Estimation in branching processes with restricted observations

Ronald Meester

Vrije Universiteit, Amsterdam, The Netherlands

Pieter Trapman

Utrecht University, Utrecht, The Netherlands

Vrije Universiteit, Amsterdam, The Netherlands

Summary. In this paper we are interested in consistent estimators for (functions of the) parameters in supercritical Galton-Watson branching processes, when we only have partial observations.
If the full generation sizes are observed, it is known that one can only estimate two moments of the offspring distribution consistently on the explosion set of the Galton-Watson process. In our context we observe a random subset of each generation, in such a way that each element is observed with unknown probability μ, independently of each other. The offspring distribution of an observed element changes upon being observed. This idea has an interpretation and application in epidemiology, where the start of a simple epidemic can be described by a Galton-Watson process, and where the only observations are those individuals that are recognized as being infected, and which subsequently produce no further offspring, or produce a reduced offspring in the time interval of detection.
We prove that under certain conditions, we can estimate the mean of the offspring distribution of an individual and some other function which can be interpreted as a second moment when we ignore dependencies.
Somewhat surprisingly, if our observations do have further offspring, we are able to estimate a third function of the parameters consistently. This often enables us to estimate μ consistently. We also discuss rates of convergence of the various estimators.
We apply the estimators to data of real epidemics and computer simulations, and find that we can estimate the first moment of the offspring distribution very well, while the estimator for the second and third function of the parameters converges very slowly. The conclusion is that what is theoretically possible, is not always practically feasible.

1. Introduction

It is known that it is impossible to consistently estimate more than two moments of the offspring-distribution in a supercritical Galton-Watson process if only the generation sizes X_n of the process are observed, (Theorem 1.3 of Guttorp it et al. (1991)). However, it is not a-priori clear what one can estimate consistently in a Galton-Watson process if only a random subset of each generation is observed, and the observation of an element (possibly) changes its offspring distribution.

For X_n, the generation sizes and μ_n, a known sequence, converging to μ, Jacob it et al. (1998) have shown that if the observations are binomially distributed with parameters

address for correspondence: Pieter Trapman, Faculty of Veterinary Medicine, Utrecht University, Yalelaan 7, 3584 CL Utrecht, The Netherlands. E-mail: **trapman@vet.uu.nl**

X_n and μ_n and the offspring distribution has a finite fourth moment, it is also possible to estimate the first two moments of the offspring distribution consistently on the explosion set (i.e. the set where $\lim_{n\to\infty} X_n = \infty$). Here the observed individuals may produce offspring, but the offspring of these individuals is distributed like the offspring of L unobserved individuals, where L is a non-negative integer.

Our set-up differs in two aspects from Jacob it et al. (1998). First, we are interested in the case where $\mu_n = \mu$ is constant but unknown. Furthermore, we assume that the offspring distribution of unobserved and observed individuals have a finite fourth moment, but no further assumptions are made about the offspring distributions. Our methods are also quite different; our martingales are based on observable quantities. Besides proving some results analogous to Jacob it et al. (1998), we also show that μ, unknown in our case, can under certain circumstances be estimated consistently.

The kind of partial observations we are dealing with is especially interesting for estimation in epidemics of infectious diseases. If the number of susceptible individuals (where an individual may also refer to a herd instead of an individual animal) is very large, we may describe the start of an idealized epidemic by a Galton-Watson process, where discrete points in time index the generations (see e.g. Andersson it et al. (2000) or Trapman it et al. (2004)). As soon as an infectious disease is observed in an individual, it stops being infective because of isolation (in the case of human infections) or culling (in the case of very contagious animal diseases like classical swine fever, foot and mouth disease or avian influenza). So in the time interval of detection, the individual was only infectious during a fraction of the interval length which implies that observed individuals have reduced offspring. In this epidemiological setting, an individual that is infectious but not detected in a certain generation, will still be infectious in the next one, so a surviving individual will cause at least one infective individual in the next generation, namely itself. Our task is to estimate parameters of the offspring distribution using only this partial information. This interpretation should be compared to the work in Becker it et al. (1997). In that paper, estimation takes place under the assumption that one also observes, in addition to individuals without further offspring, the total number of infectious individuals at the beginning, and at the end of the observation period.

Our main interest is in estimating the offspring mean and the parameter μ, because those parameters are extremely important for decisions about measures to be taken to stop an epidemic. We are able to estimate the offspring mean very efficiently. On the explosion set, we are (under certain conditions) able to consistently estimate two other functions of μ and the parameters of the offspring distribution. These three estimators will lead to a system of three equations. For many models we also have three unknowns, namely the offspring mean, the offspring variance and μ. In principle, we can therefore often estimate these parameters.

In the next section we set things up formally. In Section 3 we introduce two consistent estimators for the offspring mean, and in Section 4 we estimate a second function of the parameters. This function can be interpreted as a second moment, as we will explain. In Section 5 we give the estimator for a third function of parameters. Finally, we apply our results to simulated data, as well as to real data from the 1997 epidemic of classical swine fever in The Netherlands. It turns out that what is theoretically possible, is not always practically feasible due to extreme slow convergence of estimators.

2. Formal set-up

Throughout, (X_n) represents the generation sizes of a supercritical Galton-Watson branching process. We interpret this branching process as the number of infected individuals at the discrete time instants n, $n = 0, 1, 2, \ldots$. The X_n are not observable. Between time n and $n+1$, a certain (random) number of the infected individuals is detected; we assume that each infected individual is detected with probability μ during this time interval, independently of each other. The parameter μ is unknown. The number of such detected individuals between time n and time $n+1$ is denoted by Z_{n+1}, and this random quantity is observable. So, given X_n, Z_{n+1} has a binomial distribution with parameters X_n and μ. The Z-individuals produce no further offspring, but the X-individual producing the Z-individual may produce offspring in the interval it is detected (and so produce some new X-individuals and precisely one Z-individual). Each infected individual at time n leads to at most one Z-individual at the next time $n+1$, and to a random number of X-individuals at time $n+1$. If we assume that an infectious individual can only be removed from the process by being detected, the offspring of an undetected X-individual consists of at least one element, namely itself.

We denote by m_+ the expected number of X-offspring of an infected individual at time n, given that it is not detected between time n and time $n+1$. Similarly, m_- is the expected number of X-offspring of an infected individual given that it is detected. The corresponding variances are denoted by σ_+^2 and σ_-^2 respectively. In formulas, this reads as follows:

$$
\begin{aligned}
m_+ &:= \mathbb{E}(X_1|X_0 = 1, Z_1 = 0), \\
m_- &:= \mathbb{E}(X_1|X_0 = 1, Z_1 = 1), \\
\sigma_+^2 &:= \mathbb{E}((X_1 - (m_+))^2|X_0 = 1, Z_1 = 0), \\
\sigma_-^2 &:= \mathbb{E}((X_1 - (m_-))^2|X_0 = 1, Z_1 = 1).
\end{aligned}
$$

Finally, we write m for the unconditional expected number of X-individuals of an infected individual:

$$m = \mathbb{E}(X_1|X_0 = 1) = (1-\mu)m_+ + \mu m_-.$$

Similarly, the unconditional variance is denoted by σ^2. We have from Lemma 2.1 of Guttorp it et al. (1991) that for a random variable Y, an event F and F^c the complement of F,

$$
\begin{aligned}
Var(Y) &= \mathbb{P}(F)Var(Y|F) + \mathbb{P}(F^c)Var(Y|F^c) + \\
&\quad +(\mathbb{E}(Y|F) - \mathbb{E}(Y|F^c))^2\mathbb{P}(F)\mathbb{P}(F^c).
\end{aligned}
$$

Applying this with $F = \{Z_1 = 0\}$ yields

$$\sigma^2 = Var(X_1|X_0 = 1) = (1-\mu)\sigma_+^2 + \mu\sigma_-^2 + (m_+ - m_-)^2\mu(1-\mu).$$

Define $A := \{W > 0\}$, where W is defined as $\lim_{n\to\infty} \frac{X_n}{m^n}$. Assuming a finite offspring variance, we have that A differs by a null set from the explosion set $\{\lim_{n\to\infty} X_n > 0\}$ (Theorem 1.1 of Guttorp it et al. (1991)). Because A is a tail-event, conditioning on this event is strictly speaking not proper for estimation purposes, but analyzing the behavior of the process on the set A seems necessary, for only on this set we obtain infinitely many observations.

3. Estimating the offspring mean m

In this section we discuss two consistent estimators for m. Perhaps the first estimator that comes to mind is

$$\bar{m}_n := \frac{Z_{n+1}}{Z_n},$$

and our first result states that this estimator is indeed consistent on the explosion set A.

THEOREM 3.1. *We have that*

$$\bar{m}_n \to m \qquad a.s.\ on\ A.$$

Proof: The proof is based on a simple martingale argument. Let

$$M_n := \prod_{i=0}^{n} \frac{Z_{i+1}+1}{\mu X_i + 1}$$

Note that this is a (positive) martingale with respect to $\mathcal{F}_n$, the σ-algebra generated by $\{Z_{i+1}, X_i; 0 \le i \le n\}$. Since $\sup_n \mathbb{E}(M_n) \le 1$, the martingale convergence theorem implies that M_n converges almost surely to an almost surely finite random variable M.

