

**AMSI**  
**VACATION**  
**RESEARCH**  
**SCHOLARSHIPS**  

---

**2017-2018**

---



**Stochastic models for populations  
with a carrying capacity**

Jiangrong Ouyang

Supervised by Dr. Sophie Hautphenne and Mr. Peter Braunsteins  
The University of Melbourne

Vacation Research Scholarships are funded jointly by the Department of Education and Training  
and the Australian Mathematical Sciences Institute.



Australian Government  
Department of Education and Training

**AMSI** AUSTRALIAN  
MATHEMATICAL  
SCIENCES  
INSTITUTE



**Abstract**

In this report, we focus on stochastic models for populations with a carrying capacity. Firstly, we will analyse different discrete-time and continuous-time population-size dependent branching processes for the modelling of the black robin population. Then we will develop parameter estimation methods to fit these models to the data. In particular, the approximation by linear fractional branching processes will give us lower and upper bounds for the value of the carrying capacity for the black robins.

**1 Introduction**

**1.1 Discrete-time models**

Consider a standard Galton-Watson process, where the population size at generation  $n$ ,  $Z_n$ , satisfies

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_i$$

where the offspring distribution  $\xi$  is independent and identically distributed.

Many biological populations experience a logistic growth, that is, the population per capita growth rate decreases as the population size approaches a maximum imposed by limited resources in the environment, known as the carrying capacity. We now introduce the carrying capacity,  $K$ , in our model and make the offspring distribution,  $\xi_i$  dependent on the population size:

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_i(Z_n),$$

where, in our case, we consider binary splitting, that is,

$$\mathbb{P}(\xi_i(z) = 2) = \frac{vK}{K + (2v - 1)z} = 1 - \mathbb{P}(\xi_i(z) = 0), \tag{†}$$

where  $v$  is a “stretching” parameter to control the splitting probability of the chain.

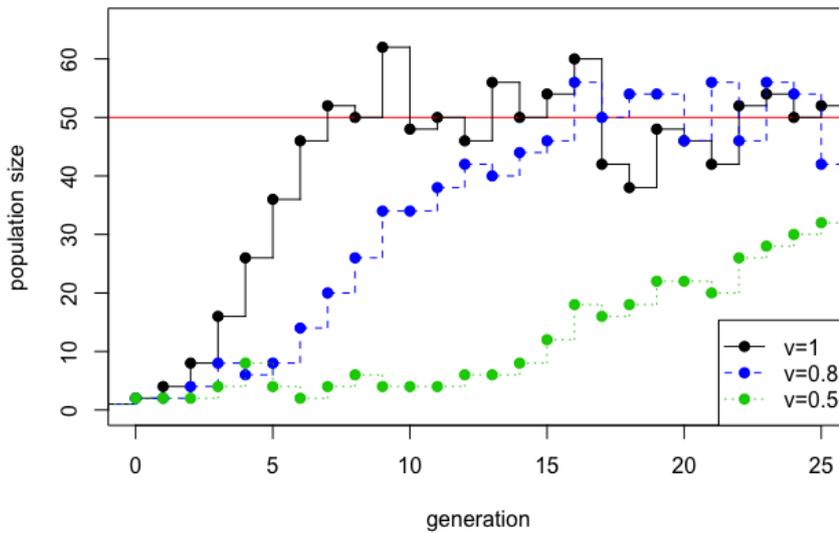


Figure 1: Effects of different parameter values  $v$  with  $K = 50$



All three trajectories approach the carrying capacity  $K$  along time. The trajectory with a larger value of parameter  $v$  has larger splitting probability, and thus it takes less time to reach the carrying capacity  $K$ .

The success rate  $\frac{vK}{K+(2v-1)z}$  is chosen to make  $K$  the threshold value at which the growth rate changes from positive to negative if the initial population size is less than  $K$ . When the population size reaches  $K$ , i.e.:  $z = K$ , then  $\mathbb{P}(\xi_i(K) = 2) = \frac{vK}{K+(2v-1)K} = \frac{1}{2}$ . This implies that the mean of the offspring distribution  $E(\xi_1) = m = 1$  at this time. We call the process *critical* if  $m = 1$ . When the population size is under  $K$ , i.e.:  $z < K$ , the mean of the offspring distribution is greater than 1, and we call the process *supercritical*. On the other hand,  $z > K$  implies  $m < 1$ , and the process is called *subcritical*. Therefore, any trajectory tends to approach the carrying capacity,  $K$ , and lingers around  $K$  until it dies out eventually.

