The Annals of Statistics
1979, Vol. 7, No. 3, 680–685

ESTIMATION PROBLEMS IN BRANCHING PROCESSES WITH RANDOM ENVIRONMENTS

BY J. P. DION[1] AND W. W. ESTY

Université du Québec à Montréal and Montana State University

Estimators are given for the two main parameters of the branching process with random environments, namely, the mean of the environmental mean (the average number of offspring per individual) and the mean of the logarithm of the environmental mean. The latter parameter indicates the certain extinction or possible explosion of the process. The consistency and asymptotic normality of the estimators are shown in the context of Smith and Wilkinson. Consistent estimators are also given for the corresponding variances and thus confidence intervals are obtained for the two main parameters of the system.

1. Introduction. The branching process with random environment (B.P.R.E.) as introduced by Smith (1968) and Smith and Wilkinson (1969) and generalized by Athreya and Karlin (1971) is particularly well adapted to describe population growth. It has been considered, for example, in the case of the growth of the placenta of a pregnant woman (Winkel et al., 1976) or for the study of the whooping crane population of North America (Keiding, 1976).

More recently, Becker (1977) used it to model the growth of an epidemic and proposed a strongly consistent estimator for the criticality parameter. Apart from that, estimation problems had yet received no systematic treatment. In this work, estimators are given for two of the main parameters of the B.P.R.E., namely the reproduction average per individual and the mean of the logarithm of the "environmental mean," this latter parameter the criticality parameter, indicating the certain extinction or possible explosion of the process. The consistency and asymptotic normality of the estimators are shown in the general context of Smith and Wilkinson (1969). Consistent estimators are furthermore provided for the associated variances.

Throughout the text, the notation is that of Athreya and Karlin (1971). Assume that $\{\zeta_n, n = 0, 1, \cdots\}$, the environmental sequence, is formed by i.i.d. random variables, with values in Θ. For every $\zeta \in \Theta$, consider the p.g.f., $\varphi_\zeta(s) = \Sigma_{j=0}^{\infty} p_j(\zeta)s^j$. For each realization of the process $\{\zeta_n\}$ and the associated random sequences of p.g.f.'s, there evolves a population Z_n, governed by the laws of the standard temporally nonhomogeneous branching process.

The problems studied here are the estimation of $\mu = E\varphi_\zeta'(1)$ and of $\pi = E(\log \varphi_\zeta'(1))$. Since $E(Z_n|Z_0 = 1) = \mu^n$ and since, modulo mild integrability conditions, extinction is certain iff $\pi \leqslant 0$, there is considerable interest in knowing the

Received April 1977; revised December 1977.

[1]Research supported by NRCC under grant A8852.

AMS 1970 *subject classifications.* Primary 62M05; secondary 60J80.

Key words and phrases. Branching process, random environment, estimation, consistency and asymptotic normality.

value of these parameters. Point estimators for μ and π are given in Theorems 2 and 4, and asymptotic confidence intervals in Corollaries 2 and 3.

We will assume throughout that the process is *supercritical* (i.e., $E(\log \varphi_\zeta'(1)) > 0$), that $E[|\log \varphi_\zeta'(1)|] < \infty$ and $E(-\log(1 - \varphi_{\zeta_0}(0)) < \infty$, that $\mu \equiv E(\varphi_\zeta'(1)) < \infty$, $0 < \sigma^2 \equiv \mathrm{Var}(\varphi_\zeta'(1)) < \infty$ and $0 < \gamma^2 \equiv E(\mathrm{Var}(Z_1|\zeta_0, Z_0 = 1)) < \infty$. We will first suppose that $p_0(\zeta) = 0$, $\forall\zeta$, and later remove that assumption.

2. Estimation of μ and σ^2. Let us consider the sample $\{Z_0, Z_1, \cdots, Z_n\}$ formed by the first n generations of the B.P.R.E., with $Z_0 \equiv 1$.

Of course, if one knew $(\varphi_{\zeta_0}'(1), \cdots, \varphi_{\zeta_n}'(1))$, one would use the arithmetic mean $n^{-1}\Sigma_{k=0}^{n-1}\varphi_{\zeta_k}'(1)$ as an estimator for μ. This would give an unbiased, strongly consistent and asymptotically normal estimator for μ, since the $\{\varphi_{\zeta_k}'(1)\}$ are i.i.d. with finite mean and variance. This suggests $\tilde{\mu} = n^{-1}\Sigma_{k=0}^{n-1}(Z_{k+1}/Z_k)$ as an estimator for μ.

