A BRANCHING PROCESS VERSION OF THE BELL-ANDERSON CELL POPULATION MODEL

Ziad Taib

Department of Mathematics
Chalmers University of Technology
and the University of Göteborg
S-412 96 Göteborg, Sweden
e-mail: ziad@math.chalmers.se

Abstract. We describe a branching process version of the Bell-Anderson model and discuss the existence of a stable cell size distribution. Moreover, some quantities of interest for the biologists such as the α-curve and the β-curve are shown to have natural interpretations within the context of the model. We also show that some empirical facts about the α- and the β- curves as well as about the correlations between sizes of near relatives (siblings, mother-daughter, etc.) are direct consequences of the model.

Key words: Multitype branching process, stable size distribution, α-curve, β-curve

1 Introduction

In this article, we consider the asymptotic behaviour of a multitype branching process version of the Bell-Anderson cell population model. The type of questions we investigate are similar to those in (Lasota & MacKey, 1984) and (Sennerstam & Strömberg, 1995) just to mention two references, one where mathematical results are obtained and one where answers are arrived at using simulations. Since many of these questions concern the genealogy of a randomly sampled cell, it turns out that the branching process approach is a natural alternative to deterministic modeling as well as to simulation studies.

2 The model

We assume that every cell inherits, at birth, a size $r \in S$, where $S = [\frac{\epsilon}{2}, L)$ is equipped with its Borel algebra $\mathcal{E}$. The basic outcome space at the individual level is denoted by Ω (equipped with some σ-algebra $\mathcal{A}$) and a cell of (birth-) size r chooses an outcome $\omega \in \Omega$ using $\mathbb{P}_r$, the life law of cells of type r. After having completed its cell cycle type, λ, a cell divides into two equal daughter cells. $\lambda : S \times \Omega \to R_+$ is but one example of some important aspect of the life of a cell. Any other aspect can be handled similarly.

If a cell population is issue from one single founder cell o, we can label the set of all possible cells by $I = \{o\} \cup \bigcup_{n \geq 1} \{1, 2\}^n$, where $\{o\}$ is the zeroth generation consisting of the founder cell. The whole population is defined on $(S \times \Omega^I, \mathcal{E} \times \mathcal{A}^I)$. As explained in (Jagers, 1989), this together with the assumption that the fates of different cells are independent defines a unique probability measure P_r on the entire population process. The development of the individual cells is assumed to follow the following rules. At birth a cell has an initial size, r which increases with time according to some growth function g. Growth is thus completely deterministic. The size at age t is denoted by $m(r, t)$. m and G are related by the initial value problem $\frac{dm}{dt} = g(m)$, where $m(r, 0) = r$. We also introduce $T(x) = \int_0^x \frac{1}{g(x)} dy$. Then $m(r, t) = T^{-1}(T(r) + t)$ and $T(x) - T(r)$ can be interpreted as the time it takes to grow from size r to size x. Eventually a cell will split into two equal daughters. Notice that this means we disregard cell death. The age λ, at division is usually referred to as the cell cycle time, or the generation time. Its distribution is given through its hazard rate function $b(s), s \in S$.

The basic parameter of the process is the so-called reproduction kernel $\mu(r, ds \times dt)$ which gives the expected number of children with size in ds born to a mother of (birth-) size r in the age interval dt. Using this notation, it is not difficult to see that the reproduction kernel takes the form

$$\begin{aligned} \mu(r, ds \times dt) &= 2E_r \left[1(\lambda \in dt) 1 \left(\frac{m(r, \lambda)}{2} \in ds \right) \right] \\ &= \int_{\epsilon/2} 1(u \in dt) 1 \left(\frac{m(r, u)}{2} \in ds \right) b(m(r, u)) \\ &\quad e^{-\int_0^u b(m(r,v)) dv} du \end{aligned}$$

and making a change of $*$ variable $m(r, u) = x$, we see that

$$\mu(r, ds \times dt) = 21(T(2s) - T(r) \in dt) \frac{b(2s)}{g(2s)} e^{-\int_r^{2s} \frac{b(v)}{g(v)} dv}.$$