The model above is a classic binary splitting model, where one individual either splits into 2 individuals or dies. The random variable  $\xi_i(z)$  has the same distribution as  $Y = 2X$ , where  $X \sim \text{Ber}(\frac{vK}{K+(2v-1)z})$ . And thus

$$(Z_n | Z_{n-1} = z_{n-1}) \sim \text{Bi} \left( z_{n-1}, \frac{vK}{K + (2v-1)z_{n-1}} \right)$$

Hence, by the Markov property,

$$\begin{aligned} \mathbb{P}[(Z_0, Z_1, \dots, Z_n) = (z_0, z_1, \dots, z_n)] \\ &= \prod_{i=1}^n \mathbb{P}(Z_i = z_i | Z_{i-1} = z_{i-1}) \\ &= \prod_{i=1}^n \binom{z_{i-1}}{\frac{z_i}{2}} p_2^{\frac{z_i}{2}} (1-p_2)^{z_{i-1} - \frac{z_i}{2}} \end{aligned}$$

where  $p_2 = p_2(z_i) = \frac{vK}{K+(2v-1)z_i}$ .

Then the log-likelihood is, (up to a constant difference)

$$\begin{aligned} l(K, v) &= \sum_{i=1}^n \left( \log(p_2^{\frac{z_i}{2}}) + \log(1-p_2)^{z_{i-1} - \frac{z_i}{2}} \right) \\ &= \sum_{i=1}^n \left( \log \left( \frac{p_2}{1-p_2} \right)^{\frac{z_i}{2}} + \log(1-p_2)^{z_{i-1}} \right) \\ &= \sum_{i=1}^n \left( \log \left( \frac{vK}{(1-v)K + (2v-1)z_{i-1}} \right)^{\frac{z_i}{2}} + \log \left( \frac{(1-v)K + (2v-1)z_{i-1}}{K + (2v-1)z_{i-1}} \right)^{z_{i-1}} \right). \end{aligned}$$

Note that, we keep the powers  $\frac{z_i}{2}$  and  $z_{i-1}$  instead of making them factors in the expression to prevent the occurrence of  $\log 0$  when we run the optimisation algorithm to find the maximum likelihood.

## 1.2 Continuous-time models

Consider a standard linear birth-death process with birth rate  $\lambda_i = i\lambda$  and death rate  $\mu_i = i\mu$ , where  $i$  is the population size. In this process, the population stays in state  $i$  for an exponential time with rate  $\lambda_i + \mu_i$ . Then it will jump up to state  $i + 1$  with probability  $\frac{\lambda_i}{\lambda_i + \mu_i}$  or down to state  $i - 1$  with complementary probability  $\frac{\mu_i}{\lambda_i + \mu_i}$ . Without carrying capacity  $K$ , the chain will grow exponentially with rate  $\omega = \lambda - \mu$  if  $\lambda > \mu$ .

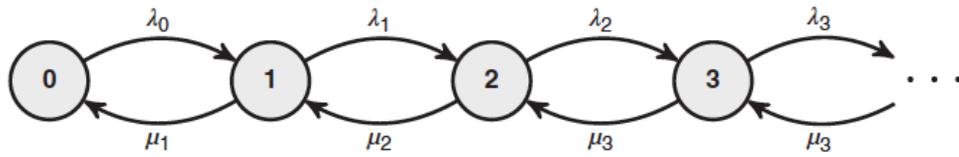


Figure 2: transition diagram of birth-death process

As we did for the discrete-time model, we can introduce the carrying capacity,  $K$  into the model, so that the population has a logistic growth. First we consider two functional forms of birth rate and death rate. The first one has linear birth rate, and the death rate increases as the population size increases. While the second one has linear death rate, and the birth rate decreases as the population size increases.

- $\lambda_i = i\lambda$  and  $\mu_i = i(\mu + \beta i)$
- $\lambda_i = i\lambda e^{-\beta i}$  and  $\mu_i = i\mu$

In order to have carrying capacity  $K$  being the value at which the process is critical, solve  $\beta$  in the equation

$$\lambda_K = \mu_K,$$

then

- $\lambda_i = i\lambda$  and  $\mu_i = i(\mu + \frac{\lambda - \mu}{K}i)$
- $\lambda_i = i\lambda(\frac{\mu}{\lambda})^{\frac{i}{K}}$  and  $\mu_i = i\mu$ .