We also need to show that M is strictly positive on A. To do this, we define

$$\bar{M}_n := \prod_{i=0}^{n} \frac{\mu(X_i+1)}{Z_{i+1}+1}.$$

Elementary computations yield

$$\begin{aligned} \mathbb{E}\Big[\tfrac{1}{Z_1+1}|X_0=k\Big] &= \textstyle\sum_{i=0}^{k} \frac{1}{i+1}\binom{k}{i}\mu^i(1-\mu)^{k-i} \\ &= \tfrac{1}{k+1}\textstyle\sum_{i=0}^{k}\frac{1}{\mu}\binom{k+1}{i+1}\mu^{i+1}(1-\mu)^{k-i} \\ &= \tfrac{1}{\mu(k+1)}(1-(1-\mu)^{k+1}) \\ &\le \tfrac{1}{\mu(k+1)}, \end{aligned} \tag{1}$$

so $\bar{M}_n$ is a supermartingale with respect to $\mathcal{F}_n$. By the martingale convergence theorem we know that $\bar{M}_n$ converges almost surely to an almost surely finite random variable $\bar{M}$. Now write

$$M_n\bar{M}_n = \prod_{i=0}^{n}\frac{\mu(X_i+1)}{\mu X_i+1} = \prod_{i=0}^{n}\Big(1-\frac{1-\mu}{\mu X_i+1}\Big).$$

The X_i's almost surely grow exponentionally on A, so

$$\sum_{i=1}^{\infty}\frac{1-\mu}{\mu X_i+1} < \infty$$

almost surely on A, which implies that

$$M\bar{M} = \lim_{n\to\infty} M_n\bar{M}_n = \prod_{i=0}^{\infty}\Big(1-\frac{1-\mu}{\mu X_i+1}\Big) > 0, \tag{2}$$

almost surely on A. Because $\bar{M}$ is almost surely finite, (2) is only possible if M is almost surely positive on A.

Now since $\frac{X_{i+1}}{X_i} \to m$ a.s. on A (Theorem 2.1 of Guttorp it et al. (1991)), we have that $\frac{X_{n+1}+\mu^{-1}}{X_n+\mu^{-1}} \to m$ a.s. on A. So we have

$$\frac{Z_{n+1}+1}{Z_n+1} = \frac{\frac{M_n}{M_{n-1}}}{\frac{M_{n-1}}{M_{n-2}}} \frac{\mu X_n + 1}{\mu X_{n-1}+1} \to m, \qquad \text{a.s. on } A,$$

because from $0 < M < \infty$ a.s. on A, it follows that

$$\frac{M_n}{M_{n-1}} \to \frac{M}{M} = 1, \qquad \text{a.s. on } A.$$

On A, Z_n will almost surely tend to infinity, so $\frac{Z_{n+1}+1}{Z_n+1}$ will have the same limit as $\frac{Z_{n+1}}{Z_n}$, proving the theorem. □

Clearly, $\bar{m}_n$ does not use all the available information. To this end, we introduce another estimator, namely

$$\tilde{m}_n := \frac{\sum_{i=2}^{n+1} Z_i}{\sum_{i=1}^{n} Z_i}.$$

We now prove that $\tilde{m}_n$ also is a consistent estimator for m:

THEOREM 3.2. *On the explosion set A we have a.s., as $n \to \infty$,*

$$\tilde{m}_n \to m.$$

To prove this theorem, we start with a lemma from Hall it et al. (1980).

LEMMA 3.3. *Let $(S_n = \sum_{i=1}^n \xi_i, \mathcal{F}_n, n \geq 1)$ be a martingale and let $(U_n, n \geq 1)$ be a nondecreasing sequence of positive random variables such that U_n is $\mathcal{F}_{n-1}$-measurable. Then $U_n^{-1} S_n \to 0$ a.s. on the set*

$$\left\{ \lim U_n = \infty, \sum_{i=1}^{\infty} U_i^{-2} \mathbb{E}(\xi_i^2 | \mathcal{F}_{i-1}) < \infty \right\}.$$

COROLLARY 3.4. *On the explosion set A we have a.s. as $n \to \infty$,*

$$\frac{\sum_{i=1}^n Z_i}{\sum_{i=0}^{n-1} X_i} \to \mu.$$

This corollary is intuitively obvious because the denominator is the total number of individuals in the first n generations (including generation 0), while the numerator is the number of observed individuals in these generations. Here is a formal proof.

Proof of Corollary 3.4: Write

$$\frac{\sum_{i=1}^n Z_i}{\sum_{i=0}^{n-1} X_i} = \frac{\sum_{i=1}^n (Z_i - \mu X_{i-1})}{\sum_{i=0}^{n-1} X_i} + \mu.$$

We define $U_n = \sum_{i=0}^{n-1} X_i$, $\xi_i = Z_i - \mu X_{i-1}$ and $S_n = \sum_{i=1}^{n} \xi_i$. Note that U_n is $\mathcal{F}_{n-1}$-measurable, where $\mathcal{F}_{n-1}$ is the σ-algebra generated by $X_0, X_1, \ldots, X_{n-1}, Z_1, \ldots, Z_{n-1}$. Furthermore,

$$\begin{aligned}\sum_{i=1}^{\infty} U_i^{-2}\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1}) &= \sum_{i=1}^{\infty} \frac{\mathbb{E}((Z_i-\mu X_{i-1})^2|\mathcal{F}_{i-1})}{(\sum_{j=0}^{i-1} X_j)^2} \\ &= \sum_{i=1}^{\infty} \frac{\mu(1-\mu)X_{i-1}}{(\sum_{j=0}^{i-1} X_j)^2} \le \sum_{i=1}^{\infty} \frac{\mu(1-\mu)}{X_{i-1}}.\end{aligned}$$

This last sum is almost surely finite on A, because X_i is strictly positive and almost surely grows exponentially in i. So the set $\{\lim U_n = \infty, \sum_{i=1}^{\infty} U_i^{-2}\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1}) < \infty\}$ contains A up to a set of measure zero. Now we may apply Lemma 3.3 and conclude that

$$\frac{\sum_{i=1}^{n} Z_i}{\sum_{i=0}^{n-1} X_i} \to 0 + \mu = \mu \qquad \text{a.s. on } A.$$

as $n \to \infty$. □

Proof of Theorem 3.2: From Theorem 2.1 of Guttorp it et al. (1991) we have $\frac{\sum_{i=1}^{n} X_i}{\sum_{i=0}^{n-1} X_i} \to m$ a.s. on A. We apply Corollary 3.4, giving

$$\frac{\sum_{i=2}^{n+1} Z_i}{\sum_{i=1}^{n} Z_i} = \frac{\frac{\sum_{i=2}^{n+1} Z_i}{\sum_{i=1}^{n} X_i}}{\frac{\sum_{i=1}^{n} Z_i}{\sum_{i=0}^{n-1} X_i}} \frac{\sum_{i=1}^{n} X_i}{\sum_{i=0}^{n-1} X_i} \to \frac{\mu}{\mu} m = m, \qquad \text{a.s. on } A.$$

□

We are also interested in the rate of convergence of $\tilde{m}_n$. This rate follows from the next theorem.

Theorem 3.5. *The random variables*

$$\Big(\sum_{i=1}^{n} Z_i\Big)^{\frac{1}{2}} (\tilde{m}_n - m)$$

converge in distribution to a sum of three normal random variables with zero mean and finite variance.

This theorem implies that $(\tilde{m}_n - m)$ converges with a rate of order $\Big(\sum_{i=1}^{n} Z_i\Big)^{-\frac{1}{2}}$.

We use a part of Theorem 2.3 of Guttorp it et al. (1991) as a lemma to prove Theorem 3.5.

Lemma 3.6. *Assume that $m > 1$ and let Y be a standard normal random variable, independent of (X_n). For any x we have*

$$\mathbb{P}\Big[\frac{1}{\sigma}\Big(\sum_{i=1}^{n} X_{i-1}\Big)^{\frac{1}{2}} \Big(\frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} X_{i-1}} - m\Big) \le x | X_n > 0\Big] \to \mathbb{P}(Y \le x).$$

In the same way one can prove that

$$\mathbb{P}\Big[\frac{1}{\sqrt{\mu(1-\mu)}}\Big(\sum_{i=1}^{n} X_{i-1}\Big)^{\frac{1}{2}}\Big(\frac{\sum_{i=1}^n Z_i}{\sum_{i=1}^n X_{i-1}} - \mu\Big) \leq x | X_n > 0\Big] \to \mathbb{P}(Y \leq x).$$

Proof of Theorem 3.5: First we rewrite $\Big(\sum_{i=1}^n Z_i\Big)^{\frac{1}{2}}(\tilde{m}_n - m)$ as

$$\begin{aligned} & \Big(\sum_{i=1}^n Z_i\Big)^{\frac{1}{2}}(\tilde{m}_n - m) \\ = & \Big(\sum_{i=1}^n Z_i\Big)^{-\frac{1}{2}}\sum_{i=1}^n (Z_{i+1} - mZ_i) \\ = & \Big(\sum_{i=1}^n Z_i\Big)^{-\frac{1}{2}}\Big[\sum_{i=1}^n (Z_{i+1} - \mu X_i) + \mu \sum_{i=1}^n (X_i - mX_{i-1}) \\ & - m\sum_{i=1}^n (Z_i - \mu X_{i-1})\Big] \end{aligned} \tag{3}$$

We already know that on A, $\frac{\sum_{i=1}^n Z_i}{\sum_{i=1}^n X_{i-1}}$ converges a.s. to the constant μ and $\frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n X_{i-1}}$ converges a.s. to the constant m. Now the second term on the righthand side of (3) can be rewritten as

$$\begin{aligned} & \mu\Big(\frac{\sum_{i=1}^n Z_i}{\sum_{i=1}^n X_{i-1}}\Big)^{-\frac{1}{2}}\Big(\sum_{i=1}^n X_{i-1}\Big)^{-\frac{1}{2}}\sum_{i=1}^n (X_i - mX_{i-1}) \\ & = \mu\Big(\frac{\sum_{i=1}^n Z_i}{\sum_{i=1}^n X_{i-1}}\Big)^{-\frac{1}{2}}\Big(\sum_{i=1}^n X_{i-1}\Big)^{\frac{1}{2}}\Big(\frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n X_{i-1}} - m\Big). \end{aligned}$$

From Lemma 3.6 we see that the second term converges in distribution to a normal distribution with zero mean and finite variance. We can treat the other terms on the righthand side in the same way, which proves the theorem. □

4. Estimating a second function of the parameters

Next we want to find a consistent estimator for another function of σ^2, μ and m. We first do this for the special case where $m_- = \sigma_-^2 = 0$. After that we treat the general case, and finally we interpret the function of the parameters that we can estimate.