With $Q(x) = b(x)/g(x)$, this can be written as

$$\mu(r, ds \times dt) = 21(T(2s) - T(r) \in dt) Q^(2s) e^{-\int_r^{2s} Q(v) dv}$$

$$u = T(x) - T(r), \quad du = \frac{1}{g(x)dx}$$

3 The assumptions

Let $\hat{\mu}_\alpha(r, ds) = \int_0^\infty \bar{e}^{\alpha t}\mu(r, ds \times dt)$. We will call the real number $\rho(\hat{\mu}_\alpha)$ defined by

$$\rho(\hat{\mu}_\alpha)^{-1} = \sup\{\lambda \geq 0; \sum_{n \geq 0} \lambda^n \hat{\nu}_\alpha^n \quad \text{is } \sigma\text{-finite}$$

the Penon root of the kernel $\hat{\mu}_\alpha$. Here $\hat{\mu}_\alpha^n$ is defined by the recursion

$$\hat{\mu}_\alpha^n(r, ds) = \int_S \hat{\mu}^{n-1}(r, dv)\hat{\mu}_\alpha(v, ds).$$

By the kernel $\hat{\mu}_\alpha$ being σ-finite it is meant that there exists a strictly positive $\mathcal{E} \times \mathcal{E}$-measurable function f such that

$$\int_S f(r, s)\hat{\mu}_\alpha(r, ds) < \infty \quad \forall r \in S.$$

α will be chosen as to give the kernel $\hat{\mu}_\alpha$ the Penon root one, and for that choice the subscript α will be omitted. Notice that the Penon root may be infinity or zero.

When $\mathcal{E}$ is countably generated, we can use the following definition of a conservative kernel. A kernel $\hat{\mu}$ is said to be conservative iff there exists a σ-finite non-zero measure m on $(S, \mathcal{E})$ such that $m(A) > 0 \Rightarrow \sum_{n \geq 0} \hat{\mu}^n(r, A) = \infty$, $\forall A \in \mathcal{E}$ and $r \in S$.

When $\hat{\mu}$ is conservative, it is possible to show by abstract Perron-Frobenius theory that there exists a σ-finite eigen measure π on $(S, \mathcal{E})$ and a strictly positive a.e. finite measurable eigen-function h satisfying

$$h(r) = \int_S h(r)\hat{\mu}(r, ds) \quad \text{and}$$
$$\pi(ds) = \int_S \mu(r, ds)\pi(dr)$$

Both h and π are unique up to a multiplicative constant. When $\inf h > 0$, we can norm π into a probability measure.

Four our model, we have

$$h(r) = 2e^{\alpha T(r)} \int_{e^0}^{r} Q(v)dv \int_{r/2} e^{-\alpha T(2s)} \int_{e^{-0}}^{2s} Q(v)dv$$
$$Q(2s)h(s)ds$$

and

$$\pi(ds) = 2\int_0^{2s} e^{-\alpha(T(s)-T(r))} Q(2s) e^{-\int_r^{2s} Q(v)dv} \pi(dr)$$

Notice that

$$\hat{\mu}(r, ds) = 2Q(2s) e^{-\alpha(T(2s)-T(r))} e^{-\int_r^{2s} Q(v)dv} ds,$$

for $s > \frac{r}{2}$ and $r \in S$.

Once α has been fixed by the Perron root requirement, the existence of π will follow once we can show that the kernel $(\hat{\mu}(r, ds))$ is conservative.

We will make the following assumptions

- (i) g is continuous and strictly positive on $[\frac{\epsilon}{2}, 2L)$.
- (ii) $g(2x) \neq 2g(x)$ which excludes exponential growth for individual cells.
- (iii) $b(x0 > 0 \quad \epsilon \leq x > 2L$
- (iv) $\int_\epsilon^{2L} b(x)dx = \infty$

We will also assume that $L > \infty$.

Lemma 1. A sufficient condition for the conservativity of the kernel $\hat{\mu}(r, ds)$ is the existence of a σ-finite non-zero measure m on $(S, \mathcal{E})$ such that $\hat{\mu}(r, A) \geq m(A) \quad \forall r \in S$ and $A \in \mathcal{E}$.