THEOREM 1. *Let* $\tilde{\mu} = n^{-1}\Sigma_{k=0}^{n-1}(Z_{k+1}/Z_k)$ *and assume that* $p_0(\zeta) = 0$, $\forall\zeta$, $\mu < \infty$, $0 < \sigma^2, \gamma^2 < \infty$. *Then*

(i) $E(\tilde{\mu}) = \mu$;

(ii) $\mathrm{Var}(\tilde{\mu}) = n^{-1}\sigma^2 + n^{-2}\gamma^2\Sigma_{k=0}^{n-1}E(Z_k^{-1})$.

PROOF. (i) Since $E(Z_{k+1}/Z_k|Z_k, \zeta_k) = \varphi_{\zeta_k}'(1)$, it follows that $E(\tilde{\mu}) = \mu$.

(ii) $\mathrm{Var}(\tilde{\mu}) = n^{-2}[\Sigma_{k=0}^{n-1}\mathrm{Var}(Z_{k+1}/Z_k) + 2\Sigma_{j<k}\mathrm{Cov}(Z_{k+1}/Z_k, Z_{j+1}/Z_j)]$. Since $E((Z_{k+1}/Z_k)(Z_{j+1}/Z_j)|Z_k, Z_j, Z_{j+1}, \zeta_k) = \varphi_{\zeta_k}'(1) \cdot Z_{j+1}/Z_j$ one has $E((Z_{k+1}/Z_k)(Z_{j+1}/Z_j)|Z_k, Z_{j+1}, Z_j) = \mu Z_{j+1}/Z_j$ or $E((Z_{k+1}/Z_k)(Z_{j+1}/Z_j)) = \mu^2$, implying $\mathrm{Cov}(Z_{k+1}/Z_k, Z_{j+1}/Z_j) = 0$. Also $\mathrm{Var}(E(Z_{k+1}/Z_k|Z_k, \zeta_k)) = \mathrm{Var}\varphi_{\zeta_k}'(1) = \sigma^2$ and $E(\mathrm{Var}(Z_{k+1}/Z_k|Z_k, \zeta_k)) = E(\gamma^2(\zeta_k)Z_k^{-1})$, where $\gamma^2(\zeta_k)$ represents the conditional variance of the number of offspring per individual, given the environment ζ_k. Since ζ_k is independent of Z_k (but not of Z_{k+1}, however) and since the $\{\zeta_k\}$ are i.i.d., $E(\gamma^2(\zeta_k)Z_k^{-1}) = E(\gamma^2(\zeta_k))E(Z_k^{-1}) = \gamma^2 \cdot E(Z_k^{-1})$. It follows that $\mathrm{Var}(Z_{k+1}/Z_k) = \sigma^2 + \gamma^2 \cdot E(Z_k^{-1})$, a fact which emphasizes the interplay between environmental and demographic variability (cf. Keiding, 1976, page 149). Thus $\mathrm{Var}(\tilde{\mu}) = n^{-1}\sigma^2 + n^{-2}\gamma^2\Sigma_{k=0}^{n-1}E(Z_k^{-1})$.

COROLLARY 1. *Under the hypotheses of Theorem* 1,

$$\tilde{\mu} \to_p \mu.$$

Intuitively, one expects strong consistency and asymptotic normality to hold for $\tilde{\mu}$, *since* Z_{k+1}/Z_k *is a good estimator for* $\varphi_{\zeta_k}'(1)$, *given* ζ_k, *and since the sequence* $\{\varphi_{\zeta_k}'(1)\}$ *obeys the strong law of large numbers and the central limit theorem. The proof of the next theorem makes only more precise this heuristic derivation.*

THEOREM 2. *Under the hypotheses of Theorem* 1,

(i) $\tilde{\mu} \to_{\mathrm{a.s.}} \mu$ *as* $n \to \infty$;

(ii) $n^{\frac{1}{2}}(\tilde{\mu} - \mu)/\sigma \to_D N(0, 1)$ *as* $n \to \infty$.

PROOF. Put $\bar{\mu} = n^{-1}\Sigma_{k=0}^{n-1}\varphi'_{\zeta_k}(1)$. The part (i) will be proved by showing that $(\tilde{\mu} - \bar{\mu}) \to_{a.s.} 0$ and the part (ii) by proving that $n^{\frac{1}{2}}(\tilde{\mu} - \bar{\mu})/\sigma \to_P 0$.