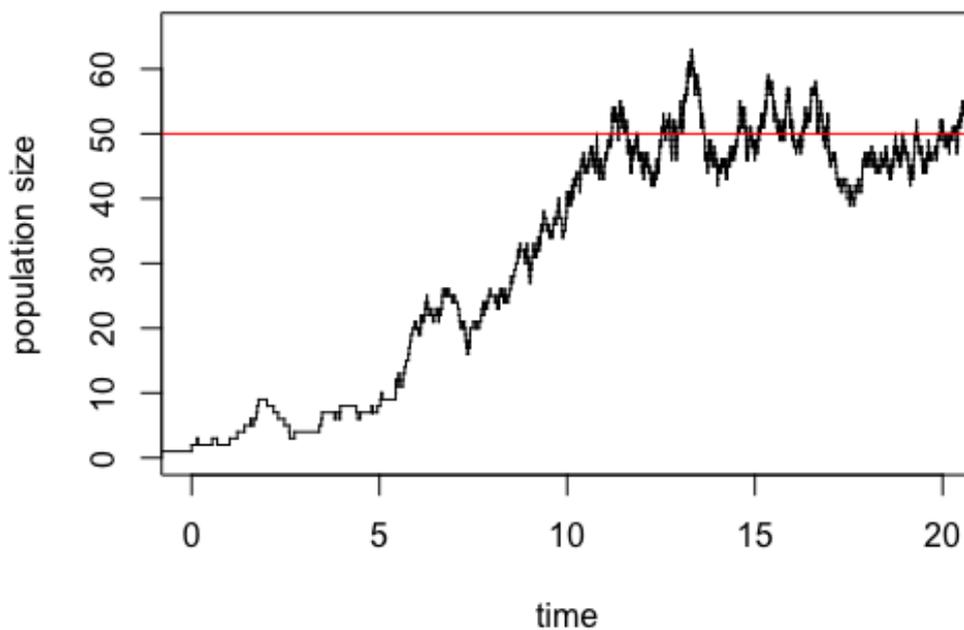


Figure 3: A trajectory with carrying capacity  $K = 50$



Recall how the process evolves, the chain stays in state  $i$  for an exponentially distributed time with rate  $\lambda_i + \mu_i$ , and then moves up or down with probability  $\frac{\lambda_i}{\lambda_i + \mu_i}$  and  $\frac{\mu_i}{\lambda_i + \mu_i}$ , respectively. Hence we can write out the likelihood function of this process,

$$L(\lambda, \mu, K) = \prod_{j=1}^n \left( \frac{\lambda_{z_j}}{\lambda_{z_j} + \mu_{z_j}} \right)^{\mathbb{1}(z_{j+1} > z_j)} \left( \frac{\mu_{z_j}}{\lambda_{z_j} + \mu_{z_j}} \right)^{\mathbb{1}(z_{j+1} < z_j)} (\lambda_{z_j} + \mu_{z_j}) \exp(-(\lambda_{z_j} + \mu_{z_j})t_{z_j}),$$

where  $t_{z_j}$  is the holding time for how long the process stays in state  $z_j$  before it jumps to  $z_{j+1}$ .

Let  $up_i$  and  $down_i$  be the total number of times the process jumps up and down while in state  $i$ , respectively. And let  $t_i$  be the total time that the process stays in state  $i$ . Let  $S$  be all states the process visited. Then, the likelihood can be rewritten as,

$$L(\lambda, \mu, K) = \prod_{i \in S} \lambda_i^{up_i} \mu_i^{down_i} \exp\left(-\sum_{i \in S} (\lambda_i + \mu_i)t_i\right).$$

## 2 Parameter estimation

### 2.1 Maximum likelihood estimation

For discrete-time models, maximum likelihood estimation is easy as we generally have observations of the population size at each generation. And the likelihood function can be optimized easily. However, in the continuous-time case, data of continuous observation is rare as recording population size at any time is expensive. Usually, we only have population size in annual basis. Nevertheless, it is sometimes possible to have full observations. For theoretical purpose and to compare with other estimators, we will examine maximum likelihood estimators on both discrete-time and continuous-time models.

In the discrete-time case, we used the binary splitting model given in (†). We took  $v$  from  $\{0.8, 1\}$  and  $K$  from  $\{50, 100, 200\}$ . We considered all combinations, and simulated 500 generations for each trajectory. 1000 simulated trajectories were conducted for each pair. Then we estimated  $v$  and  $K$  by maximum likelihood estimation. The result is given by Table 1.

		Mean $\hat{v}$	Std.Dev. $\hat{v}$	Mean $\hat{K}$	Std.Dev. $\hat{K}$
$v = 0.8$	$K = 50$	0.8105	0.0403	49.9802	0.8486
	$K = 100$	0.8096	0.0360	99.9962	1.1993
	$K = 200$	0.8074	0.0316	200.0397	1.6996
$v = 1$	$K = 50$	0.9840	0.0267	50.0273	0.6563
	$K = 100$	0.9855	0.0234	100.0364	0.8832
	$K = 200$	0.9891	0.0183	200.0046	1.2477

Table 1: Summary statistics for Discrete-time models

For continuous-time models, we chose the first setting above, i.e.:  $\lambda_i = i\lambda$  and  $\mu_i = i(\mu + \frac{\lambda - \mu}{K}i)$ . We set  $\lambda = 1.5, \mu = 0.5$ , and took  $K$  from  $\{50, 80, 150\}$ , and we again performed 1000 simulations for each parameters set. The maximum likelihood estimators computed by using full observations are close to the true parameters. The result is given by Table 2.



