Stochastic Processes and their Applications 31 (1989) 307-314
North-Holland

WHAT CAN OR CAN'T BE ESTIMATED IN BRANCHING AND RELATED PROCESSES?

Ellen MAKI

Department of Biostatistics, Ontario Cancer Institute, 500 Sherbourne St., Toronto, Ontario, Canada

Philip McDUNNOUGH

Department of Statistics, University of Toronto, Toronto, Ontario, Canada

Received 23 April 1987
Revised 25 October 1988

Estimation of the underlying distribution is considered for the incompletely observed random walk and the incompletely observed Galton-Watson branching tree. Based on infrequent observation of a random walk, parameters not completely determined by the first few moments of the underlying distribution cannot be consistently estimated. A similar result is given for the branching tree when observations are sums of family sizes. When the offspring distribution belongs to the power series family MLE's are obtained from an approximate likelihood.

random walk * Galton-Watson branching tree * consistent estimation

1. Introduction

A Galton-Watson branching tree is constructed as follows. A fixed number, N, of individuals, ancestors of the tree, act independently to produce the next generation of the tree. Specifically, each of the N ancestors dies, leaving behind a random number of offspring who make up the first generation. This process continues, with members of one generation dying and leaving offspring to form the next generation. Let $p(\cdot|1)$ denote the offspring distribution and Z_n the total number of individuals born into generation n. The Galton-Watson process $\{Z_n; n=0, 1, 2, \ldots\}$ is known to be Markov with transition probabilities $P(Z_n=i|Z_{n-1}=j)=p(i|j)$, where $p(\cdot|j)$ is the j-fold convolution of $p(\cdot|1)$. In this article, only processes with $p(0|1)=0$ and offspring mean $\mu>1$ are considered (i.e. we restrict ourselves to processes that explode w.p. 1.).

We can go about making observations on the Galton-Watson tree in a number of ways. One possibility is to observe the process $\{Z_n\}$. For this scheme of observations, estimation of the mean, μ, and variance, σ^2, of $p(\cdot|1)$ has been considered by many authors (cf. Dion and Keiding (1977)). Also, Lockhart (1982) has shown under fairly general conditions that when the observations are $\{Z_n\}$ the only parameters of the offspring distribution that can be consistently estimated are μ and σ^2. At the other extreme is the case where the whole tree is observed and the entire offspring distribution can be estimated. Of course, this leads us to ask what happens in intermediate cases, including situations in which family sizes, rather than offspring numbers, are observed.

The observation of an individual's family size rather than their number of offspring is useful in the study of kin number problems (cf. Waugh (1981), Joffe and Waugh (1982)). The incomplete observational scheme where individuals are sampled from one or several consecutive generations and their family sizes recorded has been studied by the authors. The probability structure of this model was examined and estimates of μ and $p(\cdot|1)$ were obtained in Maki and McDunnough (1989).

In this article we consider the case where a random sample is taken at each of several consecutive generations and a total of family sizes is recorded for the sample from each generation. A similar type of problem arises in an incompletely observed random walk. Here observations consist of sums of i.i.d. random variables; the number of random variables comprising the sums may vary. Guttorp and Siegel (1985) have given a condition for the existence of consistent estimates of moments in the incompletely observed random walk. We will give a separate condition for the nonexistence of consistent estimates of certain parameters in the incompletely observed random walk. The results of Maki and McDunnough (1989) show that family sizes are asymptotically i.i.d. and that sums of family sizes are therefore asymptotically sums of i.i.d. random variables. Making use of this fact and using the results for the incompletely observed random walk, we will obtain separate conditions for the existence and nonexistence of certain parameters of the offspring distribution on the basis of observing sums of family sizes.

2. The incompletely observed random walk

Consider a random walk, $S_n = Y_1 + Y_2 + \cdots + Y_n$, where the random variables Y_i are i.i.d. We will consider only the case where the Y_i are non-negative integer-valued random variables and have probability function $p(\cdot|1)$.