4.1. *The case* $m_- = \sigma_-^2 = 0$

We will prove the following result, yielding a consistent (in probability) estimator of $(1-\mu)m + \mu\sigma^2 + \mu m^2 + m^2$.

THEOREM 4.1. *As* $n \to \infty$, *we have*

$$\frac{1}{n}\sum_{i=1}^{n}(Z_i+1)\left(\frac{Z_{i+1}}{Z_i+1} - \tilde{m}_n\right)^2 \to (1-\mu)m + \mu\sigma^2 + \mu m^2 + m^2$$

in probability on A.

This subsection is devoted to a proof of this result. We first compute $\mathbb{E}\left[(Z_1+1)\left(\frac{Z_2}{Z_1+1} - m\right)^2 | X_0 = k\right]$. To do this, we remember (1):

$$\mathbb{E}\Big[\frac{1}{Z_1+1}|X_0 = k\Big] = \frac{1}{\mu(k+1)}(1-(1-\mu)^{k+1}); \tag{4}$$

and note that elementary computations yield

$$\begin{aligned}
\mathbb{E}\Big[\frac{X_0 - Z_1}{Z_1+1}|X_0 = k\Big] &= \sum_{i=0}^{k} \frac{k-i}{i+1}\binom{k}{i}\mu^i(1-\mu)^{k-i} \\
&= \sum_{i=0}^{k} \frac{1-\mu}{\mu}\binom{k}{i+1}\mu^{i+1}(1-\mu)^{k-i-1} \\
&= \frac{1-\mu}{\mu}(1-(1-\mu)^k);
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}\Big[\frac{(X_0 - Z_1)(X_0-Z_1-1)}{Z_1+1}|X_0 = k\Big] &= \sum_{i=0}^{k} \frac{(k-i)(k-i-1)}{i+1}\binom{k}{i}\mu^i(1-\mu)^{k-i} \\
&= \sum_{i=0}^{k} \frac{(1-\mu)^2 k}{\mu}\binom{k-1}{i+1}\mu^{i+1}(1-\mu)^{k-i-2} \\
&= \frac{(1-\mu)^2 k}{\mu}(1-(1-\mu)^{k-1}).
\end{aligned}$$

Now we can compute the desired expectation:

$$\begin{aligned}
&\mathbb{E}\left[(Z_1+1)\Big(\frac{Z_2}{Z_1+1} - m\Big)^2|X_0 = k\right] \\
= \ & \mathbb{E}[(Z_1+1)^{-1}(Z_2 - m(Z_1+1))^2|X_0 = k] \\
= \ & \mathbb{E}[(Z_1+1)^{-1}([Z_2 - \mu X_1] + [\mu X_1 - m(Z_1+1)])^2|X_0 = k] \\
= \ & \mathbb{E}[(Z_1+1)^{-1}(Z_2 - \mu X_1)^2|X_0 = k] + \mathbb{E}[(Z_1+1)^{-1}(\mu X_1 - m(Z_1+1))^2|X_0 = k] \\
= \ & \mathbb{E}\Big[\frac{\mu(1-\mu)X_1}{Z_1+1}|X_0 = k\Big] + \mathbb{E}[(Z_1+1)^{-1}(\mu(X_1 - m_+(X_0-Z_1))^2|X_0 = k] \\
& +\mathbb{E}[(Z_1+1)^{-1}(\mu m_+(X_0-Z_1) - m(Z_1+1))^2|X_0 = k] \\
= \ & \mathbb{E}\Big[\frac{\mu((1-\mu)m_+ + \mu\sigma_+^2)(X_0-Z_1)}{Z_1+1}|X_0 = k\Big] \\
& +\mathbb{E}[(Z_1+1)^{-1}(\mu m_+(X_0-Z_1))^2|X_0 = k] \\
& +\mathbb{E}[m^2(Z_1+1) - 2\mu m m_+(X_0-Z_1)|X_0 = k] \\
= \ & (1-\mu)(m + \mu\sigma_+^2 + \mu(m_+)^2)(1-(1-\mu)^k) \\
& +\mathbb{E}[(Z_1+1)^{-1}(\mu^2(m_+)^2(X_0-Z_1)(X_0-Z_1-1)|X_0 = k] + m^2 - \mu m^2 k \\
= \ & ((1-\mu)m + \mu\sigma^2 - \frac{\mu^2}{1-\mu}m^2 + \frac{\mu}{1-\mu}m^2)(1-(1-\mu)^k) + m^2 - \mu m^2 k(1-\mu)^{k-1} \\
= \ & ((1-\mu)m + \mu\sigma^2 + \mu m^2)(1-(1-\mu)^k) + m^2(1 - k\mu(1-\mu)^{k-1}) \\
= \ & ((1-\mu)m + \mu\sigma^2 + \mu m^2)\mathbb{P}(Z_1 \neq 0|X_0 = k) + m^2\mathbb{P}(Z_1 \neq 1|X_0 = k).
\end{aligned}$$

With this expression in our hand, we can identify a suitable martingale. We denote by $\mathcal{F}_n$ the σ-algebra generated by $\{Z_i; 1 \leq i \leq 2n\}$.

LEMMA 4.2.

$$
\begin{aligned}
M_n \quad := \quad & \sum_{j=1}^{n}\Big((Z_{2j-1}+1)\left(\frac{Z_{2j}}{Z_{2j-1}+1}-m\right)^2 \\
& -\left[((1-\mu)m+\mu\sigma^2+\mu m^2)\mathbb{1}_{\{Z_{2j-1}>0\}}+m^2\mathbb{1}_{\{Z_{2j-1}\neq 1\}}\right]\Big)
\end{aligned}
$$

is a martingale with respect to $\mathcal{F}_n$.

Proof: It is clear that M_n is measurable with respect to $\mathcal{F}_n$. Let $\xi_{n+1} := M_{n+1} - M_n$ be the increments and note that $\mathbb{E}(\xi_{n+1}|X_{2n}, \mathcal{F}_n) = \mathbb{E}(\xi_{n+1}|X_{2n}) = 0$, where the last equality follows from the previous computation. Hence,

$$
\mathbb{E}(M_{n+1}|\mathcal{F}_n) = M_n + \mathbb{E}(\mathbb{E}(\xi_{n+1}|X_{2n}, \mathcal{F}_n)|\mathcal{F}_n) = M_n.
$$

□

THEOREM 4.3. *We have, as* $n \to \infty$, *a.s. on* A,

$$
\frac{1}{n}\sum_{j=1}^{n}(Z_{2j-1}+1)\left(\frac{Z_{2j}}{Z_{2j-1}+1}-m\right)^2 \to (1-\mu)m+\mu\sigma^2+\mu m^2+m^2.
$$

Furthermore, writing

$$
\tilde{S}_n(m) := \sum_{i=1}^{n}(Z_i+1)\left(\frac{Z_{i+1}}{Z_i+1}-m\right)^2,
$$

we have a.s. on A,

$$
\frac{1}{n}\tilde{S}_n(m) \to (1-\mu)m+\mu\sigma^2+\mu m^2+m^2.
$$

Proof: From Lemma 4.2 we have that M_n is a martingale with respect to $\mathcal{F}_n$, with increments,

$$
\begin{aligned}
\xi_j \quad := \quad & (Z_{2j-1}+1)\left(\frac{Z_{2j}}{Z_{2j-1}+1}-m\right)^2 \\
& -\left[((1-\mu)m+\mu\sigma^2+\mu m^2)\mathbb{1}_{\{Z_{2j-1}>0\}}+m^2\mathbb{1}_{\{Z_{2j-1}\neq 1\}}\right].
\end{aligned}
$$