(i) Recall (Athreya & Karlin, 1971, Theorem 1, page 1845) that $Z_n/\Sigma_{k=0}^{n-1}\varphi'_{\zeta_k}(1) \to_{a.s. n\to\infty} W > 0$. Hence $Z_n/Z_{n-1} - \varphi'_{\zeta_{n-1}}(1) \to_{a.s.} 0$, which implies that $(n^{-1}\Sigma_{k=0}^{n-1}(Z_{k+1}/Z_k) - n^{-1}\Sigma_{k=0}^{n-1}\varphi'_{\zeta_k}(1)) \to_{a.s.} 0$.

(ii) Since $E(\tilde{\mu} - \bar{\mu}) = 0$, we will show that $\text{Var}(n^{\frac{1}{2}}(\tilde{\mu} - \bar{\mu})) \to 0$, then appeal to Tchebyshev's inequality to conclude that $n^{\frac{1}{2}}(\tilde{\mu} - \bar{\mu}) \to_p 0$.

$$\text{Var}\big(n^{\frac{1}{2}}(\bar{\mu} - \tilde{\mu})\big) = n^{-1}\Big[\Sigma_{k=0}^{n-1}\text{Var}(\varphi'_{\zeta_k}(1) - Z_{k+1}/Z_k) + 2\Sigma_{j<k}\text{Cov}\big((\varphi'_{\zeta_k}(1) - Z_{k+1}/Z_k), (\varphi'_{\zeta_j}(1) - Z_{j+1}/Z_j)\big)\Big].$$

But $E((\varphi'_{\zeta_k}(1) - Z_{k+1}/Z_k)(\varphi'_{\zeta_j}(1) - Z_{j+1}/Z_j)|Z_k, Z_{j+1}, Z_j, \zeta_k, \zeta_j) = 0$ and thus $\text{Cov}((\varphi'_{\zeta_k}(1) - Z_{k+1}/Z_k), (\varphi'_{\zeta_j}(1) - Z_{j+1}/Z_j)) = 0$. Furthermore, $\text{Var}(\varphi'_{\zeta_k}(1) - Z_{k+1}/Z_k) \to_{k\to\infty} 0$, since $\text{Var}(\varphi'_{\zeta_k}(1) - Z_{k+1}/Z_k) = \gamma^2 E(Z_k^{-1})$ as shown previously. So $n^{-1}\Sigma_{k=0}^{n-1}\text{Var}(\varphi'_{\zeta_k}(1) - Z_{k+1}/Z_k) \to 0$, which completes the proof.

NOTE. It is interesting that an asymptotic confidence interval for μ can be deduced if $\sigma^2 > 0$ is known. Knowledge of γ^2 would not be required, a situation different from the classical Galton-Watson case.

If σ^2 is unknown, it can be estimated by its corresponding empirical variance s_φ^2, where $s_\varphi^2\ (n-1)^{-1}\Sigma_{k=0}^{n-1}(Z_{k+1}/Z_k - \tilde{\mu})^2$. That it is strongly consistent but will tend to slightly overestimate σ^2 is proved in the next theorem.

THEOREM 3. *Under the hypotheses of Theorem* 1,
(i) $E(s_\varphi^2) = \sigma^2 + n^{-1}\gamma^2\Sigma_{k=0}^{n-1}E(Z_k^{-1})$; *and*
(ii) $s_\varphi^2 \to_{a.s. n\to\infty} \sigma^2$.

PROOF. Note that $s_\varphi^2 = (n-1)^{-1}\Sigma_{k=0}^{n-1}(Z_{k+1}/Z_k - \mu)^2 - n(n-1)^{-1}(\tilde{\mu} - \mu)^2$. By taking the expectation and invoking Theorem 1, one obtains part (i). For part (ii), one uses further the decomposition

$$(Z_{k+1}/Z_k - \mu)^2 = \big(Z_{k+1}/Z_k - \varphi'_{\zeta_k}(1)\big)^2 + \big(\varphi'_{\zeta_k}(1) - \mu\big)^2 + 2\big(Z_{k+1}/Z_k - \varphi'_{\zeta_k}(1)\big)\big(\varphi'_{\zeta_k}(1) - \mu\big).$$

Recall that if $\{x_n\}$ and $\{y_n\}$ are two sequences of real numbers with $x_n \to 0$ and $n^{-1}\Sigma_{k=1}^{n}|y_k| \to C < \infty$, one has $n^{-1}\Sigma_{k=1}^{n}x_k y_k \to 0$.