Suppose that observations are made on the random walk at non-random (integer-valued) time points $n_0 \equiv 0$, n_1, $n_2, \ldots$. Denote the interobservation times by $r_i = n_i - n_{i-1}$, $i \geq 1$. Guttorp and Siegel (1985) have shown that if $\sum_i r_i^{2-k} = \infty$, then the first k moments of $p(\cdot|1)$ can be consistently estimated. For $k=3$ this condition is both necessary and sufficient. We will show here that if $\sum_i r_i^{-(k-2)/2} < \infty$, then we cannot consistently estimate any moment of order greater than k. While this does not provide a single condition which is both necessary and sufficient, it does provide a partial solution to the question of what can and cannot be consistently estimated for any given sequence of interobservation times $\{r_i\}$.

Define a doubly infinite sequence $\{r_{N,i}\}$ by putting $r_{N,i} = r_{N+i-1}$. For any given N, consider an incompletely observed random walk with interobservation times $r_{N,1}$, $r_{N,2}, \ldots$. If we define U_i to be $U_i = S_{n_i} - S_{n_{i-1}}$ where $n_i = \sum_{j=1}^{i} r_{N,j}$, then U_i is the sum of $r_{N,i}$ i.i.d. random variables. Recall that $p(\cdot|1)$ is the distribution underlying the random walk, and use $p(\cdot|j)$ to denote its j-fold convolution. The following lemma will be useful in proving a lack of consistent estimates. In what follows, lattice offset is defined as the smallest integer k for which $p(k|1)>0$ and lattice size is the greatest integer j for which $P(Y=k \bmod j)=1$.

Lemma 1. *Let $p(\cdot|1)$ and $q(\cdot|1)$ be two non-singular probability functions belonging to some family $\mathscr{P}$. Suppose that p and q have the same (finite) first k moments, the same lattice size and lattice offset. Then for the incompletely observed random walk with $\sum_i r_i^{-(k-2)/2}<\infty$ we have*

$$\lim_{N\to\infty}\sup_A |P^N(A)-Q^N(A)|=0,$$

where $P^N(A)=P((U_1, U_2, \ldots)\in A \mid$ interobservations times are $\{r_{N,i}\})$, and P, Q are probability measures corresponding to p, q.

Proof. We have $\sup_A |P^N(A)-Q^N(A)|=\lim_{j\to\infty}\sup_A |P^{N|j}(A)-Q^{N|j}(A)|$, where $P^{N|j}$ is the restriction of P^N to the sigma-field $\sigma(U_1, U_2, \ldots)$. Now

$$\begin{aligned}\sup_A &|P^{N|j}(A)-Q^{N|j}(A)|\\ &=\tfrac{1}{2}\sum_{m_1}\cdots\sum_{m_j}|P^{N|j}(U_1=m_1,\ldots,U_j=m_j)-Q^{N|j}(U_1=m_1,\ldots,U_j=m_j)|\\ &=\tfrac{1}{2}\sum_{m_1}\cdots\sum_{m_j}\left|\prod_{i=1}^{j}p(m_i|r_{N,i})-\prod_{i=1}^{j}q(m_i|r_{N,i})\right|. \qquad (1)\end{aligned}$$

From Theorem 22.1 of Bhattacharya and Ranga Rao (1976), a local limit theorem for lattice distributions, we know that there is a constant c, dependent on p and q, for which

$$\sum_{m_i=0}^{\infty}|p(m_i|r_{N,i})-q(m_i|r_{N,i})|\leq c r_{N,i}^{-(k-2)/2}.$$

From here it follows that

$$\sup_A |P^N(A)-Q^N(A)|\leq \lim_{j\to\infty}\frac{c}{2}\sum_{i=1}^{j} r_{N,i}^{-(k-2)/2}=\frac{c}{2}\sum_{i=N}^{\infty} r_i^{-(k-2)/2}.$$

The sum $\sum_{i=1}^{\infty} r_i^{-(k-2)/2}$ is finite, and so $\lim_{N\to\infty}\sum_{i=N}^{\infty} r_i^{-(k-2)/2}=0$, from which we conclude that $|P^N(A)-Q^N(A)|\to 0$ as $N\to\infty$.

