APPENDIX A

Multivariate Probability Generating Functions

In this section, we will collect some results which are referred to throughout the book. Suppose $X = (X_1, \ldots, X_n) \sim \left\{p_{i_1 i_2 \ldots i_n}\right\}_{i_1, i_2, \ldots, i_n \geq 0}$ is a finite vector of non-negative random variables, or a Z_+^n-valued rv.

Definition 6 (Definition of the multivariate pgf)**.** The pgf f_X of a Z_+^n-valued rv X is the function

$$f_X(s) = \mathrm{E}\left(s_1^{X_1} s_2^{X_2} \cdots s_n^{X_n}\right) = \sum_{i_1, i_2, \ldots, i_n \geq 0} p_{i_1 i_2 \cdots i_n} s_1^{i_1} s_2^{i_2} \cdots s_n^{i_n}, \tag{A.1}$$

well defined if $s = (s_1, s_2, \ldots, s_n) \in U_n \equiv [0, 1]^n$.

Theorem 30 (Multivariate pgf theorem)**.** *Suppose X is a Z_+^n-valued rv with pgf f_X. Let us denote by (N_i) the nontriviality condition for the ith coordinate of X, that is, $P[X_i \leq 1] < 1$.*

1. *f_X is non-negative and continuous with all derivatives. Under (N_i), it is increasing and convex as a function of s_i.*
2. *The marginal laws for subsets of X_i's can be obtained by setting respective arguments of the pgf equal to 1 [e.g., $f_X(s)_{|s_j=1,\ j\neq i} = f_{X_i}(s_i)$, etc.]; $f_X(e) = 1$, where $e = (1, \ldots, 1)$.*
3. *$\partial^{k_1+\cdots+k_n} f_X(0)/\partial s_1^{k_1} \cdots \partial s_n^{k_n} = k_1! \cdots k_n!\ p_{k_1 \cdots k_n}$.*
4. *The $(k_1, \ldots, k_n)$th mixed factorial moment of X, $\mu_{k_1,\ldots,k_n} = \mathrm{E}\left[\prod_{i=1}^n \prod_{j=0}^{k_i-1} (X_i - j)\right]$, is finite if and only if $\partial^{k_1+\cdots+k_n} f_X(e-)/\partial s_1^{k_1} \cdots \partial s_n^{k_n} = \lim_{s \uparrow e} \partial^{k_1+\cdots+k_n} f_X(s)/\partial s_1^{k_1} \cdots \partial s_n^{k_n}$ is finite. In such a case, $\mu_{k_1,\ldots,k_n} = \partial^{k_1+\cdots+k_n} f_X(e-)/\partial s_1^{k_1} \cdots \partial s_n^{k_n}$.*
5. *If X and Y are two independent Z_+^n-valued rv's, then $f_{X+Y}(s) = f_X(s) f_Y(s)$.*

6. *If Y is a Z_+^n-valued rv and $\left\{X_j^{(i)};\ i \geq 1\right\}$, $j = 1, \ldots, n$, are sequences of Z_+^m-valued rv's, then $V = \sum_{j=1}^{n} \sum_{i_j=1}^{Y_j} X_j^{(i_j)}$ is a Z_+^m-valued rv with pgf $f_V(s) = f_Y\left[f_{X_1^{(1)}}(s), \ldots, f_{X_n^{(1)}}(s)\right]$, $s \in U_m$.*
7. *Suppose $\{X_i;\ i \geq 1\}$ is a sequence of Z_+^n-valued rv's. The limit $\lim_{i\to\infty} f_{X_i}(s) = f_X(s)$ exists for each $s \in U^n$ if and only if the sequence $\{X_i;\ i \geq 1\}$ converges in distribution (i.e., when $\lim_{i\to\infty} \mathrm{P}[X_{i,1} = k_1, \ldots, X_{i,n} = k_n] = \mathrm{P}[X_1 = k_1, \ldots, X_n = k_n]$). Then, $f_X(s)$ is the pgf of the limit rv X.*

A further generalization to the denumerable infinite case is possible. Suppose that $X = (X_1, \ldots, X_n, \ldots) \sim \left\{\left\{p_{i_1 i_2 \cdots i_n}\right\}_{i_1, i_2, \ldots, i_n \geq 0}\right\}_{n \geq 1}$ is an infinite vector of non-negative random variables, with the σ-algebra generated by the finite-dimensional truncations of the sequence. Also, we may consider X a Z_+^∞-valued rv.