	Mean $\hat{\lambda}$	Std.Dev. $\hat{\lambda}$	Mean $\hat{\mu}$	Std.Dev. $\hat{\mu}$	Mean $\hat{K}$	Std.Dev. $\hat{K}$
$K = 50$	1.4985	0.0430	0.4436	0.1598	49.7685	3.1809
$K = 80$	1.4968	0.0358	0.4637	0.1278	79.7099	4.0910
$K = 150$	1.4976	0.0248	0.4745	0.0988	149.5125	5.7603

Table 2: Summary statistics for Continuous-time models

In order to deal with discrete observation of a continuous-time model, we developed a parameter estimation method via an embedded discrete-time linear fractional branching process. We then compared the maximum likelihood estimators based on full observations with the new estimators.

## 2.2 Linear fractional branching process

Consider a linear birth-death process, with birth rate  $\lambda$  and death rate  $\mu$  per individual. Each individual in this process waits for an exponential time with rate  $\lambda + \mu$ , and then either gives birth with probability  $\frac{\lambda}{\lambda + \mu}$  or dies with complementary probability  $\frac{\mu}{\lambda + \mu}$ . From the point of view of population size, the chain jumps up from  $i$  to  $i + 1$  with rate  $i\lambda$ , or decreases to  $i - 1$  with rate  $i\mu$ , i.e.:  $\lambda_i = i\lambda$ ,  $\mu_i = i\mu$ .

Suppose we can only observe a trajectory at discrete times instead of continuously. Although we lose information of the process that lies between any two consecutive observed times, we can get the maximum likelihood estimators as long as we know the distribution that the discretely observed data follows.

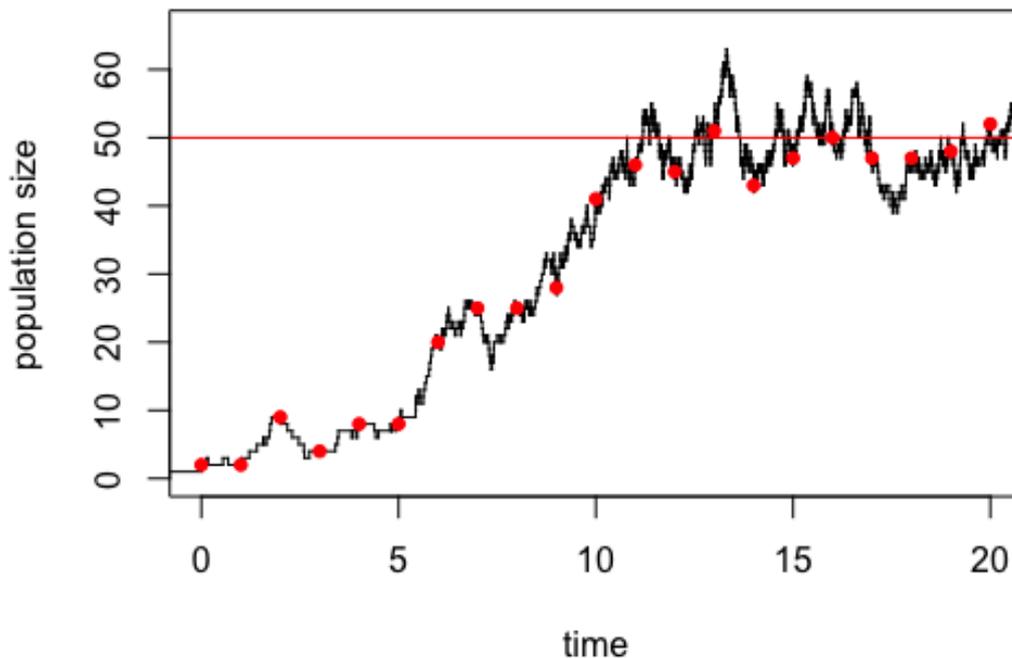


Figure 4: A trajectory of a continuous process observed at discrete times (red dots)



Let  $Z(t)$  be the population size at time  $t$  in a linear birth-death process. Define  $p_{a,k}(t) = \mathbb{P}(Z(t) = k | Z(0) = a)$  for  $k = 0, 1, \dots$ . The corresponding probability generating function (PGF) is

$$F_a(s, t) = \sum_{k \geq 0} p_{a,k}(t) s^k = F_1(s, t)^a.$$

The last equality follows since the offspring of all individuals are independent and identically distributed, and thus the population size at time  $t$  with the initial population size  $Z(0) = a$ , is equivalent to the sum of the population sizes at time  $t$  of  $a$  branching processes with initial population size  $Z(0) = 1$ .

