

Chris Heyde on Branching Processes and Population Genetics

Eugene Seneta¹

When Chris Heyde returned to Australia in September 1968 to take up a position as Reader in Ted Hannan's teaching Department of Statistics in the-then School of General Studies at the Australian National University, there was still a strong trend of research into population genetics emanating from Pat Moran's purely research department in the Institute of Advanced Studies, a direction synthesized in Moran (1962). There was also strong activity in both departments in the theory of branching processes. With its common focus with population genetics on reproduction and extinction, it was the primary interest at the time of the author of this commentary. A strong impetus to branching process theory had occurred with the appearance of the book of Harris (1963).

The dominant theme of Chris's work prior to his arrival had been refinement of classical limit theory for sums of independent random variables. A principal theme of his work since 1970 has been the theory and use of martingale methods in situations of statistical dependence. This seems to have its origin in his paper [M31], in the context of a simple (Bienaymé-Galton-Watson) branching process $\{Z_n\}, n \geq 0; Z_0 = 1$. Suppose the non-degenerate offspring distribution has probability generating function

$$F(s) = \sum_{j=0}^{\infty} s^j P(Z_1 = j), s \in [0, 1].$$

When $1 < m = E(Z_1) < \infty$ the process is called supercritical, and in this case the probability of extinction, q , is the unique real number in $[0, 1)$ satisfying $F(q) = q$. The random sequence $\{Z_n/m^n\}$ is clearly a non-negative martingale, which in the supercritical case under the-then standard condition $E(Z_1^2) < \infty$ was known to converge almost surely (a.s.) to a random variable W for which $P(W=0) = q$ and which has a continuous density on the set of positive real numbers. After a number of refinements of this convergence result, it was shown in Seneta (1968) that in the non-degenerate supercritical case without further conditions, the choice $c_n = 1/h_n(s_0)$, where s_0 is any fixed number in $(0, -\log q)$, and $h_n(s_0)$ is the n -th functional iterate of $h(s)$, which is the inverse function of $k(s) = -\log F(e^{-s})$, results in convergence in distribution of $\{Z_n/c_n\}$ to a proper non-degenerate random variable, W . While it was noticed that $\{W_n = h_n(s_0)Z_n\}$ was a submartingale, a.s. convergence could not be deduced from the usual theorem since $E(W_n)$ is bounded if and only if $E(Z_1 \log Z_1) < \infty$. Chris's insight was that $\{\exp(-h_n(s_0)Z_n)\}$ is a martingale, with (obviously) bounded mean, thus giving the sharpening to a.s. convergence. Reciprocally, this strengthening immediately led to an analogous result for the supercritical branching process with immigration $\{X_n\}$ (Seneta, 1970), which then led to collaboration with Chris in [M42] on analogues of the Central Limit Theorem and the Law of the Iterated Logarithm as refinements of the results

$$X_n / m^n \rightarrow V, \text{ a.s.,}$$

¹School of Mathematics and Statistics, FO7, University of Sydney, N.S.W. 2006, Australia
e-mail: eseneta@maths.usyd.edu.au

and

$$X_{n+r} / X_n \rightarrow m^r, \text{ a.s.}$$

The note [M31] stimulated Chris to develop theory for the supercritical process analogous to the classical rate of convergence theorems for the difference $\{Z_n/m^n - W\}$ under the classical condition $E(Z_1^2) < \infty$. The central limit theorem analogue developed in [M34] and [M38] is beautifully attained with random norming and the necessary conditioning:

$$\lim_{n \rightarrow \infty} Pr\{(m^2 - m)^{1/2} \sigma^{-1} Z_n^{-1/2} (m^n W - Z_n) \leq x \mid Z_n > 0\} = \Phi(x), \quad (1)$$

where $\sigma^2 = \text{Var}Z_1$. This result is followed in [M37] by a corresponding iterated logarithm analogue, using an intermediate result from [M39], under the condition $E(Z_1^3) < \infty$, later relaxed in [M41] to $E(Z_1^2) < \infty$. These papers display a finely-honed appreciation of, and ability to adapt in a non-trivial way, and to a quite different context, tools of the classical limit theory.