Definition 7 (Denumerable pgf definition)**.** The pgf f_X of a Z_+^∞-valued rv X is a function

$$f_X(s) = \mathrm{E}\left(s_1^{X_1} s_2^{X_2} \cdots s_n^{X_n} \cdots\right) = \sum_{i_1, i_2, \ldots, i_n \geq 0} p_{i_1 i_2 \cdots i_n} s_1^{i_1} s_2^{i_2} \cdots s_n^{i_n} \tag{A.2}$$

defined for

$$s \in \bigcup\nolimits_{n\geq 1} U_n \equiv \bigcup\nolimits_{n\geq 1} \{(s_1, s_2, \ldots, s_n, 1, 1, \ldots) : s_1, s_2, \ldots, s_n \in [0, 1]\} \tag{A.3}$$

(i.e., for arguments $s \in [0, 1]^\infty$ with only finite number of coordinates not equal to 1).

Properties 1 through 5 stated in the multivariate pgf theorem carry over to the finite-dimensional restrictions of the denumerable pgf. An important difference is that Property 6 does not necessarily hold for infinite n, as the resulting sum may be improper (if it is proper, then Property 6 holds). Also, the convergence property (Property 7) requires an additional continuity requirement:

Denumerable pgf convergence. Suppose $\{X_i, i \geq 1\}$ is a sequence of Z_+^∞-valued rv's. A necessary and sufficient condition for convergence in distribution of this sequence to a Z_+^∞-valued rv X is that $\lim_{i\to\infty} f_{X_i}(s) = f_X(s)$ exists for each $s \in \bigcup_{n\geq 1} U_n$, and that f_X is pointwise continuous for all sequences $\{s^{(i)}, i \geq 1\}$ with $s^{(i)} \in U_n$. Then, $f_X(s)$ is the pgf of the limit rv X.

APPENDIX B

Probability Distributions for the Bellman–Harris Process

B.1 Construction

We start with a rigorous construction of the probability space of the process, following Chapter 6 of Harris (1963). The elements of the probability space are family histories of the particles.

B.1.1 The families

Let $\mathcal{I}$ be the collection of elements ι, where ι is either 0 or a finite sequence of positive integers $i_1, i_2, \ldots, i_k$. The collection $\mathcal{I}$ is denumerably infinite. The elements ι are enumerated in a sequence $\iota_1, \iota_2, \ldots$, starting, for example, with $0, 1, 2, 11, 3, 21, 12, 111, \ldots$. The ancestor or founder is denoted by $< 0 >$, whereas $< i_1, i_2, \ldots, i_k >$ denotes the i_kth child of the i_{k-1}th child of $\ldots$, of the i_2th child of the i_1th child of the ancestor.

The family history ω is the sequence $\omega = (l, \nu; l_1, \nu_1; l_2, \nu_2; l_{11}, \nu_{11}; \ldots)$, where l_ι is a non-negative real and represents the length of life of ι, and ν_ι is a non-negative integer and represents the number of children of ι. The collection of all family histories is denoted by Ω. Family history is a redundant description of the particles pedigree in the sense that it enumerates even "nonexistent" children; for example, if $\nu_{ij} = 5$ (the jth child of the ith child of the ancestor has five children), then none of the pairs l_{ijk}, ν_{ijk} for $k > 5$ corresponds to any members of the pedigree.

For each $\omega \in \Omega$, we define a sequence $I_0(\omega), I_1(\omega), \ldots$, where I_k is a collection of objects $< \iota >$ called the kth generation. The 0th generation $I_0(\omega)$ is the ancestor $< 0 >$ and $I_1(\omega)$ is the set of all objects $< i >$ with $1 \leq i \leq \nu(\omega)$. The succeeding generations are defined inductively: $I_k(\omega)$ is the set of all objects $< i_1 i_2 \ldots i_k >$ such that $< i_1 i_2 \ldots i_{k-1} >$ belongs to $I_{k-1}(\omega)$ and $i_k \leq \nu_{i_1 i_2 \ldots i_{k-1}}(\omega)$. The set of

objects $\bigcup_{k=0}^{\infty} I_k(\omega)$ is called the family $I(\omega)$. In view of remarks in the preceding paragraph, more than one family history ω may, in general, correspond to the same family $I(\omega)$.