Let  $F(s, t) = F_1(s, t) = \sum_{k \geq 0} p_{1,k}(t) s^k$ . Consider the backward Kolmogorov equation,

$$p'_{i,j}(t) = \mu_i p_{i-1,j} - (\mu_i + \lambda_i) p_{i,j} + \lambda_i p_{i+1,j}.$$

Then,

$$\begin{aligned} \frac{\partial F}{\partial t} &= \sum_{k \geq 0} p'_{1,k}(t) s^k \\ &= \sum_{k \geq 0} (\mu_1 p_{0,k} - (\mu_1 + \lambda_1) p_{1,k} + \lambda_1 p_{2,k}) s^k \\ &= \mu - (\mu + \lambda) F(s, t) + \lambda F^2(s, t) \end{aligned}$$

This is an ordinal differential equation with respect to  $t$ . It has the form of a Riccati differential equation. The explicit solution is

$$F(s, t) = \begin{cases} 1 + \frac{(\lambda - \mu)(s - 1)}{(\lambda s - \mu) \exp((\mu - \lambda)t) - \lambda(s - 1)}, & \lambda \neq \mu \\ 1 + \frac{s - 1}{1 - \lambda t(s - 1)}, & \lambda = \mu \end{cases}.$$

The probability generating function above corresponds to a modified geometric distribution, which is a mixture of a point mass at 0 and geometric distribution,

$$p_{1,k}(t) = p_0(t) \mathbb{1}\{k = 0\} + (1 - p_0(t))(1 - p(t))p(t)^{k-1} \mathbb{1}\{k > 0\},$$

where

$$p_0(t) = \begin{cases} \frac{\mu(e^{\omega t} - 1)}{\lambda e^{\omega t} - \mu}, & \lambda \neq \mu \\ \frac{\lambda t}{1 + \lambda t}, & \lambda = \mu \end{cases}, \quad p(t) = \begin{cases} 1 - \frac{\lambda}{\mu} p_0(t), & \lambda \neq \mu \\ 1 - p_0(t), & \lambda = \mu \end{cases}$$

Conditionally on  $Z(0) = a$ ,  $\mathbb{P}(Z(t) = k | Z(0) = a) = p_0(t)^a$  by independence of the individuals. For  $k > 0$ , by conditioning on the number of ancestors whose descendance becomes extinct by time  $t$ , we have

$$\begin{aligned} &\mathbb{P}(Z(t) = k | Z(0) = a) \\ &= \sum_{j=\max(0, a-k)}^{a-1} \binom{a}{j} \binom{k-1}{a-j-1} p_0^j \{(1 - p_0)p\}^{a-j} (1 - p)^{k-a+j}, \end{aligned}$$

where  $p_0 = p_0(t)$ ,  $p = p(t)$ .

### 2.3 Estimators via linear fractional branching process

Now that we have the distribution that the discretely observed data follows, we can treat the observed process as a standard Galton-Watson process. Such a process is called the *embedded* linear fractional branching process.



However, the underlying continuous linear birth-death process is population size independent. Once we introduce carrying capacity,  $K$ , the theory above does not apply.

To deal with this issue, one way is to introduce carrying capacity,  $K$  in a different way. Let us use the first case of the setting above,  $\lambda_i = i\lambda$  and  $\mu_i = i(\mu + \frac{\lambda - \mu}{K}i)$ . However, in this time, the birth and death rates per individual in the underlying model do not change whenever there is a jump. Instead, they would only be updated at the observed time, but still according to the same formulae,  $\lambda_i = i\lambda$  and  $\mu_i = i(\mu + \frac{\lambda - \mu}{K}i)$ .

Hence, between any two consecutive observed times, the birth and death rates per individual are kept constant. Equivalently, the chain is a concatenation of a sequence of consecutive standard linear birth-death processes with different rates. It can also be viewed as if the population stayed constant and jumped instantly at the observed times, and the birth and death rates change as usual.