Now we apply Lemma 3.3 with the given ξ_j and $U_n = n$. On the set A, we have $U_n \to \infty$. To show that $\sum_{i=1}^{\infty} U_i^{-2}\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1}) < \infty$ on A, we claim that there exists a constant $C < \infty$ such that

$$
Var\left[(Z_1+1)\left(\frac{Z_2}{Z_1+1}-m\right)^2 |X_0 = k\right] < C,
$$

uniformly in k. The computation that justifies this claim is tedious, elementary and lengthy, and is omitted. (The computations can be obtained from the authors on request.) Now we have

$$
\begin{aligned}
\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1}) \quad = \quad & Var\left[(Z_{2i-1}+1)\left(\tfrac{Z_{2i}}{Z_{2i-1}+1}-m\right)^2 |\mathcal{F}_{i-1}\right] \\
= \quad & \mathbb{E}\Big(Var\left[(Z_{2i-1}+1)\left(\tfrac{Z_{2i}}{Z_{2i-1}+1}-m\right)^2 |X_{2(i-1)}\right] |\mathcal{F}_{i-1}\Big) < C,
\end{aligned}
$$

and we conclude that

$$\frac{M_n}{n} \to 0.$$

Now write $\frac{1}{n}\bar{M}_n = \frac{1}{n}\sum_{j=1}^{n} \bar{\xi}_j$, where

$$\bar{\xi}_j = (Z_{2j}+1)\left(\frac{Z_{2j+1}}{Z_{2j}+1} - m\right)^2 - \left[((1-\mu)m + \mu\sigma^2 + \mu m^2)1\!\!1_{\{Z_{2j}>0\}} + m^2 1\!\!1_{\{Z_{2j}\neq 1\}}\right]$$

are the martingale increments. Define $\bar{\mathcal{F}}_j$ as the σ-algebra generated by $\{Z_1, \dots, Z_{2j+1}\}$. Now with the same arguments as for the a.s. convergence of $\frac{1}{n}M_n$ we may prove that $\frac{1}{n}\bar{M}_n \to 0$ a.s. on A. Finally note that

$$\frac{1}{2n}\sum_{i=1}^{2n}(Z_i+1)\left(\frac{Z_{i+1}}{Z_i+1} - m\right)^2 - \left[(1-\mu)m + \mu\sigma^2 + \mu m^2\right]1\!\!1_{\{Z_{2j-1}>0\}} - m^2 1\!\!1_{\{Z_{2j-1}\neq 1\}}$$

is equal to

$$\frac{1}{2n}(M_n + \bar{M}_n),$$

and the second result of the theorem follows. □

Remark: Observe that $\sum_{i=1}^{n}(Z_i+1)(\frac{Z_{i+1}}{Z_i+1} - m)^2 - [(1-\mu)m + \mu\sigma^2 + \mu m^2 1\!\!1_{\{Z_i>0\}} + m^2 1\!\!1_{\{Z_i\neq 1\}}]$ is not a martingale itself, so we can not use Lemma 3.3 directly.

Because we do not know m, we cannot use $\tilde{S}_n(m)$ for estimation purposes, and we also need to analyse the behavior of $\tilde{S}_n(\tilde{m}_n)$. Some algebra yields

$$\frac{1}{n}\left(\tilde{S}_n(m) - \tilde{S}_n(\tilde{m}_n)\right) = \frac{1}{n}(m - \tilde{m}_n)^2 \sum_{i=1}^{n} Z_i + (m^2 - \tilde{m}_n^2). \tag{5}$$

From Theorem 3.5 we know that the square root of $(m - \tilde{m}_n)^2 \sum_{i=1}^{n} Z_i$ is the sum of three random variables, each converging in distribution to a normally distributed random variable with finite variance. So the square root of the first term on the right hand side converges in distribution to 0. Because 0 is a constant, the convergence also is in probability. If $A_n \to 0$ in probability, then $A_n^2 \to 0$ in probability, so $\frac{1}{n}(m - \tilde{m}_n)^2 \sum_{i=1}^{n} Z_i$ converges in probability to 0 on A. Together with Theorem 4.3 this proves Theorem 4.1.

4.2. The general case

Until now we considered the situation where the observed individuals have no further offspring. We now allow observed individuals to have some X-offspring in the generation after the observation. So in terms of epidemics, in this section we allow detected individuals to infect other individuals during the interval of detection. We assume that the fourth moment of the offspring distributions are finite. Our main result in this more general case reads as follows.

THEOREM 4.4. *As $n \to \infty$, we have*

$$\frac{1}{n}\sum_{i=1}^{n}(Z_i+1)\left(\frac{Z_{i+1}}{Z_i+1}-\frac{\sum_{j=2}^{n+1}Z_j}{\sum_{j=1}^{n}Z_j}\right)^2 \to (1-\mu)m+\mu\sigma^2+\mu m^2+m^2-2\mu m m_-,$$

in probability on A.

This gives us a consistent estimator (in probability) of $(1-\mu)m+\mu\sigma^2+\mu m^2+m^2-2\mu m m_-$ on the explosion set A.

To prove this result, one can compute that (again we omit the lengthy details)

$$\begin{aligned}
&\mathbb{E}\Big[(Z_1+1)\Big(\frac{Z_2}{Z_1+1}-m\Big)^2 - \\
&[(1-\mu)m+\mu\sigma^2+\mu m^2]1\!\!1_{\{Z_1>0\}} - m^2 1\!\!1_{\{Z_1\neq 1\}} + 2\mu m(m_-)1\!\!1_{\{Z_1>1\}}|X_0=k\Big] \\
+\quad &\mathbb{E}\Big[\frac{\mu[\mu\sigma_-^2+(1-\mu)(m_-)]}{Z_1+1}1\!\!1_{\{Z_1>0\}} - \frac{\mu^2(m_-)^2}{Z_1+1}(1\!\!1_{\{Z_1>1\}}-1\!\!1_{\{Z_1=1\}})|X_0=k\Big]=0.
\end{aligned}$$

This leads to the following lemma, which can be proved as Lemma 4.2:

LEMMA 4.5. *Let $\mathcal{F}_n$ be the σ-algebra generated by $\{Z_i; 1\le i\le 2n\}$. Then*

$$\begin{aligned}
M_n \quad := \quad & \sum_{j=1}^{n}\Big[(Z_{2j-1}+1)\left(\frac{Z_{2j}}{Z_{2j-1}+1}-m\right)^2 - \\
&[(1-\mu)m+\mu\sigma^2+\mu m^2]1\!\!1_{\{Z_{2j-1}>0\}} - m^2 1\!\!1_{\{Z_{2j-1}\neq 1\}} \\
&+2\mu m(m_-)1\!\!1_{\{Z_{2j-1}>1\}} + \frac{\mu[\mu\sigma_-^2+(1-\mu)(m_-)]}{Z_{2j-1}+1}1\!\!1_{\{Z_{2j-1}>0\}} \\
&-\frac{\mu^2(m_-)^2}{Z_{2j-1}+1}(1\!\!1_{\{Z_{2j-1}>1\}}-1\!\!1_{\{Z_{2j-1}=1\}})\Big]
\end{aligned}$$

is a martingale with respect to $\mathcal{F}_n$.

Using this lemma we now prove

THEOREM 4.6. *As $n\to\infty$, we have a.s. on A,*

$$\frac{1}{n}\sum_{j=1}^{n}(Z_{2j-1}+1)\Big(\frac{Z_{2j}}{Z_{2j-1}+1}-m\Big)^2 \to (1-\mu)m+\mu\sigma^2+\mu m^2+m^2-2\mu m m_-.$$

Furthermore, writing

$$\tilde{S}_n(m) := \sum_{i=1}^{n}(Z_i+1)\Big(\frac{Z_{i+1}}{Z_i+1}-m\Big)^2,$$

we have a.s. on A,

$$\frac{1}{n}\tilde{S}_n(m) \to (1-\mu)m+\mu\sigma^2+\mu m^2+m^2-2\mu m m_-.$$

Proof: Again we use the same proof as for the special case of the previous section. The only extra thing to prove is that there exists a $C < \infty$ such that

$$Var\left[(Z_1+1)\left(\frac{Z_2}{Z_1+1}-m\right)^2|X_0=k\right] < C,$$

for all k. This is equivalent to proving that

$$\mathbb{E}\left[(Z_1+1)^2\Big(\frac{Z_2}{Z_1+1}-m\Big)^4|X_0=k\right] < C_1 \tag{6}$$

for some $C_1 < \infty$ and all k. Z_2 consists of, say, Z_2^+ individuals that are stemming from individuals not detected in generation 1 and Z_2^- individuals stemming from individuals detected in generation 1. Now we can write (6) as

$$\mathbb{E}\left[(Z_1+1)^2\Big(\frac{Z_2^+ + Z_2^-}{Z_1+1}-(1-\mu)m_+ - \mu m_-\Big)^4|X_0=k\right] < C_1.$$

Minkowski's inequality yields

$$\begin{aligned}&\mathbb{E}\left[(Z_1+1)^2\left(\tfrac{Z_2^+ + Z_2^-}{Z_1+1}-(1-\mu)m_+ - \mu m_-\right)^4|X_0=k\right]\\ &\leq 16\mathbb{E}\left[(Z_1+1)^2\left(\tfrac{Z_2^+}{Z_1+1}-(1-\mu)m_+\right)^4|X_0=k\right]\\ &+16\mathbb{E}\left[(Z_1+1)^2\left(\tfrac{Z_2^-}{Z_1+1}-\mu m_-\right)^4|X_0=k\right].\end{aligned}$$

Note that the first expression at the righthand side is exactly the expression we used in Section 4.1, so we only have to analyse the second expression. We write this as

$$16\mathbb{E}\Big[\frac{(Z_2^- - \mu m_- Z_1 - \mu m_-)^4}{(Z_1+1)^2}|X_0=k\Big].$$

Now $(Z_2^- - \mu m_- Z_1)$ is the sum of Z_1 i.i.d. random variables. Because the fourth moment of the binomial distribution is finite and because we have assumed that the offspring distribution of observed individuals has a finite fourth moment, we have by a simple computation that

$$\mathbb{E}[(Z_2^- - \mu m_- Z_1)^4|Z_1] < C_2 Z_1^2,$$

for some constant $C_2 < \infty$. In a similar way we have

$$\mathbb{E}[(Z_2^- - \mu m_- Z_1)^3|Z_1] < C_3 Z_1,$$

and

$$\mathbb{E}\left[(Z_2^- - \mu m_- Z_1)^2|Z_1\right] < C_4 Z_1,$$

for some $C_3 < \infty$ and $C_4 < \infty$. Hence

$$\mathbb{E}\left[\frac{(Z_2^- - \mu m_- Z_1 - \mu m_-)^4}{(Z_1+1)^2}|X_0=k\right] < \frac{C_5(Z_1^2+Z_1+1)}{(Z_1+1)^2} \leq C_5$$

for some $C_5 < \infty$, which proves our theorem. □

The argument can now be finished exactly as in the previous section.