Since $(Z_{k+1}/Z_k - \varphi'_{\zeta_k}(1)) \to_{a.s.} 0$, $n^{-1}\Sigma_{k=0}^{n-1}(Z_{k+1}/Z_k - \varphi'_{\zeta_k}(1))(\varphi'_{\zeta_k}(1) - \mu) \to_{a.s.} 0$. Since also $\tilde{\mu} - \mu \to_{a.s.} 0$, it follows that $s_\varphi^2 - (n-1)^{-1}\Sigma_{k=0}^{n-1}(\varphi'_{\zeta_k}(1) - \mu)^2 \to_{a.s.} 0$ and by the strong law of large numbers, that

$$s_\varphi^2 \to_{a.s.} E\big(\varphi'_{\zeta_k}(1) - \mu\big)^2 = \sigma^2.$$

COROLLARY 2. *Under the hypothesis of Theorem* 1,

$$n^{\frac{1}{2}}(\tilde{\mu} - \mu)/s_\varphi \to_D N(0, 1).$$

3. Estimation of π and Var log $\varphi_\zeta'(1)$. Consider now $\pi = E \log \varphi_\zeta'(1)$. If $(\varphi_{\zeta_0}'(1), \cdots, \varphi_{\zeta_{n-1}}'(1))$ is known, an obvious and nice estimator for π is the corresponding arithmetic mean

$$\bar{X}_{\log} = n^{-1}\Sigma_{k=0}^{n-1}\log \varphi_{\zeta_k}'(1).$$

Its properties are summarized below:

PROPOSITION A. *If* Var log $\varphi_\zeta'(1)$ *is positive and finite,*

(i) $E(\bar{X}_{\log}) = \pi$;

(ii) $\bar{X}_{\log} \to_{\text{a.s.}} \pi$ *as* $n \to \infty$; *and*

(iii) $n^{\frac{1}{2}}(\bar{X}_{\log} - \pi) \to_D N(0, \text{Var} \log \varphi_\zeta'(1))$.

Using the sample $\{Z_0, Z_1, \cdots, Z_n\}$, one would thus expect $\tilde{\pi} = n^{-1}\Sigma_{k=0}^{n-1}\log(Z_{k+1}/Z_k) = (\log Z_n)/n$ to behave nicely for estimation purposes. This estimator was studied by Heyde (1975, page 50) in a different context. In fact, the following properties hold here:

THEOREM 4. *Assume* Var log $\varphi_\zeta'(1)$ *is positive and finite. Let* $\tilde{\pi} = (\log Z_n)/n$. *Under the hypotheses of Theorem* 1,

(i) $E(\tilde{\pi}) \leqslant \pi$;

(ii) $\tilde{\pi} \to_{\text{a.s.}} \pi$ *as* $n \to \infty$; *and*

(iii) $n^{\frac{1}{2}}(\tilde{\pi} - \pi) \to_D N(0, \text{Var} \log \varphi_\zeta'(1))$.

PROOF. By Jensen's inequality, $E(\log(Z_{k+1}/Z_k)|\zeta_k) \leqslant \log E(Z_{k+1}/Z_k|\zeta_k) = \log \varphi_{\zeta_k}'(1)$. This implies $E \log(Z_{k+1}/Z_k) \leqslant E \log \varphi_{\zeta_k}'(1) = \pi$. Hence $E(\tilde{\pi}) \leqslant \pi$. Furthermore, since $(Z_n)/\Sigma_{k=0}^{n-1}\varphi_{\zeta_k}'(1) \to_{\text{a.s.}} W > 0$, then $\log Z_n - \Sigma_{k=0}^{n-1}\log \varphi_{\zeta_k}'(1) \to_{\text{a.s.}} \log W$, from which $\forall \alpha > 0$, $(\log Z_n)/n^\alpha - \Sigma_{k=0}^{n-1}\log \varphi_{\zeta_k}'(1)/n^\alpha \to_{\text{a.s.}} 0$. Invoking Proposition A, one readily obtains (ii) and (iii).

NOTES. 1. The properties (i) and (ii) tell us that $\tilde{\pi}$ will tend to underestimate π although it is strongly consistent. An equivalent version of (ii) was given by Becker (1977).

2. The property (iii) may be of independent interest since it does not depend on the limiting distribution of W. If π were known, it could be used to predict the population size at time n, giving a partial solution to a problem encountered by Keiding (1976) while studying the whooping crane population. The property (iii) is also strongly related to part (b) of the theorem in Keiding & Nielson (1973).