Intuitively, what Lemma 1 tells us is that random walks having the same first k moments and defined on the lattice are indistinguishable on the basis of "infrequent" observations. This idea is presented formally in Theorem 1, where we show that parameters of the underlying distribution cannot be consistently estimated if they are not determined entirely by the first k moments and lattice size and offset.

Theorem 1. *Suppose that p and q are two distributions which satisfy the conditions of Lemma 1. Let θ be some functional of the distribution which is not completely determined by the first k moments, lattice size and lattice offset. If T is a consistent estimate of θ for the incompletely observed random walk with $\sum_{i=1}^{\infty} r_i^{-(k-2)/2} < \infty$, then $\theta(P) = \theta(Q)$, where P and Q are probability measures corresponding to p and q.*

Proof. Suppose that $\theta(P) \neq \theta(Q)$, and that $T(U_1, U_2, \ldots)$ is a consistent estimator of θ. For any fixed N we can choose n large enough so that

$$\sup_A |P^{N+n}(A) - Q^{N+n}(A)| < \varepsilon \tag{2}$$

where $0 < \varepsilon < 1$. Since p and q are not singular, there is some integer x for which $p(x|1)$ and $q(x|1)$ are both greater than zero. Define the event B as

$$B = \{U_1 = xr_{N,1}, \ldots, U_{n-1} = xr_{N,n-1}\}.$$

Clearly, each of $P^N(A)$ and $Q^N(A)$ is greater than zero.

Since T is a consistent estimator of θ, we must have $P^N(\{T(U_1, U_2, \ldots) = \theta(P)\}) = 1$. Now, we can write

$$\begin{aligned} &P^N(\{T(U_1, U_2, \ldots) = \theta(P)\} \cap B) \\ &\quad = P^N(B) P^N(T(xr_{N,1}, \ldots, xr_{N,n-1}, U_n, U_{n+1}, \ldots) = \theta(P)), \end{aligned}$$

from which we see that $P^N(T(xr_{N,1}, \ldots, xr_{N,n-1}, U_n, U_{N+1}, \ldots) = \theta(P)) = 1$. This last statement, however, is equivalent to $P^{N+n}(T(xr_{n,1}, \ldots, xr_{N,n-1}, U_1, U_2, \ldots) = \theta(P)) = 1$. Similarly, since $Q^N(B) > 0$ and $Q^N(T(U_1, U_2, \ldots) = \theta(P)) = 0$, we must have $Q^{N+n}(T(xr_{N,1}, \ldots, xr_{N,n-1}, U_1, U_2, \ldots) = \theta(P)) = 0$. It follows then that

$$\begin{aligned} &P^{N+n}(T(xr_{N,1}, \ldots, xr_{N,n-1}, U_1, U_2, \ldots) = \theta(P)) \\ &\quad - Q^{N+n}(T(xr_{N,1}, \ldots, xr_{N,n-1}, U_1, U_2, \ldots) = \theta(P)) = 1 \end{aligned}$$

which is a contradiction of (2). Therefore we must have $\theta(P) = \theta(Q)$.

As pointed out by Guttorp and Siegel (1985), a Galton–Watson process with offspring mean $\mu > 1$ grows at least as fast as an incompletely observed random walk with $r_i = \theta^i$ for any θ in the interval $(0, \mu)$. Thus, since $\sum_i \theta^{-i} < \infty$, moments of order greater than two cannot be consistently estimated from the generation sizes of a Galton–Watson tree. This is the result obtained by Lockhart (1982).