It seems rather difficult to establish the rate of convergence for the estimator of the second function of parameters, but the following theorem gives a bound for this rate.

Theorem 4.7. *Define* $\gamma := (1-\mu)m + \mu\sigma^2 + \mu m^2 + m^2 - 2\mu m m_-$. *Then for all* $\delta > 0$ *we have*

$$n^{\frac{1}{2}-\delta}\left(\frac{1}{n}\tilde{S}_n(\tilde{m}_n) - \gamma\right) \to 0,$$

in probability on A.

Proof: From Corollary 3.1 in Hall it et al. (1980) it follows that if $\{S_{ni}, \mathcal{F}_{n,i}, 1 \le i \le n\}$ is a square integrable martingale array with differences $X_{ni} := S_{ni} - S_{ni-1}$ (where S_{n0} is defined as 0) and such that

$$\sum_{i=1}^{n} \mathbb{E}(X_{ni}^2 | \mathcal{F}_{n,i-1}) \to 0 \qquad \text{in probability,}$$

and $\mathcal{F}_{n,i} \subseteq \mathcal{F}_{n+1,i}$ for $1 \le i \le n$, then $S_{nn} = \sum_{i=1}^{n} X_{ni} \to 0$ in probability.

We use this with $S_{nn} = n^{-(\frac{1}{2}+\delta)} M_n$, where M_i is as in Lemma 4.5, and $X_{ni} = n^{-(\frac{1}{2}+\delta)} T_i$, where T_i are the summands of M_n. We define $\mathcal{F}_{n,i}, 1 \le i \le n$, as the σ-algebra generated by $\{Z_i; 1 \le i \le 2n\}$, so $\mathcal{F}_{n,i} \subseteq \mathcal{F}_{n+1,i}$ holds for $1 \le i \le n$. We have already shown that $\mathbb{E}(T_i^2) < C$ for some C, uniformly in i, hence

$$\sum_{i=1}^{n} \mathbb{E}(X_{ni}^2 | \mathcal{F}_{n,i-1}) \le n^{-(1+2\delta)} nC \to 0 \qquad \text{in probability,}$$

so $n^{-(\frac{1}{2}+\delta)} M_n \to 0$ in probability. We can prove in the same way that $n^{-(\frac{1}{2}+\delta)} \bar{M}_n \to 0$, where $\bar{M}_n$ is as in the proof of Theorem 4.3. Now by the definitions of M_n and $\bar{M}_n$, we may see that on the explosion set A,

$$n^{\frac{1}{2}-\delta}(\frac{1}{n}\tilde{S}_n(m) - \gamma) = n^{-(\frac{1}{2}+\delta)} M_n + n^{-(\frac{1}{2}+\delta)} \bar{M}_n + n^{-(\frac{1}{2}+\delta)} \sum_{i=1}^{n} f(Z_i),$$

where $f(x) = O(\frac{1}{x})$. Since $\frac{Z_{i+1}}{X_i} \to \mu$ a.s. on A and $\frac{X_i}{m^i}$ converges almost surely to a finite random variable, we know that $\frac{Z_i}{m^i}$ almost surely converges to an almost surely finite and positive random variable, $\bar{W}$ say. Now we use the Toeplitz Lemma (Lemma 1.2 in Guttorp it et al. (1991)) to see that

$$\sum_{i=1}^{\infty} f(Z_i) < \left(\sum_{i=1}^{\infty} m^{-i}\right) C \frac{1}{\bar{W}} < \infty,$$

for some C. So $n^{\frac{1}{2}-\delta}(\frac{1}{n}\tilde{S}_n(m) - \gamma) \to 0$ in probability on the explosion set A.

By using the rate of convergence of $\tilde{m}_n$, equation (5) and the arguments following that equation we see that for all $\delta_1 > 0$, $\frac{1}{n^{\delta_1}}\tilde{S}_n(m) - \frac{1}{n^{\delta_1}}\tilde{S}_n(\tilde{m}_n) \to 0$ in probability on A, and now with $\delta_1 = \frac{1}{2} + \delta$ the theorem follows. □

4.3. Interpretation of γ

The expression

$$\gamma = (1-\mu)m + \mu\sigma^2 + \mu m^2 + m^2 - 2\mu m m_-$$

which appeared as the limit in the previous subsection, turns out to have a somewhat surprising interpretation: it appears as a second moment if we treat our Z-observations as a Galton-Watson process itself. To explain what we mean by this, we compute $\mathbb{P}(X_0 = l|Z_1 = k)$, when the a-priori distribution of X_0 is uniform on the integers between k and N, where $N >> k$. After this we let N tend to infinity.

$$\mathbb{P}(X_0 = l|Z_1 = k) = \frac{\mathbb{P}(Z_1 = k|X_0 = l)\mathbb{P}(X_0 = l)}{\sum_{i=k}^{N} \mathbb{P}(Z_1 = k|X_0 = i)\mathbb{P}(X_0 = i)} = \frac{\mathbb{P}(Z_1 = k|X_0 = l)}{\sum_{i=k}^{N} \mathbb{P}(Z_1 = k|X_0 = i)}.$$

First we compute the denominator for $N \to \infty$

$$\begin{aligned}\lim_{N\to\infty} \sum_{i=k}^{N} \mathbb{P}(Z_1 = k|X_0 = i) &= \sum_{i=k}^{\infty} \binom{i}{k} \mu^k (1-\mu)^{i-k} \\ &= \frac{1}{\mu} \sum_{j=0}^{\infty} \binom{k+j}{j} \mu^{k+1}(1-\mu)^j.\end{aligned}$$

The summands are exactly the probabilities of a negative binomial distribution, with parameters $k+1$ and μ, so $\sum_{j=0}^{\infty} \binom{k+j}{j} \mu^{k+1}(1-\mu)^j = 1$. Therefore we know that the denominator converges to μ^{-1}.

With some abuse of notation, we write a superscript $*$ when we discuss probabilities and accompanying expectations after taking the limit for $N \to \infty$. This leads to

$$\mathbb{P}^*(X_0 = l|Z_1 = k) = \binom{l}{k} \mu^{k+1}(1-\mu)^{l-k}.$$

Now it is easy to compute that

$$\begin{aligned}\mathbb{E}^*(X_0|Z_1 = k) &= \frac{k+1}{\mu} - 1, \\ Var^*(X_0|Z_1 = k) &= \frac{(1-\mu)(k+1)}{\mu^2}.\end{aligned}$$

Now we may compute $\mathbb{E}^*(Z_2|Z_1 = k)$ and $Var^*(Z_2|Z_1 = k)$ by

$$\begin{aligned}\mathbb{E}^*(Z_2|Z_1 = k) &= \mu\mathbb{E}^*(X_1|Z_1 = k) \\ &= \mu m_+ \mathbb{E}^*(X_0 - Z_1|Z_1 = k) + \mu m_- k \\ &= \mu m_+ \frac{k+1}{\mu} - \mu m_+ - \mu(m_+ - m_-)k \\ &= (1-\mu)m_+(k+1) + \mu m_- k \\ &= (1-\mu)m_+ + mk,\end{aligned}$$

$$\begin{aligned}Var^*(Z_2|Z_1 = k) &= \mathbb{E}[\Big(Z_2 - (1-\mu)m_+ - mk\Big)^2|Z_1 = k] \\ &= \mathbb{E}[\mu(1-\mu)X_1 + \Big(\mu X_1 - (1-\mu)m_+ - mk\Big)^2|Z_1 = k] \\ &= \mu(1-\mu)m_+\mathbb{E}^*(X_0 - Z_1|Z_1 = k) + \mu(1-\mu)m_- k\end{aligned}$$

$$
\begin{aligned}
&+\mu^2\sigma_+^2\mathbb{E}^*(X_0-Z_1|Z_1=k)+\mu^2\sigma_-^2 k\\
&+\mathbb{E}^*([\mu m_+(X_0-Z_1)+\mu m_- Z_1-(1-\mu)m_+-mk]^2|Z_1=k)\\
=\;&(1-\mu)m_+(k+1-\mu)-\mu(1-\mu)(m_+-m_-)k\\
&+\mu\sigma_+^2(k+1-\mu)+\mu^2(\sigma_-^2-\sigma_+^2)k\\
&+\mu^2(m_+)^2\mathbb{E}^*([X_0-\frac{k+1}{\mu}-1]^2|Z_1=k)\\
=\;&(1-\mu)m(k+1)-(1-\mu)\mu m_-+\mu\sigma^2 k\\
&+\mu^2(1-\mu)(m_+-m_-)^2k+\mu(\sigma^2-\mu\sigma_-^2-\mu(1-\mu)(m_+-m_-)^2)\\
&+\frac{\mu^2}{(1-\mu)^2}(m-\mu m_-)^2\frac{(1-\mu)(k+1)}{\mu^2}\\
=\;&[(1-\mu)m+\mu\sigma^2+(1+\mu)m^2-2\mu m m_-]k\\
&+(1-\mu)m-(1-\mu)\mu m_-+\mu(\sigma^2-\mu\sigma_-^2-\frac{\mu}{1-\mu}(m-m_-)^2)\\
&+\frac{(m-\mu m_-)^2}{(1-\mu)}\\
=\;&[(1-\mu)m+\mu\sigma^2+(1+\mu)m^2-2\mu m m_-](k+1)\\
&-(1-\mu)\mu m_- -\mu^2\sigma_-^2.
\end{aligned}
$$

We see that $\mathbb{E}^*(Z_2|Z_1=k)=mk+O(1)$ for $k\to\infty$ and $Var^*(Z_2|Z_1=k)=\gamma k+O(1)$ for $k\to\infty$. In this sense, we again estimate a first and second moment, just as in the classical case where the full generation sizes are observed.

