- a survey. *Adv. Statist. Anal.* **92**, 319–341 (2008)
25. Weiß, C.H.: Fully observed INAR(1) processes. *J. Appl. Stat.* **39**, 581–598 (2011)