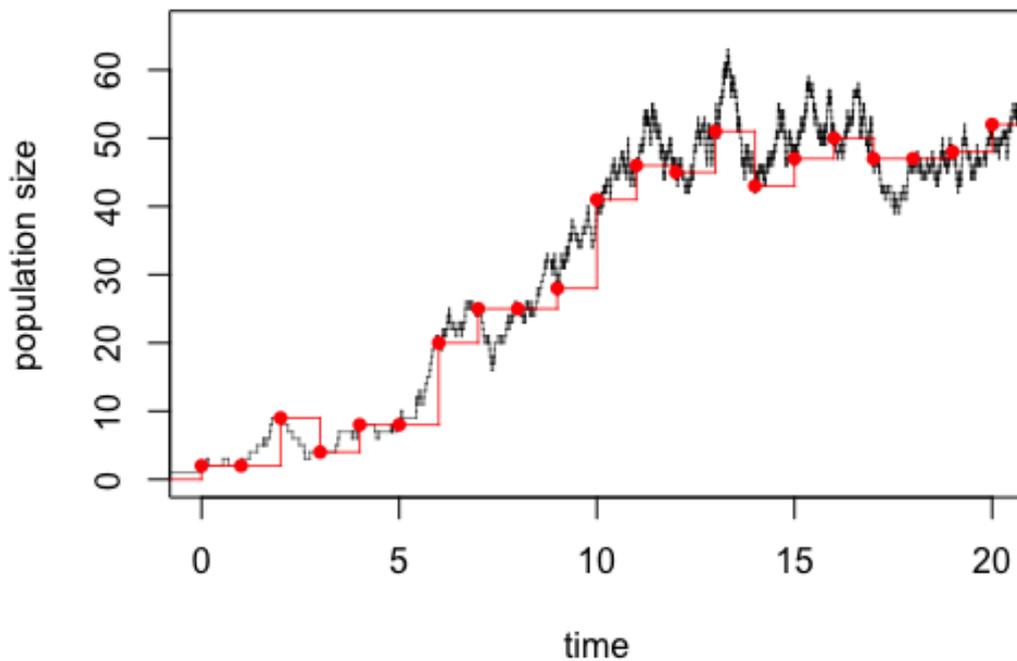


Figure 5: The birth and death rates are kept constant

In the above setting, the carrying capacity,  $K$ , can be estimated by maximum likelihood estimation using discretely-observed data. And in this setting, there is no approximation involved since each inter-observed period is a linear fractional process.

But if we insist on the original setting that the birth and death rates per individual change in the underlying model whenever there is a jump, and implement the estimation method suggested above, then it will be an approximation since we treat each inter-observed period as a linear fractional process while it is not.

To be consistent with the usual setting of population size dependent branching process that the birth rate  $\lambda$  and death rate  $\mu$  per individual change whenever there is a jump, we decided to use the regular setting and said the estimation method suggested above is an approximation.



There are different schemes of approximation. For one inter-period of the process, as we only observed the start and end points of the trajectory, a reasonable approximation would set the population in this period being the average of two observed population sizes. If the blind period between two consecutive observation time is small enough in the sense that the population size has monotonic trend, then the approximation will be close to maximum likelihood estimation using data of full trajectory. If the blind period is too wide, then the population size might experience large increase or decrease so that the average of two end points cannot represent the period accurately.

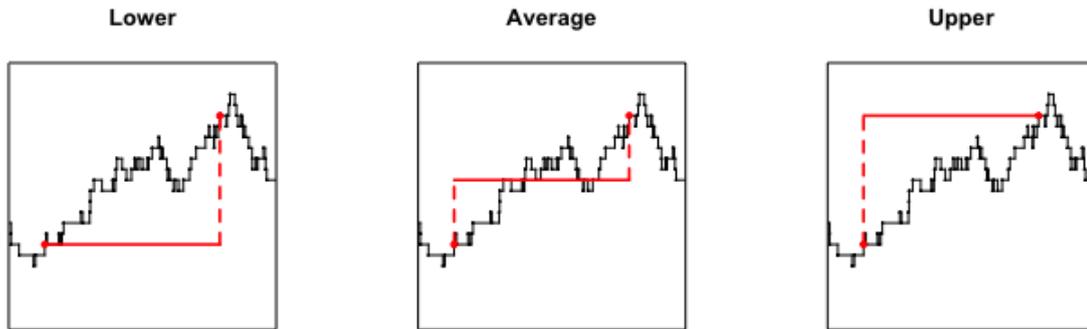


Figure 6: Different schemes of approximation

If we approximate the population size by the smaller value of the two end points, then the value of  $K$  tends to be underestimated. While if we approximate the population size by the larger value of the two end points, then the value of  $K$  tends to be overestimated. The intuition is that, if we use smaller population size, then we would expect a larger growth rate. But the population size in the observed data grows with smaller rate as population size is larger in reality. To compensate this smaller growth, we need a smaller carrying capacity  $K$ . Since the carrying capacity,  $K$  has an adverse effect on the growth rate. the smaller  $K$ , the stronger effect it suppress the growth process. Therefore, the approximation by smaller population size would imply a smaller carrying capacity  $K$ . It is easier to see when we use an specific example. In our setting,  $\mu_i = i(\mu + \frac{\lambda - \mu}{K}i)$ . To retain the same death rate (or the growth rate since the birth rate is independent of  $K$  in this case), a smaller population size  $i$  results in a smaller  $K$ , and *vice versa*.

The argument above suggests a lower and upper bound for the value of the carrying capacity  $K$ . More specifically, let  $\hat{K}_{\text{lower}}$  and  $\hat{K}_{\text{upper}}$  be the estimators arised by approximation using smaller and larger population sizes respectively, and let  $\hat{K}_{\text{MLE}}$  be the maximum likelihood estimator computed using full continuous observation. We expected to have

$$E(\hat{K}_{\text{lower}}) \leq E(\hat{K}_{\text{MLE}}) \leq E(\hat{K}_{\text{upper}})$$

The theoretical result has not been proved, yet here we just show the simulation result.