Proof.

$$\begin{aligned} \hat{\mu}^n(r, A) &= \int_S \hat{\mu}(s, A) \hat{\mu}^{n-1}(r, ds) \\ &\geq m(A) \int_S \hat{\mu}^{n-1}(r, ds) \\ &= m(A) \end{aligned}$$

This means that $\hat{\mu}^n(r, A) \geq m(A) \quad \forall n \geq 1$. If $m(A) > 0$, we get

$$\sum_{n \geq 0} \hat{\mu}^n(r, A) = \infty \quad \forall r \in S.$$

□

It is therefore enough to exhibit a candidate which can play the role of the measure m above. It turns out that it is possible to take

$$m(A) = \int_{A\cap[\frac{1}{2},L)} 2Q(2s)e^{-\alpha T(2s)}e^{-\int_0^{2s} Q(v)dv}ds.$$

Since

$$\begin{aligned} m(S) &\geq \int_{\frac{L}{2}}^{L} 2Q(2s)e^{-\int_0^{2s} Q(v)dv}ds \\ &= e^{-\int_0^{L} Q(v)dv} > 0, \end{aligned}$$

m is a non-zero measure, and it is not difficult to verify that

$$\begin{aligned} \hat{\mu}(r,A) &= \int_{A\cap[\frac{r}{2},L]} 2Q(2s)e^{-\alpha(T(2s)-T(r))}e^{-\int_r^{2s} Q(v)dv}ds \\ &\geq m(A), \quad 0 \leq r \leq L \end{aligned}$$

The discussion above settles the problem of the existence of the eigen measure π which has the interpretation (when properly normed) of being the stable (birth-) size distribution.

To proceed further, we can either use convergence theorems for the composition of a branching population or the so-called stable pedigree measure (cf. Jagers & Nerman, 199?), which describes functionals of the family tree of a randomly sampled cell from a very old population.

The results we will be interested in concern

(1) The α-curve, i.e., the proportion of undivided cells older than some preassigned age.

(2) The β-curve, i.e., the tail of the distribution of the absolute value of the difference between the cell cycle times of two sibling cells.

(3) The correlation between the cell cycle times of a mother and daughter.

We will also show that α and β have the same asymptotic logarithmic slope and that the correlation referred to in (3) is necessarily negative.

4 The asymptotic composition

The easiest way of describing different aspects of the asymptotic composition of the population is to use the stable population measure $\tilde{P}$. Since this is a quite complicated object (cf. Definition 1 in Jagers & Nerman, 1996), it will not be reproduced here. Fortunately, we only need a few of its many coordinates. The stable population measure describes the lives of the near relatives of a randomly samples cell, ego chosen from a very old population. Under P ego's type is distributed according to π.

First, we derive an expression for the α-curve. Since $\alpha(t) =$ the fraction of cells still undivided at age t, it can be interpreted as the probability that ego is still undivided at age t. With T_o denoting ego's cell cycle time, we have

$$\begin{aligned}\alpha(t) = \tilde{P}(T_0 > t) &= \int_0^\infty P_r(T_0 > t)\pi(dr) \\ &= \int_0^\infty e^{-\int_0^t b(m(r,u))du}\pi(dr)\end{aligned}$$

Next, we derive an expression for the β-curve. $\beta(t)$ can be interpreted as the probability that ego's cell cycle time T_0 differ from her sister's T_0' by at least t times units, i.e.,

$$\begin{aligned}\beta(t) &= \tilde{P}(|T_0 - T_0'| > t) \\ &= \int_0^\infty P_r(|T_0 - T_0'| > t)\pi(dr) \\ &= \int_0^\infty \int \int_{|x-y|>t} f(r,y)f(r,x)dxdy\pi(dr)\end{aligned}$$

where $f(r,x) = b(m(r,x))e^{-\int_0^x b(m(r,v))dv}$. It is straightforward to see that

$$\begin{aligned}\beta(t) &= 2\int_0^\infty \int_0^\infty \int_{y+t} f(r,x)f(r,y)dxdy\pi(dr) \\ &= 2\int_0^\infty \int_0^\infty b(m(r,y))e^{-\int_0^y b(m(r,v))dv} \\ &\quad e^{-\int_0^{y+t} b(m(r,v))dv}dy\pi(dr).\end{aligned}$$