5. Estimating a third function of parameters

From Theorem 1.3 of Guttorp it et al. (1991) we know that we cannot estimate more than two functions of the parameters consistently if only the generation sizes are observed. However, in our context we can under certain conditions estimate a third function of parameters, as we will now show.

We have already shown that for $k\to\infty$

$$
\begin{aligned}
\mathbb{E}\Big(\frac{Z_2}{Z_1+1}|X_0=k\Big) &\;\to\; m,\\
\mathbb{E}\Big((Z_1+1)\left(\frac{Z_2}{Z_1+1}-m\right)^2|X_0=k\Big) &\;\to\; \gamma.
\end{aligned}
$$

We next compute $\mathbb{E}[(Z_1+1)(\frac{Z_3}{Z_1+1}-m^2)^2|X_0=k]$:

$$
\begin{aligned}
&\mathbb{E}[(Z_1+1)(\tfrac{Z_3}{Z_1+1}-m^2)^2|X_0=k]\\
=\;&\mathbb{E}[(Z_1+1)\Big[\Big(\tfrac{Z_3-\mu X_2}{Z_1+1}\Big)^2+\mu^2\Big(\tfrac{X_2-mX_1}{Z_1+1}\Big)^2+m^2\Big(\tfrac{\mu X_1}{Z_1+1}-m\Big)^2\Big]|X_0=k]\\
=\;&\mu(1-\mu)m\mathbb{E}\Big(\tfrac{X_1}{Z_1+1}|X_0=k\Big)+\mu^2\sigma^2\mathbb{E}\Big(\tfrac{X_1}{Z_1+1}|X_0=k\Big)\\
&+m^2\mathbb{E}\Big((Z_1+1)(\tfrac{Z_2}{Z_1+1}-m)^2|X_0=k\Big)-m^2\mu(1-\mu)\mathbb{E}\Big(\tfrac{X_1}{Z_1+1}|X_0=k\Big)\\
=\;&((1-\mu)m+\mu\sigma^2-m^2(1-\mu))\mathbb{E}\Big(\tfrac{Z_2}{Z_1+1}|X_0=k\Big)\\
&+m^2\mathbb{E}\Big((Z_1+1)(\tfrac{Z_2}{Z_1+1}-m)^2|X_0=k\Big).
\end{aligned}
$$

Now note that on A,

$$\begin{aligned}
& \mathbb{E}[(Z_1+1)(\tfrac{Z_3}{Z_1+1}-m^2)^2|X_0=k] \\
= \; & ((1-\mu)m+\mu\sigma^2-m^2(1-\mu))\mathbb{E}\Big(\tfrac{Z_2}{Z_1+1}|X_0=k\Big) \\
& +m^2\mathbb{E}\Big((Z_1+1)(\tfrac{Z_2}{Z_1+1}-m)^2|X_0=k\Big) \\
\to \; & (\gamma-2m^2+2\mu m m_-)m+m^2\gamma \\
= \; & (m^2+m)\gamma-2(1-\mu)m^2m_+,
\end{aligned}$$

where the convergence is almost surely. We define

$$\gamma_* := (m^2+m)\gamma-2(1-\mu)m^2m_+,$$

and

$$S_n^*(m) := \sum_{i=1}^{n}(Z_i+1)\left(\frac{Z_{i+2}}{Z_i+1}-m^2\right)^2.$$

We can use the same martingale argument as used for the proof of Theorem 4.4 to prove the following result.

THEOREM 5.1. *The quantity*

$$\frac{1}{n}\sum_{i=1}^{n}\Big[\Big(\Big[\frac{\sum_{j=2}^{n+1}Z_j}{\sum_{j=1}^{n}Z_j}\Big]^2+\frac{\sum_{j=2}^{n+1}Z_j}{\sum_{j=1}^{n}Z_j}\Big)(Z_i+1)\left(\frac{Z_{i+1}}{Z_i+1}-\frac{\sum_{j=2}^{n+1}Z_j}{\sum_{j=1}^{n}Z_j}\right)^2\Big]-\frac{1}{n}S_n^*(m)$$

converges (on A) in probability to $2(1-\mu)m^2m_+$. Hence,

$$\frac{1}{n}S_n^*(m)\to\gamma_*,$$

in probability on A.

In order to obtain an observable quantity (m is not), we need to bound $\frac{1}{n}(S_n^*(\tilde{m}_n)-S_n^*(m))$:

$$
\begin{aligned}
& \tilde{S}_n^*(m) - \tilde{S}_n^*(\tilde{m}_n) \\
= \quad & \textstyle\sum_{i=1}^n (Z_i + 1) \left(\frac{Z_{i+2}}{Z_i+1} - m^2 \right)^2 - \sum_{i=1}^n (Z_i + 1) \left(\frac{Z_{i+2}}{Z_i+1} - (\tilde{m}_n)^2 \right)^2 \\
= \quad & \textstyle\sum_{i=1}^n \Big[(Z_i + 1)(m^4 - (\tilde{m}_n)^4) - 2Z_{i+2}[m^2 - (\tilde{m}_n)^2] \Big] \\
= \quad & \textstyle\sum_{i=1}^n \Big[Z_i(m^4 - (\tilde{m}_n)^4) - 2\tilde{m}_{n+1}\tilde{m}_n Z_i[m^2 - (\tilde{m}_n)^2] \Big] \\
& + n(m^4 - (\tilde{m}_n)^4) + 2(Z_2 - \tilde{m}_{n+1} Z_1)[m^2 - (\tilde{m}_n)^2] \\
= \quad & 2(Z_2 - \tilde{m}_{n+1} Z_1)[m^2 - (\tilde{m}_n)^2] \\
& + n(m^4 - (\tilde{m}_n)^4) + (m^2 - (\tilde{m}_n)^2)^2 \left(\textstyle\sum_{i=1}^n Z_i \right) \\
& - 2(\tilde{m}_{n+1} - \tilde{m}_n)\tilde{m}_n [m^2 - (\tilde{m}_n)^2] \left(\textstyle\sum_{i=1}^n Z_i \right) \\
= \quad & n(m^4 - (\tilde{m}_n)^4) + 2(Z_2 - \tilde{m}_{n+1} Z_1)[m^2 - (\tilde{m}_n)^2] \\
& + (m + \tilde{m}_n)^2 (m - \tilde{m}_n)^2 \left(\textstyle\sum_{i=1}^n Z_i \right) \\
& + 2(m + \tilde{m}_n)\tilde{m}_n \left(\textstyle\sum_{i=1}^n Z_i \right) \Big[(m - \tilde{m}_{n+1})(m - \tilde{m}_n) - (m - \tilde{m}_n)^2 \Big] \\
= \quad & n(m^4 - (\tilde{m}_n)^4) + 2(Z_2 - \tilde{m}_{n+1} Z_1)[m^2 - (\tilde{m}_n)^2] \\
& + (m + \tilde{m}_n)(m - \tilde{m}_n)^3 \left(\textstyle\sum_{i=1}^n Z_i \right) \\
& + 2(m + \tilde{m}_n)\tilde{m}_n (m - \tilde{m}_{n+1})(m - \tilde{m}_n) \left(\textstyle\sum_{i=1}^n Z_i \right) \\
\leq \quad & n(m^4 - (\tilde{m}_n)^4) + 2(Z_2 - \tilde{m}_{n+1} Z_1)[m^2 - (\tilde{m}_n)^2] \\
& + (m + \tilde{m}_n)(m - \tilde{m}_n)^3 \left(\textstyle\sum_{i=1}^n Z_i \right) \\
& + (m + \tilde{m}_n)\tilde{m}_n [(m - \tilde{m}_{n+1})^2 + (m - \tilde{m}_n)^2] \left(\textstyle\sum_{i=1}^n Z_i \right) \\
= \quad & n(m^4 - (\tilde{m}_n)^4) + 2(Z_2 - \tilde{m}_{n+1} Z_1)[m^2 - (\tilde{m}_n)^2] \\
& + m(m + \tilde{m}_n)(m - \tilde{m}_n)^2 \left(\textstyle\sum_{i=1}^n Z_i \right) \\
& + (m + \tilde{m}_n)(m - \tilde{m}_{n+1})^2 \left(\left(\textstyle\sum_{i=1}^{n+1} Z_i \right) - \tilde{m}_n Z_1 \right).
\end{aligned}
$$

Now $\frac{Z_1}{n}$ and $\frac{Z_2}{n}$ converge almost surely to 0. We can use the same arguments as used in Section 4 for the proof of convergence of $\frac{1}{n}(S_n(\tilde{m}_n) - S_n(m))$ to see that $\frac{1}{n}(S_n^*(\tilde{m}_n) - S_n^*(m))$ converges to zero in probability.

Note that $2(1-\mu)m^2 m_+$ is not necessary in the plain spanned by m and γ, so we are able to estimate a third parameter.