To give a confidence interval for π, one needs now a consistent estimator for Var log $\varphi_\zeta'(1)$, in the case when it is unknown. A natural estimator which has this property is the corresponding sample variance:

$$s_{\log}^2 = (n-1)^{-1}\Sigma_{k=0}^{n-1}(\log(Z_{k+1}/Z_k) - \tilde{\pi})^2.$$

THEOREM 5. *Under the hypotheses of Theorem* 4,

$$s_{\log}^2 \to_{\text{a.s.}} \text{Var} \log \varphi_\zeta'(1).$$

PROOF. The proof is entirely analogous to the one in Theorem 3 (ii), using the fact that $\log(Z_{k+1}/Z_k) - \log \varphi'_{\zeta_k}(1) \to_{\text{a.s.}} 0$.

COROLLARY 3. *Under the hypotheses of Theorem* 4,

$$n^{\frac{1}{2}}(\tilde{\pi} - \pi)/s_{\log} \to_D N(0, 1).$$

4. Extension to the case $p_0(\zeta) \geqslant 0$. Assume now that $p_0(\zeta) \geqslant 0$, $\forall \zeta$ and let A be the set of nonextinction of the supercritical B.P.R.E. It is known that $P(A) > 0$. Intuitively, one expects the asymptotic results of Sections 2 and 3 to hold on the set A. And of course if they hold on the set A, they will hold conditionally on $Z_n > 0$, which is what is really needed for inference purposes.

If the B.P.R.E. is defined on some probability space (Ω, B, P), denote by $P_A(\cdot) = P(\cdot|A)$. Then it is clear that with respect to P_A,

$$Z_n/\textstyle\prod_{k=0}^{n-1}\varphi'_{\zeta_k}(1) \to_{\text{a.s.}} W > 0.$$

Also note that if under P, the environmental process is formed by i.i.d. random variables, then the environmental process will still obey the strong law of large numbers and the central limit theorem under any $Q \ll P$, in particular $Q \equiv P_A$ (cf. Renyi, 1958), provided second moments exist. These are the basic tools needed to extend the asymptotic results of Section 2 and 3 to the case $p_0(\zeta) \geqslant 0$. Minor obvious modifications make the proofs go through, with the exception of Theorem 2, part (ii), which does not seem to admit an extension by the techniques of this paper.

5. Conclusion. The estimators for μ and π have been shown to be consistent and asymptotically normal in the context of Smith and Wilkinson. It seems worthwhile to extend these results to the case of random but varying environment and to study the multitype case.

The authors found at the "Conference on Galton-Watson processes and related topics" a stimulating "random" environment. Thanks are due to S. Durham and N. Keiding for interesting discussions and to the referee for many comments and suggestions and for pointing out the existence of Becker's paper.

REFERENCES

ATHREYA, K. B. and KARLIN, S. (1971). On branching processes with random environments; I-II. *Ann. Math. Statist.* **42** 1499–1520 and 1843–1858.

ATHREYA, K. B. and NEY, P. E. (1972). *Branching processes*. Springer, Berlin.

BECKER, N. (1977). Estimation for discrete time branching processes with application to epidemics. *Biometrics* **33** 515–522.

HEYDE, C. C. (1975). Remarks on efficiency in estimation for branching processes. *Biometrika* **62** 49–55.

KEIDING, N. (1975). Extinction and exponential growth in random environments. *Theor. Pop. Biol.* **8** 49–63.

KEIDING, N. (1976). Population growth and branching processes in random environments. *Proc. 9, Biom. Int. Conf., Inv. Papers II* 149–165.

KEIDING, N. and NIELSON, J. E. (1973). The growth of supercritical branching processes with random environments. *Ann. Probability* **1** 1065–1067.

RENYI, A. (1958). On mixing sequences of sets. *Acta Math. Acad. Sci. Hungar.* **9** 215–228.

SMITH, W. L. (1968). Necessary conditions for almost sure extinction of a branching process with random environment. *Ann. Math. Statist.* **39** 2136–2140.

SMITH, W. L. and WILKINSON, W. (1969). On branching processes in random environments. *Ann. Math. Statist.* **40** 814–827.

WINKEL, P., GAEDE, P. and LYNGBYE, J. (1976). A method for the monitoring of hormone levels in pregnancy. With a statistical appendix by S. L. Lauritzen. *J. Clin. Chem.* **22** 422–428, 196.

DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITÉ DU QUÉBEC À MONTRÉAL
CASE POSTALE 6128, SUCCURSALE "A"
MONTRÉAL, P.Q. H3C 3J7 CANADA