We now show that α and β have the same asymptotic logarithmic slopes. To see this, we consider $\frac{d}{dt}\ln\beta(t)$ and $\frac{d}{dt}\ln\alpha(t)$, as $t \to \infty$.

$$\begin{aligned}\frac{d}{dt}\ln\beta(t) &= \frac{\beta'(t)}{\beta(t)} \\ &= \frac{-\int_0^\infty \int_0^\infty b(m(r,y))e^{-\int_0^y b(m(r,v))dv} b(m(r,y+t))e^{-\int_0^{y+t} b(m(r,v))dv}dy\pi(dr)}{\int_0^\infty \int_0^\infty b(m(r,y))e^{-\int_0^y b(m(r,v))dv}e^{-\int_0^{y+t} b(m(r,v))dv}dy\pi(dr)}\end{aligned}$$

and this converges to $\lim_{t\to\infty} b(m(r,y+t)) = -b(M)$.

Compare this to

$$\begin{aligned}\frac{d}{dt}\ln\alpha(t) &= \frac{\alpha'(t)}{\alpha(t)} \\ &= \frac{-\int_0^\infty b(m(r,t))e^{-\int_0^t b(m(r,v))dv}\pi(dv)}{\int_0^\infty e^{-\int_0^t b(m(r,v))dv}\pi(dr)}\end{aligned}$$

As $t \to \infty$ this will have $-b(m(r,\infty)) = -b(M)$ as limit, i.e., α and β have the same asymptotic logarithmic slope.

Finally, we consider the mother-daughter cell cycle correlation. First, we derive the bivariate distribution for T_0 and T_1, the cell cycle times of ego and one of her daughters, under the stable population law.

$$\begin{aligned}&\tilde{P}(T_1 \in A_1, T_0 \in A_0) \\ &= \int P_r(T_0 \in A_0, T_1 \in A_1)T((dr)\end{aligned}$$

where

$$P_r(T_0 \in A_0, T_1 \in A_1) = P_r(T_1 \in A_1 | T_0 \in A_0)P_r(T_0 \in A_0)$$

Since

$$P_r(T_0 \in A_0) = \int_{A_0} b(m(r,u))e^{-\int_0^u b(m(r,v))dv}du$$

and

$$\begin{aligned}&P_r(T_1 \in A_1 | T_0 = t_0) = P_{m(r,t_0)}(T_1 \in A_1) \\ &= \int_{A_1} b\left(\frac{m(m(r_1 t_0), ix)}{2}\right)e^{-\int_0^4 b(m(\frac{r,t_0)}{2},2_v)dv}du\end{aligned}$$

we get

$$\begin{aligned}&\tilde{P}(T_0 \in A_0, T_1 \in A_1) \\ &= \int_{A_0}\int_{A_1}\int b(m(r,t_0))e^{-\int_0^{t_0} b(m(r,v))dv}b\Big(m\frac{(r,t_0)}{2},t_1\Big) \\ &e^{-\int_0^{t_1} b\left(\frac{m(r,t_0)}{2},v\right)dv}\pi(dr)dt_1dt_0 \\ &= \int_{A_0}\int_{A_1} g(t_0,t_1)dt_1dt_0\end{aligned}$$

where $g(t_0,t_1)$ can be interpreted as the bivariate density of (T_0,T_1). Using this and the marginal distributions which can be derived from it, we can now calculate the covariance

$$\tilde{E}[T_0,T_1] = \tilde{E}[T_0]\tilde{E}[T_1]$$

and retain its sign, which turns out to be negative.

5 References

1. Jagers, P. and Nerman, O. (1996). The asymptotic composition of multitype branching populations. To appear in Mar Yor, ed., séminaire de Probabilités, Springer Verlag. Lecture Notes in Mathematics.

2. Lasota, A. and MacKey, M. C. (1984). Globally asymptotic properties of proliferating cell populations. J. Math. Biol. **19**, 43-62.

3. Sennerstam, R. and Strömber, J. O. (1995). J. theor. Biol. **175**, 177-189.

4. Taib, Z. (1995). Branching processes and functional differential equations determining steady-size distributions in cell populations. J. Appl. Prob. **32** 1-10.