The note [M31] had several other important aspects. The chief of these was the reinvigoration of the theory of statistical estimation of m on the basis of a single finite observed realization, including a focus on estimation theory for a stochastic process in which extinction may occur. In [M62] Chris returns to the investigation of m under the second moment constraint, and under the simplifying assumption that $P(Z_1=0)=0$ so that $P(W>0)=1$. The starting point is a comparison of the maximum-likelihood estimator $\hat{m}_n = (Y_n - 1)Y_{n-1}^{-1}$, where $Y_n = 1 + Z_1 + \dots + Z_n$, and the simple ratio estimator $\tilde{m}_n = Z_n/Z_{n-1}$. There are different non-random normings of each of $\hat{m}_n - m$, $\tilde{m}_n - m$, which result in a limit law of form $W^{-1/2}Z$, where $Z \sim \mathcal{N}(0,1)$, on the basis of which he concludes that asymptotic relative efficiency of \tilde{m}_n relative to \hat{m}_n is $1 - m^{-1}$. Thus the limit law is normal variance-mixing. Chris then shows that under a power-series condition, \hat{m}_n is asymptotically efficient (according to an extension of Rao's concept of efficiency). Here a central limit theorem for martingales is a dominant tool, arising from the fact that $\{Z_n/m^n\}$ is a martingale.

Chris's inclination towards a framework of time-series type results in which his martingale work could be applied was hugely influential, and the extensive use of martingale central limit theory as in [M54] and [M62] for inferential purposes, was at least partly stimulated by appearance of a paper of (Bruce) Brown (1971). An early manifestation in [M46] stems from noticing that $\{Y_n/m^n\}$ is a martingale, where $Y_n = X_n - \mu$, and $\{X_n\}$ is a subcritical ($0 < m < 1$) and hence (under moment conditions) an asymptotically stationary branching process with immigration. Thus $Y_{n+1} = mY_n + \delta_n$, $n=0, 1, 2, \dots$, where $E\delta_n = 0$, and δ_n is uncorrelated with Y_n , in analogy to the first-order autoregressive process. There is certainly a line-of-descent on estimation theory for m in branching processes with immigration, which is traced in Section 1 of Qi (2007). A related line of descent, which also continues, can be traced from Ku and Seneta (1998).

Chris's interest in population genetics models was reinvigorated by a note of the author (Seneta, 1974) which extended Malécot's Markov chain formulation $\{X_n\}$ of the Fisher-Wright diploid model for fluctuation of gene frequency, say of allele a out of the two possible alleles A, a , at a single diploid locus, to randomly fluctuating population size $\{N_n\}$, so that the proportion of A alleles in the population in generation n is given by $Y_n = X_n/(2N_n)$. The extension assumes the bivariate process $\{X_n, N_n\}$ is bivariate Markov, but retains the binomial sampling structure in the manner that

$$X_n \mid \{N_{n+1}, N_n, X_n\} \sim \mathcal{B}(2N_{n+1}, Y_n).$$

Since $\{Y_n\}$ is an (obviously bounded) martingale, by the martingale convergence theorem $Y_n \rightarrow Y$ a.s., with $0 \leq Y \leq 1$. The distribution of Y , unlike the fixed population-size model, is not necessarily concentrated at the two points 0, 1. A handle on its actual concentration is provided by the fact that $h_n = E(Y_n(1 - Y_n)) \rightarrow E(Y(1 - Y))$, equals h , say. For a deterministically varying population size, with $N_n > 1$, $n \geq 1$, we have $h > 0$ if and only if $\sum_{i=1}^{\infty} 1/N_i < \infty$. This result emanates from the fact that:

$$h_{n+1} = (1 - (2N_{n+1})^{-1})h_n. \quad (2)$$

The themes of almost sure convergence of relative frequency $\{Y_n\}$, and the concentration of probability of Y via the study of the evolution of $\{h_n\}$ by relations such as (2), using martingale methods, are the dominant themes of Chris’s subsequent work. Thus, in [M66] almost immediately, two theorems, focussed on when $h=0$ or not, are based on investigation of the zero-mean martingale with bounded increments $\{U_n, \mathcal{F}_n, n \geq 1\}$, where \mathcal{F}_n is the σ -field generated by the random vectors $\{(X_0, N_0), (X_1, N_1), \dots, (X_n, N_n)\}$, and

$$U_n = \sum_{k=1}^n [Y_k(1 - Y_k) - E(Y_k(1 - Y_k) | \mathcal{F}_{k-1})], n \geq 1. \tag{3}$$

The device underlying (3) for constructing a martingale is used repeatedly in later papers.

In [M71] the model is extended to permit selection via a parameter $\sigma \geq 0$, with $\sigma=0$ corresponding to the selective neutrality of [M66]. In this extension $\{Y_n, \mathcal{F}_n\}$ is a submartingale in general, so the martingale convergence theorem can still be used; and there is the pleasing conclusion that $\sigma > 0$ implies that $Y(1 - Y)=0$, a.s., just as in the case of fixed population size.