B.1.2 The number of objects at given time

If the object $< \iota >=< i_1 \cdots i_k >$ belongs to the family $I(\omega)$, it is born at the time $t' = l + l_{i_1} + \cdots + l_{i_1 i_2 \cdots i_{k-1}}$ and dies at the time $t'' = t' + l_{i_1 i_2 \cdots i_k}$; if $t \in [t', t'')$, then the age of the object at t is $t - t'$. Thus, if at time t we count the objects that are alive and have ages $\leq y$, then $< \iota >$ is counted if and only if the following conditions hold (with obvious modifications if $\iota = 0$)

$$\begin{aligned} & i_1 \leq \nu, i_2 \leq \nu_{i_1}, \ldots, i_k \leq \nu_{i_1 i_2 \ldots i_{k-1}}, \\ & \quad t - y \leq l + l_{i_1} + \cdots + l_{i_1 i_2 \ldots i_{k-1}} \leq t, \\ & \quad l + l_{i_1} + \cdots + l_{i_1 i_2 \ldots i_{k-1}} + l_{i_1 i_2 \ldots i_k} > t. \end{aligned} \tag{B.1}$$

The first line in conditions (B.1) means that $< \iota >$ belongs to the kth generation; the second line says that $< \iota >$ was born between $t - y$ and t; the third line says that $< \iota >$ dies after time t.

For each object, ι, let us define $Z_\iota(y, t, \omega)$ to be 1 if conditions (B.1) hold and to be 0 otherwise. Define

$$Z(y, t, \omega) = \sum_{\iota \in \mathcal{I}} Z_\iota(y, t, \omega)$$

and

$$Z(t, \omega) = Z(\infty, t, \omega) = \sum_{\iota \in \mathcal{I}} Z_\iota(\infty, t, \omega).$$

Thus, $Z_\iota(y, t, \omega)$ is 1 if $< \iota >$ is alive and of age $\leq y$ at t and 0 otherwise; $Z(y, t, \omega)$ is the total number of objects of age $\leq y$ at t; $Z(t, \omega)$ is the total number of objects at t. The possibility $Z(y, t, \omega) = \infty$ for some values of y, t, ω is admitted.

Let us note that if $Z(t_0, \omega_0) = 0$ for some t_0 and ω_0, then $Z(t, \omega_0) = 0$ for all $t > t_0$.

B.1.3 Probability measure

Definition 8. The probability measure P is built on the space Ω of family histories ω in the following way.

1. The random variables l_ι are iid with distribution

$$\mathrm{P}\{l_\iota \leq t\} = G(t),$$

where G is a right-continuous probability distribution function for which $G(0+) = 0$.

2. The ν_ι's are independent of each other and of the l's, and iid with the pgf

$$f(s) = \sum_{r=0}^{\infty} p_r s^r = \sum_{r=0}^{\infty} \mathrm{P}\{\nu_\iota = r\} s^r,$$

with the trivial cases excluded and $m \equiv f'(1-) < \infty$.

We denote the kth convolution of G with itself by G^{*k} ($G^{*1} = G$). Thus

$$G^{*k}(t) = \int_{0-}^{t+} G^{*(k-1)}(t-y)\, d_y G(y).$$

Because ω corresponds to a denumerable family of independent real-valued random variables, the basic theorem of Kolmogorov ensures that the above assumptions determine uniquely a countably additive probability measure P on the σ -algebra generated by the cylinder sets in Ω. From the definition of $Z(t, \omega)$, it is seen that Z is measurable in (t, ω), where the measurable (t, ω) sets are those generated by rectangles $A \times B$, A being a Borel t-set and B a measurable set in Ω. This conclusion is equivalent to a statement that the family of rv's $\{Z(t, \omega), t \geq 0\}$ is a stochastic process.

B.1.4 The embedded Galton–Watson process and extinction probability

Let $\zeta_k = \zeta_k(\omega)$ be the number of objects in the kth generation $I_k, k = 0, 1, \ldots$. It can be verified that the sequence of random variables $\{\zeta_k, k = 0, 1, \ldots\}$ is a Galton–Watson branching process with generating function $f(s)$ (usually called the embedded Galton–Watson process). The essence of the proof is to verify the property

$$\mathrm{E}(s^{\zeta_{k+1}} | \zeta_1, \zeta_2, \ldots, \zeta_k) = [f(s)]^{\zeta_k}, \tag{B.2}$$

which characterizes the Galton–Watson process. Equation (B.2) is a version of the forward pgf equation (3.5), conditional on ζ_k.