3. Incomplete observation of a Galton–Watson tree

Consider the Galton–Watson tree with $p(0|1) = 0$ and offspring mean $\mu > 1$. Suppose that $\{r_i\}$ is a sequence of non-negative integers with $P(Z_i > r_i, i \geq 1 | Z_0 = N) = 1$ for all $N \geq 1$, and define $r_{N,i} = r_{n+i-1}$. When the Galton–Watson process is begun with N ancestors, we will select $r_{N,i}$ individuals from those present at generation $\# i$ and observe their combined total offspring, say U_i. When $P(Z_i > r_{N,i+1}, i \geq 1 | Z_0 = N) = 1$ for all $N \geq 1$, the random variables $U_1, U_2, \ldots$ constitute an incompletely observed random walk. Thus, the results of Section 2 apply here. Extension of these results to the case where $P(Z_i > r_{N,i+1}, i \geq 1 | Z_0 = N) \to 1$ as $N \to \infty$ is straightforward, although messy, and is not presented here.

When the observations made are totals of family sizes, we will show that certain moments of the asymptotic family size distribution (denoted $p(\cdot)$) cannot be consistently estimated. Since there is a one-to-one correspondence between moments of $p(\cdot)$ and $p(\cdot|1)$ via the relation $p(i) = ip(i|1)/\mu$ (cf. Maki and McDunnough (1989)), moments of the offspring distribution $p(\cdot|1)$ will also not be estimable.

Suppose that observations S_i are made at generations $n \leq i \leq n + T_n$, each S_i being the total of r_i family sizes. In what follows, we will assume that $P(Z_i > r_i, i \geq 1 | Z_0 = N) = 1$, and that for some α in the interval $(0, -\ln E(1/Y)/\ln \mu)$ $(Y \sim p(\cdot|1))$, we have $\sum_{i=n}^{n+T_n} r_i^2 (N\mu^i)^{-\alpha} \to 0$ as $n \to \infty$. Under thes conditions, the asymptotic distribution of $S_n, S_{n+1}, \ldots, S_{n+T_n}$ is known (cf. Maki and McDunnough (1989)). The following lemma is similar to the lemma given for the incompletely observed random walk, and we will use it later to show the nonexistence of consistent estimates.

Lemma 2. *Let $p(\cdot|1)$ and $q(\cdot|1)$ be two non-singular distributions belonging to some class $\mathcal{P}$. Assume that p and q have the same mean $\mu > 1$, the same finite moments of order 2 through k, the same lattice size and lattice offset, and that $p(0|1) = q(0|1) = 0$. If $\lim_{n\to\infty} \sum_{i=n}^{n+T_n} r_i^{-(k-2)/2} = 0$, then*

$$\lim_{n\to\infty} \sup_C |P((S_n, \ldots, S_{n+T_n}) \in C) - Q((S_n, \ldots, S_{n+T_n}) \in C)| = 0.$$

Proof. We will use P_A and Q_A to denote probabilities calculated under the assumption that the family sizes are i.i.d. $ip(i|1)/\mu$ and $iq(i|1)/\mu$ respectively. Let $\mathbf{S}_n = (S_n, \ldots, S_{n+T_n})$ and $\mathbf{i}_n = (i_1, \ldots, i_{T_n+1})$. Then

$$\begin{aligned}\sup_C |P(\mathbf{S}_n \in C) - Q(\mathbf{S}_n \in C)| &= \tfrac{1}{2} \sum_{\mathbf{i}_n} |P(\mathbf{S}_n = \mathbf{i}_n) - Q(\mathbf{S}_n = \mathbf{i}_n)| \\ &\leq \tfrac{1}{2} \sum_{\mathbf{i}_n} |P(\mathbf{S}_n = \mathbf{i}_n) - P_A(\mathbf{S}_n = \mathbf{i}_n)| + \tfrac{1}{2} \sum_{\mathbf{i}_n} |Q(\mathbf{S}_n = \mathbf{i}_n) - Q_A(\mathbf{S}_n = \mathbf{i}_n)| \\ &\quad + \tfrac{1}{2} \sum_{\mathbf{i}_n} |P_A(\mathbf{S}_n = \mathbf{i}_n) - Q_A(\mathbf{S}_n = \mathbf{i}_n)|. \end{aligned} \tag{3}$$