In [M87], Chris addresses such issues by generalizing Moran’s overlapping generation model, which is also a Markov chain. Here the transition time-points are times of death of single haploid individuals, of type a or A . The evolution of population structure is again studied through the movement of the (haploid) proportion $Z_n = X_n/N_n, n \geq 0$, which is again a bounded martingale, so $Z_n \rightarrow Z, 0 \leq Z \leq 1$. The quantity $Z_n(1 - Z_n)$ is now randomly normed to be a non-negative martingale relative to the σ -field \mathcal{F}_n , and martingale theory provides complete resolution of when $h = E(Z(1 - Z)) > 0$, inasmuch as a necessary and sufficient condition is that $1 < m < \infty$, where m is the mean of the offspring distribution. This insight, progressive assimilation and evolution of earlier techniques to resolve a question completely, is characteristic of Chris’s work.

A concluding paper [M98] to the sequence [M66], [M71], [M87], [M91], [M89], [M105], reviews difficulties in studying effects of varying population size focussed on *relative* frequencies such as $\{Z_n\}$. To remedy these, Chris characteristically uses a perceived unifying feature based on the growth of *actual* frequencies $\{X_n\}$:

$$E(X_{n+1} | \mathcal{F}_n) = X_n(1 + f_n), \text{ a.s.}, \tag{4}$$

where f_n is \mathcal{F}_n -measurable, and $1 + f_n \geq 0$, a.s. There is, for the first time, no attention paid to the limiting behavior of $E(Z_n(1 - Z_n))$ as a means of information on the limiting distribution of Z .

Curiously, in a slightly earlier paper in the same year and in the same journal, Buckley and Seneta (1983) generalized the selectively neutral context of Seneta (1974) to, in essence, conditional branching processes, with the focus then being on the behavior of the (diploid) sequence $\{Y_n(1 - Y_n)\}$ in terms of the conditional expectation:

$$E(Y_{n+1}(1 - Y_{n+1}) | \mathcal{F}_n) = (1 - W_{n+1})Y_n(1 - Y_n), \tag{5}$$

where

$$W_{n+1} = (V_n / (2N_n - 1))(N_n / N_{n+1})^2,$$

and V_n is a conditional variance of an offspring distribution. This somewhat resembles both (2) and (4). Once the expression (5) is established, the theorems giving sufficient conditions for each of $E(Y(1 - Y))=0$ or $\neq 0$ follow precisely as in [M66]. For example, the former holds if $\sum_{n=0}^{\infty} W_{n+1} = \infty$, a.s., a condition later mentioned by Donnelly (1986) as equivalent to his own necessary and sufficient condition for fixation. The very recent paper of Prince and Weber (2007) achieves an ultimate tidying up of such issues. This separate line of descent from [M66] further testifies to the enduring vitality of Chris’s ideas.

References

[1] Brown, B.M. (1971) Martingale central limit theorems. *Ann. Math. Statist.*, **42**,59–66.
 [2] Buckley, M.J. and Seneta, E. (1983) The genetic balance between varying population size and selective neutrality. *J. Mathematical Biology*, **17**, 217–222.

- [3] Donnelly, P. (1986) A genealogical approach to variable-population-size models in population genetics. *J. Appl. Prob.*, **23**, 283–296.
- [4] Harris, T.E. (1963) *The Theory of Branching Processes*. Springer, Berlin.
- [5] Ku, S. and Seneta, E. (1998) Practical estimation from the sum of AR(1) processes. *Commun. Statist. – Simula.*, **27**, 981–998.
- [6] Moran, P.A.P. (1962) *The Statistical Processes of Evolutionary Theory*. Clarendon Press, Oxford.
- [7] Prince, T. and Weber, N. (2007) Fixation in conditional branching process models in population genetics. *J. Appl. Prob.*, **44**, 1103–1110.
- [8] Qi, Y. (2007) On asymptotic normality of sequential estimators for branching processes with immigration. *Journal of Statistical Planning and Inference*, **137**, 2892–2902.
- [9] Seneta, E. (1968) On recent theorems concerning the supercritical Galton-Watson process. *Ann. Math. Statist.*, **39**, 2098–2102.
- [10] Seneta, E. (1970) On the supercritical Galton-Watson process with immigration. *Math. Biosciences*, **7**, 9–14.
- [11] Seneta, E. (1974) A note on the balance between random sampling and population size. (On the 30th anniversary of G. Malécot’s paper.) *Genetics*, **77**, 607–610.