Remark: If $m_- = 0$ we are unable to estimate the third parameter this way. However, if we look at the case where we observe a binomial distributed number of individuals from each generation, but observations do not influence the offspring distribution (that is, $m_+ = m_- = m$ and $\sigma_+^2 = \sigma_-^2 = \sigma^2$), we have an estimator for $(1-\mu)m^3$ and because we have an estimator for m that converges a.s., we are in theory able to estimate μ consistently.

As far as rates are concerned, we can prove in exactly the same way as Theorem 4.7 that

THEOREM 5.2. *For all* $\delta > 0$

$$
n^{\frac{1}{2}-\delta} \left(\frac{1}{n} \tilde{S}_n^*(\tilde{m}_n) - \gamma_* \right) \to 0
$$

in probability on A.

Summerising the results of the previous sections, we have given three consistent estimators for three functions of the parameters:

$$\begin{array}{rcll} \tilde{m}_n & \to & m, & \text{a.s. on } A, \\ \frac{1}{n}\tilde{S}_n(\tilde{m}_n) & \to & \gamma = (1-\mu)m + \mu\sigma^2 + \mu m^2 + m^2 - 2\mu m m_-, & \text{in probability on } A, \\ \frac{1}{n}\tilde{S}^*_n(\tilde{m}_n) & \to & \gamma_* = (m^2+m)\gamma - 2m^3 + 2\mu m^2 m_-, & \text{in probability on } A, \end{array}$$

where

$$\begin{aligned} \tilde{m}_n &= \frac{\sum_{i=2}^{n+1} Z_i}{\sum_{i=1}^{n} Z_i}, \\ \tilde{S}_n(m) &= \sum_{i=1}^{n}(Z_i+1)\left(\frac{Z_{i+1}}{Z_i+1} - m\right)^2, \\ \tilde{S}^*_n(m) &= \sum_{i=1}^{n}(Z_i+1)\left(\frac{Z_{i+2}}{Z_i+1} - m^2\right)^2. \end{aligned}$$

In principle, this gives three equations with four unknowns, namely m, μ, σ^2 and m_-. If we have further information, or make further assumptions about the relation between m and m_-, then we can estimate all parameters consistently, in theory at least. This is illustrated with a concrete example in the next section.

6. Analyzing a concrete offspring distribution

6.1. Definition of the model

In this section, we consider a concrete example from epidemic theory. We analyze the discrete approximation of the standard SIR epidemic (see e.g. Diekmann it et al. (2000)), where we take the number of susceptible individuals to be infinite.

We assume that if an infective individual is not detected in a certain interval, the number of new infections by this infective individual is Poisson distributed with parameter λ. Since an individual that is not detected remains infective itself, this leads to

$$\begin{aligned} m_+ &= \lambda + 1, \\ \sigma_+^2 &= \lambda. \end{aligned}$$

We need to make a choice for what happens during the interval that an individual is detected. In order to keep the model rather general we assume that the detected individual was infective during a (known) fraction ϕ of the detection interval. It then follows that

$$\begin{aligned} m_- &= \phi\lambda, \\ \sigma_-^2 &= \phi\lambda, \end{aligned}$$

and hence

$$\begin{aligned} m &= (1-\mu)(\lambda+1) + \phi\lambda\mu, \\ \sigma^2 &= (1-\mu)\lambda + \phi\lambda\mu + \mu(1-\mu)((1-\phi)\lambda+1)^2. \end{aligned}$$

Remark: If we assume that the detection time is uniformly distributed over the interval of detection, m_- will be $\frac{\lambda}{2}$. The variance of the offspring in the interval of detection will be slightly larger than $\frac{\lambda}{2}$ because the randomness of the detection time will cause some extra variance. In fact the variance would be $\frac{\lambda}{2}+\frac{\lambda^2}{12}$.

We are particulary interested in estimating m, which describes the mean growth of the number of infectious individuals, and μ, because that parameter is needed to estimate the number of infectious individuals at a certain time, which is very important in order to make decisions about measures to stop the epidemic.

In the context of the present example, the results summarised at the end of the previous section give

$$\begin{array}{rcll} \tilde{m}_n & \to & (1-\mu)(\lambda+1)+\phi\lambda\mu & \text{a.s. on } A, \\ \frac{1}{n}\tilde{S}_n(\tilde{m}_n) & \to & (1-\mu)\lambda+(1-\mu)^2+(1-\mu)(\lambda+1)^2+\phi\lambda\mu & \text{in probability on } A, \\ \frac{1}{n}\tilde{S}^*_n(\tilde{m}_n) & \to & (m^2+m)\gamma+2m^2(1-\mu)(\lambda+1) & \text{in probability on } A. \end{array} \quad (7)$$

For ease of notation, we did not expand the last quantity at the right-hand side. In (7) we have, after substituting the estimates for m, γ and γ_*, three equations with two unknowns, λ and μ. At this point, it seems that the third equation doesn't help us very much. There are several ways to proceed now.

First of all, we can ignore the third equation and solve the other two for λ and μ. However, it turns out that $\frac{1}{n}\tilde{S}_n(\tilde{m}_n)$ converges very slow, and we need a huge number of generations to obtain reliable estimates (see the next subsection).

But there is a way to use the information contained in the third equation in a meaningful way. To this end, we reparametrise the epidemic process by using m and μ instead of λ and μ. The parameter m is estimated by $\tilde{m}_n$ while we may use the combination $(1+\tilde{m}_n^{-1})\frac{1}{n}\tilde{S}_n(\tilde{m}_n)-\tilde{m}_n^{-2}\frac{1}{n}\tilde{S}^*_n(\tilde{m}_n)$ to estimate $2(1-\mu)(\lambda+1)$. Since $m=(1-\mu)(\lambda+1)+\phi\lambda\mu$ we may write:

$$\lambda+1=\frac{m+\phi\mu}{(1-\mu)+\phi\mu},$$

so again we have a system of two equations with two unknowns. From simulation results it turns out that we can give reasonable estimates for μ much faster than in the case where we use only the estimators for m and γ (see the next subsection), but for practical purposes our new estimators for μ still converges too slow.

6.2. Simulation results

We used our analysis on simulated branching processes. For the simulations we use `Mathematica`. For small generation sizes ($X_n < 1000$) we used a $Binomial(X_{n-1},\mu)$ distribution for Z_n and if $Y(x)\simeq Poisson(x)$,

$$X_n=X_{n-1}-Z_n+Y(\lambda(X_{n-1}-(1-\phi)Z_n)).$$

For large X_{n-1} it is very time consuming to simulate with binomial and Poisson distributions. To deal with this we use the (integer parts of the) normal approximations for the distributions of Z_n and X_n. So Z_n has approximately a normal distribution with mean μX_{n-1} and variance $(1-\mu)\mu X_{n-1}$, while $X_n-X_{n-1}+Z_n$ has approximately a normal distribution with mean and variance $\lambda(X_{n-1}-(1-\phi)Z_n)$.

Table 1. Point estimates and 95% quantile intervals for the case where $\phi = 0.5$

m	μ	$\tilde{m}_{25}$	γ	$\frac{1}{25}\tilde{S}_{25}(\tilde{m}_{25})$	γ_*	$\frac{1}{25}\tilde{S}^*_{25}(\tilde{m}_{26})$
1.1	0.1	1.12 (1.00 – 1.50)	2.33	1.46 (0.914 – 2.60)	2.74	1.89 (0.923 – 5.34)
	0.4	1.11 (1.00 – 1.21)	2.44	2.18 (1.13 – 4.15)	3.29	3.00 (1.41 – 5.95)
	0.5	1.11 (1.01 – 1.21)	2.47	2.15 (1.03 – 4.29)	3.53	3.15 (1.43 – 5.55)
	0.6	1.11 (1.02 – 1.20)	2.46	2.25 (1.04 – 4.44)	3.75	3.50 (1.49 – 6.71)
	0.9	1.10 (1.01 – 1.18)	1.80	1.69 (0.817 – 3.31)	3.49	3.10 (1.31 – 6.18)
1.5	0.1	1.50 (1.47 – 1.52)	3.81	3.23 (1.78 – 5.43)	7.66	6.54 (3.90 – 12.0)
	0.4	1.50 (1.48 – 1.51)	3.97	3.55 (1.64 – 6.49)	9.15	8.21 (3.83 – 15.3)
	0.5	1.50 (1.49 – 1.51)	3.97	3.55 (1.88 – 5.66)	9.65	8.42 (4.05 – 15.7)
	0.6	1.50 (1.48 – 1.51)	3.90	3.58 (1.92 – 6.09)	10.0	9.02 (4.03 – 15.1)
	0.9	1.50 (1.49 – 1.51)	2.67	2.55 (1.28 – 4.10)	8.41	7.68 (3.90 – 15.0)
2.0	0.1	2.00 (2.00 – 2.00)	6.10	5.46 (2.89 – 9.69)	21.0	19.1 (10.5 – 30.8)
	0.4	2.00 (2.00 – 2.00)	6.30	6.03 (3.13 – 10.1)	24.6	22.8 (11.5 – 35.4)
	0.5	2.00 (2.00 – 2.00)	6.25	6.04 (2.97 – 11.4)	25.5	24.5 (12.5 – 42.8)
	0.6	2.00 (2.00 – 2.00)	6.08	5.70 (3.06 – 10.3)	26.0	23.5 (13.2 – 41.0)
	0.9	2.00 (2.00 – 2.00)	3.89	3.99 (2.08 – 6.53)	19.8	19.6 (9.78 – 34.0)

Table 2. Point estimates and 95%-quantile intervals for the case where $\phi = 0$.

m	μ	$\tilde{m}_{25}$	γ	$\frac{1}{25}\tilde{S}_{25}(\tilde{m}_{25})$	γ_*	$\frac{1}{25}\tilde{S}^*_{25}(\tilde{m}_{26})$
1.1	0.5	1.10 (1.01 – 1.18)	3.27	3.06 (1.41 – 6.55)	4.89	4.77 (2.02 – 9.59)
1.5	0.5	1.50 (1.49 – 1.51)	5.75	5.33 (2.44 – 9.96)	14.8	13.6 (6.57 – 26.1)
2.0	0.5	2.00 (2.00 – 2.00)	9.75	9.43 (5.20 – 15.6)	42.5	40.2 (21.8 – 63.3)

(A) First we simulated with $\phi = 0.5$, $X_0 = 1$ and we used 25 generations for estimation. We simulated - for several μ and m values - 200 exploding branching processes. The results are given in Table 6.2. In this table, we give point estimates and 95% quantile intervals for m, γ and γ_*. It was impossible to obtain estimates for μ with only 25 generations.