Denote the terms in (3) t_1, t_2, and t_3. With regard to the first term t_1, we find that

$$\begin{aligned} t_1 &= \sum_{\mathbf{i}_n} |P(\mathbf{S}_n = \mathbf{i}_n) - P_A(\mathbf{S}_n = \mathbf{i}_n)| \\ &\leq \sum_{\mathbf{i}_n} |P(\mathbf{S}_n = \mathbf{i}_n) - P(\mathbf{S}_n = \mathbf{i}_n, I(D_{n,T_n}) = 1)| \\ &\quad + \sum_{\mathbf{i}_n} |P(\mathbf{S}_n = \mathbf{i}_n, I(D_{n,T_n}) = 1) - P_A(\mathbf{S}_n = \mathbf{i}_n)| \\ &= P(D_{n,T_n}^c) + \sum_{\mathbf{i}_n} |P(\mathbf{S}_n = \mathbf{i}_n, I(D_{n,T_n}) = 1) - P_A(\mathbf{S}_n = \mathbf{i}_n)| \end{aligned} \tag{4}$$

where D_{n,T_n} is the event that each of the $R_n = r_n + \cdots + r_{n+T_n}$ individuals in the combined sample belong to different families, and I is its indicator function. The probability $P(\mathbf{S}_i = \mathbf{i}_n, I(D_{n,T_n}) = 1)$ is obtained from Maki and McDunnough (1989). We have $P(D^c_{n,T_n}) \to 0$ as $n \to \infty$ and

$$\sum_{\mathbf{i}_n} |P(\mathbf{S}_n = \mathbf{i}_n, I(D_{n,T_n}) = 1) - P_A(\mathbf{S}_n = \mathbf{i}_n)|$$

$$\leqslant \sum_{j=1}^{T_n+1} \sum_{i_j=1}^{\infty} \sum_{B_j} \prod_{\alpha=1}^{T_n+1} \prod_{\beta=1}^{r_{n+\alpha-1}} l_{\alpha\beta} p(l_{\alpha\beta}|1)$$

$$\times \left| E\left(\binom{Z_{n-1}}{r_n} \binom{U_{n,i_1}}{r_n}^{-1} \prod_{\nu=1}^{T_n} \binom{U_{n+\nu-1,i_\nu}}{r_{n+\nu}} \binom{U_{n+\nu,i_{\nu+1}}}{r_{n+\nu}}^{-1} \right) - \mu^{-R_n} \right|. \quad (5)$$

In this last expression we have used B_j as the set of all $1 \times r_{n+j-1}$ vectors $\mathbf{l}_j$ for which $|\mathbf{l}_j| = i_j$. The random variables U_{n+k-1,i_k} are defined by $U_{n,i_1} = i_1 + \sum_{k=r_n+1}^{Z_{n-1}} Y_n^{(k)}$, and $U_{n+j-1,i_j} = i_j + \sum_{k=r_{n+j-1}}^{U_{n+j-1}} Y_{n+j-1}^{(k)}$ for $2 \leqslant j \leqslant T_n + 1$, where the $Y_i^{(j)}$ are i.i.d. $p(\cdot|1)$ and $Z_i = \sum_{j=1}^{Z_{i-1}} Y_i^{(j)}$. The quantity in (5) can be expressed as

$$\sum_{j=1}^{T_n+1} \sum_{l_j} \prod_{\alpha=1}^{T_n+1} \prod_{\beta=1}^{r_{n+\alpha-1}} l_{\alpha\beta} p(l_{\alpha\beta}|1)$$

$$\times \left| E\left(\binom{Z_{n-1}}{r_n} \binom{U_{n,i_1}}{r_n}^{-1} \prod_{\nu=1}^{T_n} \binom{U_{n+\nu-1,i_\nu}}{r_{n+\nu}} \binom{U_{n+\nu,i_{\nu+1}}}{r_{n+\nu}}^{-1} \right) - \mu^{-R_n} \right|$$

$$\leqslant \sum_{j=1}^{T_n+1} \sum_{l_j} \prod_{\alpha=1}^{T_n+1} \prod_{\beta=1}^{r_{n+\alpha-1}} l_{\alpha\beta} p(l_{\alpha\beta}|1)$$