Consider a continuous time population size dependent birth-death process. Let  $\lambda = 1.5$ ,  $\mu = 0.5$  and  $K = 50$ . When the popultion size is in state  $i$ , the birth rate and death rate are given by

$$\lambda_i = i\lambda \quad \text{and} \quad \mu_i = i(\mu + \frac{\lambda - \mu}{K}i).$$



We generated 1000 trajectories. We found the estimates for  $\lambda$ ,  $\mu$  and  $K$  resulting from approximation by lower, average and upper population sizes for each trajectory, together with maximum likelihood estimate using full continuous observation.

For the carrying capacity,  $K$ , we plot the histograms of these four estimates,  $\hat{K}_{\text{lower}}$ ,  $\hat{K}_{\text{MLE}}$ ,  $\hat{K}_{\text{avg}}$ , and  $\hat{K}_{\text{upper}}$ , to show how they are distributed. Table 3 gives the summary statistics for the estimates resulting from different schemes of approximation. The means of  $\hat{K}_{\text{lower}}$ ,  $\hat{K}_{\text{MLE}}$  and  $\hat{K}_{\text{upper}}$  are 48.0558, 49.6654 and 60.8660 respectively. Hence, in this example, the relation in expectations described above,  $E(\hat{K}_{\text{lower}}) \leq E(\hat{K}_{\text{MLE}}) \leq E(\hat{K}_{\text{upper}})$ , was verified numerically.

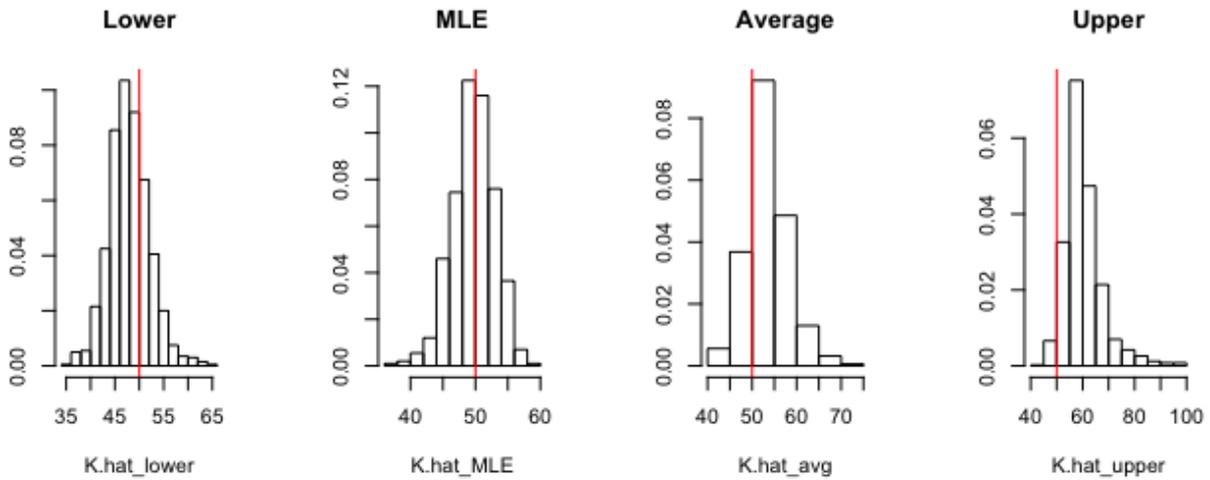


Figure 7: Histograms of  $\hat{K}$  arised from different schemes of approximation

	Mean $\hat{\lambda}$	Std.Dev. $\hat{\lambda}$	Mean $\hat{\mu}$	Std.Dev. $\hat{\mu}$	Mean $\hat{K}$	Std.Dev. $\hat{K}$
Lower	2.6668	0.7777	1.5977	0.5391	48.0558	4.2215
<b>MLE</b>	<b>1.4963</b>	<b>0.0440</b>	<b>0.4590</b>	<b>0.1473</b>	<b>49.6654</b>	<b>3.2495</b>
Average	2.9757	1.2454	1.8525	0.7084	53.6998	4.7991
Upper	2.7680	1.1678	1.8597	0.6743	60.8660	7.5803

Table 3: Summary statistics for the estimates resulting from different schemes of approximation

We see that the estimate for the birth rates  $\hat{\lambda}$  and the death rates  $\hat{\mu}$  arised from the approximations are far from the true values and maximum likelihood estimates based on full observations. This issue indicates that the birth rate  $\lambda$  and death rate  $\mu$  are sensitive to the approximation of the population sizes. This is because the birth rates and death rates have opposite effects on the population size, which will cancel each other in a period of time. And we do not have the information in the blind periods that lie between two consecutive observed time, so the birth rates  $\hat{\lambda}$  and the death rates  $\hat{\mu}$  arised from the approximations are unreliable.