(B) We also did simulations with $\phi = 0$ and $\phi = 1$. Here we took $\mu = \frac{1}{2}$ and $X_0 = 1$. For the results see Table 6.2 and Table 6.2.

(C) The estimates of m and γ at different generations are given in Table 6.2. Here we have used estimated λ and μ from the 1997 Dutch classical swine fever outbreak Meester it et al. (2002), $\lambda = 0.6$, $\mu = 0.4$, while $\phi = 0.5$. The real values of the parameters we estimate in our simulation are $m = 1.08$, $\gamma = 2.376$ and $\gamma_* = 3.098$. In the last column we have used the estimated values of m, γ and γ_* to estimate μ. We used that for the real values of m, γ and γ_* it holds that

$$(m^{-1} + m)\gamma - m^{-2}\gamma_* = 2(1 - \mu)m_+$$

and $m_+ = \frac{\phi\mu + m}{\phi\mu + (1-\mu)}$. Only after 140 time steps we got real values for the estimator of μ.

Table 3. Point estimates and 95%-quantile intervals for the case where $\phi = 1$.

m	μ	$\tilde{m}_{25}$	γ	$\frac{1}{25}\tilde{S}_{25}(\tilde{m}_{25})$	γ_*	$\frac{1}{25}\tilde{S}^*_{25}(\tilde{m}_{26})$
1.1	0.5	1.10 (1.01 – 1.19)	2.13	1.83 (0.940 – 3.03)	2.98	2.43 (1.17 – 4.74)
1.5	0.5	1.50 (1.49 – 1.51)	3.25	3.07 (1.67 – 5.31)	7.69	7.11 (3.58 – 12.5)
2.0	0.5	2.00 (2.00 – 2.00)	4.88	4.65 (2.23 – 8.41)	19.3	18.1 (9.15 – 30.6)

Table 4. Estimated values at different generations when $\lambda = 0.6$, $\mu = 0.4$ and $\phi = 0.5$.

generation (n)	$\tilde{m}_n$	$\frac{1}{n}\tilde{S}_n(\tilde{m}_n)$	$\frac{1}{n-1}\tilde{S}_{n-1}(\tilde{m}_n)$	estimated μ
2	1.400	7.200	2.888	–
3	1.000	5.600	1.100	–
4	0.9333	4.049	1.487	–
5	1.278	4.382	2.874	–
6	1.111	3.964	1.677	–
7	1.147	3.672	1.695	–
8	1.186	3.246	1.642	–
9	1.127	2.723	1.307	–
10	1.167	2.635	1.347	–
11	1.086	2.241	1.103	–
12	1.141	2.413	1.250	–
13	1.092	2.196	0.9483	–
14	1.122	2.238	1.030	–
15	1.113	2.081	0.9204	–
16	1.142	2.033	1.013	–
17	1.063	1.823	0.9735	–
18	1.049	2.086	1.514	–
19	1.128	2.261	2.362	–
20	1.112	2.907	2.745	–
25	1.087	2.766	2.703	–
50	1.082	2.849	3.142	–
100	1.082	2.928	3.750	–
150	1.080	2.613	3.431	0.2120
200	1.080	2.438	3.190	0.3665
250	1.080	2.478	3.424	0.4601
300	1.080	2.514	3.441	0.4220
350	1.080	2.529	3.404	0.3732
400	1.080	2.625	3.472	0.2302
450	1.080	2.481	3.318	0.3889
500	1.080	2.522	3.288	0.2852
550	1.080	2.502	3.225	0.2674
600	1.080	2.502	3.236	0.2776
650	1.080	2.415	3.162	0.3839
700	1.080	2.383	3.094	0.3871
750	1.080	2.398	3.059	0.3318

Table 5. Estimates from the 1997 Dutch CSF-outbreak; comparison between MLE and our current estimates.

			mle's		computed			estimates		
stage	weeks	observations	μ	λ	m	γ	γ_*	m	γ	γ_*
2	10	101	0.4	0.6	1.08	2.376	3.098	1.15	1.03	0.942
3	8	160	0.5	0.7	1.025	2.220	2.822	1.04	1.87	2.57
4	9	107	0.3	0.3	0.955	1.928	1.940	0.857	0.813	0.935
5	30	51	0.05	0.25	1.194	2.631	3.505	0.880	0.859	0.867

6.3. Real data

We also did the analysis of the data of the 1997 Dutch classical swine fever (CSF) outbreak as treated in Meester it et al. (2002). In that paper the outbreak is modelled by a Galton-Watson process. Because of changing measures of the government, the parameters μ and λ differ for different stages of the epidemic. For this reason the epidemic is divided into 5 stages. The time unit is one week and it is assumed that on average detections take place in the middle of a time interval, so ϕ is set to 0.5. In Meester it et al. (2002), an expensive algorithm is used find the maximum likelihood estimators for μ and λ in the different stages. Klinkenberg et al.(Klinkenberg (2003), Chapter 6) already showed that these maximum likelihood estimates are not reliable.

We will use our estimators to estimate m and γ for the different stages. We omit the first stage, because for that stage we have only one observation. We compare our estimates with the m and γ computed from the maximum likelihood estimates of λ and μ given in Meester it et al. (2002). The results are given in Table 6.3. We also give the duration of the stage (in weeks) and the number of observed individuals in a stage of the epidemic in this table.

In the second, third and fourth stage of the epidemic our estimate for m seems to be rather good, as we might expect. The estimate for γ doesn't seem to be very informative. In the final stage of the epidemic only few cases were observed and there were many weeks without any observation. Due to this few observations we may expect our estimators not to converge very fast. In none of the stages we could estimate μ and λ by using our estimated m and γ in (7), as the solutions of this system of equations gave no real μ between 0 and 1.

7. Conclusions

(A) From Theorem 1.3. of Guttorp it et al. (1991) we know that we cannot estimate more than two functions of parameters consistently only the generation sizes X_n are given. In Jacob it et al. (1998) it is showed that if we observe only a $Binomial(X_n, \mu)$ distributed fraction of the generation sizes, we can estimate two functions of parameters consistently, if μ is known. In this paper we have shown that, under certain conditions, we can estimate three functions of the parameters even when we do not know μ.

(B) For epidemiological purposes we want to estimate μ as well, because this parameter gives an indication of how many individuals are infectious at a certain time, which may be important for implementing measures. In order to estimate this parameter in reasonable time we apparently need more and other information. We can possibly get this information by using contact tracing, i.e. finding out what contacts are made by an individual before it was observed and which contact may have caused the infection. Sometimes it is possible to get experimental information about the time between infection and removal of an individual, from that information we may also estimate μ. Note that Becker it et al. (1997) are able to estimate μ and λ but they need information about the number of infectious individuals at the time of estimation, and this information is typically not available.

References

Andersson, H. and Britton, T. (2000) Stochastic epidemic models and their statistical analysis. *Springer lecture notes in Statistics* **151**. New York: Springer Verlag.

Becker, N. G. and Hasofer A. M. (1997) Estimation in Epidemics with Incomplete Observations. *Journal of the Royal Statistical Society. Series B(Methodological)* **59(2)** 415-429.

Diekmann, O. and Heesterbeek, J. A. P. (2000) *Mathematical Epidemiology of Infectious Diseases*. Chichester: John Wiley & Sons.

Grimmett, G. R. and Stirzaker, D. R. (1992) *Probability and Random Processes Second Edition*. New York: Oxford University Press Inc.

Guttorp, P. (1991) *Statistical Inference for Branching Processes*. New York: John Wiley & Sons.

Hall, P. and Heyde, C. C. (1980) *Martingale Limit Theory and its Applications*. New York: Academic Press.

Jacob, C. and Peccoud J. (1998) Estimation of the Parameters of a Branching Process from Migrating Binomial Observations. *Advances in Applied Probability* **30(4)** 948-967.

Jagers, P. (1975) *Branching Processes with Biological Applications*. London: John Wiley & Sons.

Klinkenberg, D. (2003) *Mathematical epidemiology and the conctrol of classical swine fever virus*. PhD Thesis Utrecht University.

Meester, R., de Koning, J., de Jong, M. C. M. and Diekmann, O. (2002) Modelling and Prediction of Classical Swine Fever Epidemics. *Biometrics* **41** 178-184.

Trapman, P., Meester, R. and Heesterbeek, J. A. P. (2004) A branching model for the spread of infectious animal diseases in varying environments. *Journal of Mathematical Biology* **49(6)** 553-576.