$$\times E\left(\left| \binom{Z_{n-1}}{r_n} \binom{U_{n,i_1}}{r_n}^{-1} \prod_{\nu=1}^{T_n} \binom{U_{n+\nu-1}}{r_{n+\nu}} \binom{U_{n+\nu,i_{\nu+1}}}{r_{n+\nu}}^{-1} - \mu^{-R_n} \right| \right)$$

$$= E\left(\prod_{\alpha=0}^{T_n} \prod_{\beta=1}^{r_{n+\alpha}} Y_{n+\alpha}^{(\beta)} \left| \prod_{\nu=0}^{T_n} \binom{Z_{n+\nu-1}}{r_{n+\nu}} \binom{Z_{n+\nu}}{r_{n+\nu}}^{-1} - \mu^{-R_n} \right| \right).$$

It can be shown, as in Maki and McDunnough (1989), that this last expression $\to 0$ if $\sum_n^{n+T_n} r_i^2 \mu^{-k\alpha} \to 0$. Thus, the term t_1 in (3) converges to 0. Similar arguments show that $t_2 \to 0$.

In order to show that $t_3 \to 0$, we make use of the fact that under P_A and Q_A the random variables S_i are independent and each is the sum of i.i.d. random variables. The results for the partially observed random walk are therefore applicable, and we have

$$t_3 \leqslant c \sum_{j=1}^{T_n+1} r_{n+j-2}^{-(k-2)/2} = c \sum_{j=n}^{n+T_n} r_j^{-(k-2)/2}. \quad (6)$$

By assumption the last sum in (6) $\to 0$ and our proof is complete.

Proceeding as for the random walk, we use Lemma 2 to demonstrate that certain parameters of the offspring distribution cannot be consistently estimated from sums of family sizes. In what follows, P, Q are probability measures corresponding to p, q.

Theorem 2. *Suppose $p(\cdot|1)$ and $q(\cdot|1)$ are two offspring distributions which satisfy the conditions of Lemma 2, and let θ be some functional of the asymptotic family size distribution not determined entirely by the first k moments, lattice size and lattice offset. If $F_n \equiv F(S_n, \ldots, S_{n+T_n})$ is a weakly consistent estimator of θ and $\sum_n^{n+T_n} r_j^{-(k-2)/2} \to 0$, then we must have $\theta(P) = \theta(Q)$.*

Proof. Suppose that $\theta(P) \neq \theta(Q)$. If F_n is a consistent estimator of θ, then $P(|F_n - \theta(P)| < \delta) \to 1$ and $Q(|F_n - \theta(Q)| < \delta) \to 1$ as $n \to \infty$. We can choose n large enough so that we have

(i) $P(|F_n - \theta(P)| < \delta) > 1 - \varepsilon$ and $Q(|F_n - \theta(P)| < \delta) < \varepsilon$, and

(ii) $\sup_C |P((S_n, \ldots, S_{n+T_n}) \in C) - Q((S_n, \ldots, S_{n+T_n}) \in C)| < \varepsilon$.

These two conditions clearly contradict each other since the first implies that

$$|P(|F_n - \theta(P)| < \delta) - Q(|F_n - \theta(P)| < \delta)| > 1 - 2\varepsilon > \varepsilon$$

for an appropriate choice of ε. Thus we must have $\theta(P) = \theta(Q)$.

4. Parametric estimation

In Maki and McDunnough (1989) the problem of estimating the offspring distribution from an approximate likelihood was considered for the case where the offspring distribution was a member of the power series family and the entire vector of family sizes was observed. We will show here that by observing family size totals and using the asymptotic likelihood function, we obtain the same estimates as when we observe individual family sizes.