However, note that the differences between the estimate for the birth rates  $\hat{\lambda}$  and the death rates  $\hat{\mu}$ , i.e. the growth rates, denoted as  $\hat{\omega}$ , are around the true value 1, in all cases. This is consistent with the growth trend of the trajectory that shows how faster the population size reaches the carrying capacity  $K$ .

Even though the birth rate  $\hat{\lambda}$  and death rate  $\hat{\mu}$  are off the track, the carrying capacity  $\hat{K}$  is not affected. This is because the carrying capacity  $K$  only controls the threshold value of population size that shows the limiting resources of the environment. Despite we only observe population size at discrete time, we still have trajectories lingering around the true value of the carrying capacity  $K$ . Therefore, the estimators of  $K$  arised from approximation should still reflect this information, and thus be close to the true value of  $K$ .

Hence, we concluded that in the parameter estimation method developed above, the difference of the birth rates  $\hat{\lambda}$  and the death rates  $\hat{\mu}$ , i.e. the growth rates  $\hat{\omega}$ , and the carrying capacity,  $\hat{K}$  provide satisfactory estimators, while the birth rates  $\hat{\lambda}$  and the death rates  $\hat{\mu}$  individually are unreliable.

## 2.4 Application to the black robin population

In the end, we applied this method to the black robin population. We obtained  $\hat{K}_{\text{lower}} = 179.97$ ,  $\hat{K}_{\text{avg}} = 198.91$  and  $\hat{K}_{\text{upper}} = 229.50$ .

$\hat{K}_{\text{lower}}$  and  $\hat{K}_{\text{upper}}$  serve as lower and upper bounds for the carrying capacity  $K$  of the black robin population. And  $\hat{K}_{\text{avg}}$  serves as a point estimate for the carrying capacity  $K$  of the black robin population.

In Figure 8, we plot a trajectory generated using the point estimates resulting from the approximation by the average point.

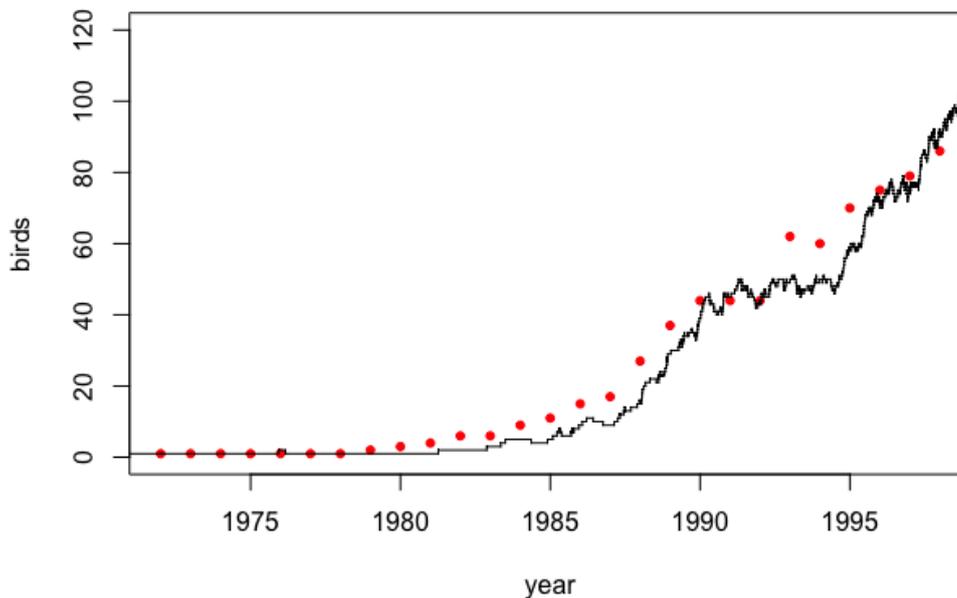


Figure 8: Trajectory generated by parameters estimates using average approximation



## References

- Davison, A. C., S. Hautphenne, and A. Kraus (2018). Parameter estimation for discretely-observed linear birth-and-death processes. *arXiv preprint arXiv:1802.05015*.
- Dobrow, R. P. (2016). *Introduction to Stochastic Processes with R*. John Wiley & Sons, Inc., Hoboken, New Jersey.
- Haccou, P., P. Jagers, and V. A. Vatutin (2005). *Branching processes: variation, growth, and extinction of populations*. Number 5. Cambridge university press.
- Sagitov, S. (2013). Linear-fractional branching processes with countably many types. *Stochastic processes and their applications* 123(8), 2940–2956.