Assume that the offspring distribution is of the form $p(i|1) = \theta^i a_i A(\theta)^{-1}$, $i = 1, 2, 3, \ldots$ where $\theta > 0$, the a_i are known constants, and $A(\theta) = \sum \theta^i a_i$. The asymptotic family size distribution, in this case, is $p(i) = (ia_i)\theta^{i-1}\{A'(\theta)\}^{-1}$, also a member of the power series family. The estimate of $f(\theta) = 1 + \theta A''(\theta)/A'(\theta)$, obtained by maximizing the asymptotic likelihood function and based on observing individual family sizes, is $\bar{X} = R_n^{-1} \sum_{i=n}^{n+T_n} \sum_{j=1}^{r_{n+i}} X_i^{(j)}$. Now, the asymptotic likelihood function can be expressed as

$$\{A'(\theta)\}^{-R_n} \theta^{R_n(\bar{X}-1)} \prod_{i=n}^{n+T_n} \prod_{j=1}^{r_{n+i}} X_i^{(j)} a_{X_i^{(j)}}$$

and so it follows from the factorization theorem that $\bar{X}$ is a sufficient statistic.

A result which is useful here is one due to Keiding and Lauritzen (1978). They have shown that if $t(X)$ is sufficient and is the MLE of some function of θ, say $\tau(\theta)$, based on observation of X, then it is also the MLE of $\tau(\theta)$ when only $t(X)$ is observed. Consequently, in the problem considered here, $\bar{X}$ is also the MLE of $f(\theta)$ when we observe only family size totals. Thus, the results given in Maki and

McDunnough (1989) for estimating θ, μ, and $p(\cdot\,|1)$ continue to hold under this scheme of incomplete observation.

While this result may also hold when $p(\cdot\,|1)$ belongs to some other parametric family, it will clearly not be true in general.

5. Conclusion

In this article, we have presented results for the incompletely observed random walk and the incompletely observed branching tree. We have shown, in Section 2, that when a random walk is observed "infrequently" (i.e. $\sum r_i^{-(2k-3)/2} < \infty$), moments of order greater than $2k-1$ cannot be consistently estimated. This result is a complement to that of Guttorp and Seigel (1985), who have shown that when "frequent" observations are made on the random walk (i.e. $\sum r_i^{-(k-2)} = \infty$), the first k moments can be consistently estimated. While these two results together do not provide a single condition which is both necessary and sufficient for the consistent estimation of the first k moments, they do give a partial answer to the question of what can or cannot be estimated in a partially observed random walk.

The incomplete observations of a branching tree considered in this article were sums of family sizes. Using the fact that family sizes are asymptotically i.i.d., we showed that the results for the partially observed random walk also hold for the incompletely observed branching tree.

When the offspring distribution is known to belong to some parametric family, we expect to be able to estimate more than in the nonparametric case. Indeed, we have shown that this is true for the power series family of offspring distributions. Using an approximate likelihood, we are able to obtain an estimate of $p(\cdot\,|1)$ which in fact coincides with that obtained from observation of individual family sizes.

References

[1] J.P. Dion, and N. Keiding, Statistical inference in branching processes, A. Joffe and P. Ney, eds., Branching Processes (Marcel Dekker, New York, 1978) 105-140.

[2] R.N. Bhattacharya, and R. Ranga Rao, Normal Approximation and Asymptotic Expansions (John Wiley & Sons, New York, 1976).

[3] P. Guttorp, and A.F. Siegel, Consistent estimation in partially observed random walks, Ann. Statist. 19 (1985) 474-494.

[4] A. Joffe, and W.A.O'N. Waugh, Exact distributions of kin numbers in a Galton-Watson process, J. Appl. Prob. 19 (1982) 767-775.

[5] N. Keiding, and S.L. Lauritzen, Marginal maximum likelihood estimates and estimation of the offspring mean in a branching process, Scand. J. Statist. 5 (1978) 106-110.

[6] R.A. Lockhart, On the nonexistence of consistent estimates in Galton-Watson processes, J. Appl. Prob. 19 (1982) 842-846.

[7] E. Maki and P. McDunnough, Sampling a branching tree, Stochastic Process. Appl. 31 (1989) 283-305 (this issue).

[8] W.A.O'N. Waugh, Application of the Galton-Watson process to the kin number problem, Adv. Appl. Prob. 13 (1981) 631-649.