The embedded Galton–Watson process is helpful in proving that the probability of extinction for the Bellman–Harris process is subject to the same rules which govern the Markov versions. Let us note, for example, that if the embedded process becomes extinct for some ω, then the time-continuous process does too, as there is only a finite number of nonvoid generations $I_k(\omega)$ which may last for only a finite time. Thus, $\lim_{k\to\infty} \zeta_k(\omega) = 0$ implies $\lim_{t\to\infty} Z(t, \omega) = 0$. The opposite is, in general, not true. An example can be a situation when all the objects in the kth generation have life lengths $\leq 2^{-k}$ and, consequently, $Z(t) = 0,\ t > 2$. The following result demonstrates that such occurrences have probability 0.

Theorem 31. *Let A be the event $\{\zeta_n > 0$, for each $n\}$ and let B be the event $\{Z(t) > 0$, for each $t \geq 0\}$. If $\mathrm{P}\{A\} > 0$, then $\mathrm{P}\{B|A\} = 1$.*

Corollary 4. *The probability of extinction [i.e., of the event $\bar{B} \equiv \{Z(t) = 0$, for all sufficiently large $t\}$], is equal to the probability of the event $\bar{A}$ [i.e., to the smallest non-negative root q of the equation $s = f(s)$].*

B.2 Integral Equation

B.2.1 Decomposition into subfamilies

If the initial object dies at or before time t, then the objects present at t are its children or their descendants. For a family history $\omega = (l, \nu; l_1, \nu_1; l_2, \nu_2; l_{11}, \nu_{11}; \ldots)$ and each $i = 1, 2, \ldots$, let us define $\omega_i = (l_i, \nu_i; l_{i1}, \nu_{i1}; l_{i2}, \nu_{i2}; l_{i11}, \nu_{i11}; \ldots)$. The ω_i may be interpreted as the family history of $< i >$ and its descendants, although if $\nu < i$, then this family is not actually realized.

For the family history ω_i, let us define the random variables $Z_\iota(y, t, \omega_i)$, $Z(y, t, \omega_i)$, and $Z(t, \omega_i)$ in a way analogous to that in which, for ω, the rv's $Z_\iota(y, t, \omega)$, $Z(y, t, \omega)$, and $Z(t, \omega)$ were previously defined. Suppose that $l(\omega) \in [0, t]$ and $\nu(\omega) > 0$. It can be formally shown using the above definitions that

$$Z(t, \omega) = \sum_{i=1}^{\nu} Z(t - l, \omega_i). \tag{B.3}$$

In view of the fact that

$$I(\omega) = < 0 > \cup \bigcup_{i=1}^{\nu(\omega)} I(\omega_i),$$

the proof of Eq. (B.3) is reduced to careful "bookkeeping" of the indicator functions $Z_\iota(y, t - l, \omega_i)$ and $Z_{i\iota}(y, t, \omega)$.

B.2.2 Generating functions

Let

$$F(s, t) = \sum_{r=0}^{\infty} \mathrm{P}\{Z(t) = r\} s^r. \tag{B.4}$$

Because the case $Z(t) = \infty$ has not yet been eliminated, it can be $F(1, t) < 1$. However, also in this case, the basic properties of the pgf's are verified. Let us note the alternative expression

$$F(s, t) = \mathrm{E}[s^{Z(t)}] \equiv \int_\Omega s^{Z(t,\omega)}\, dP(\omega), \tag{B.5}$$

where $0^0 = 1$ and $s^\infty = 0$, even if $s = 1$.

Theorem 32. *The generating function F satisfies the integral equation*

$$F(s, t) = s[1 - G(t)] + \int_{0-}^{t+} f[F(s, t - u)]\, dG(u), \tag{B.6}$$

where $t \geq 0$ and $s \in [0, 1]$.

Proof. Based on Eq. (B.5), let us write

$$F(s, t) = \int_\Omega s^{Z(t,\omega)}\, dP(\omega) = \int_{\{l>t\}} s^Z\, dP + \sum_{k=0}^{\infty} \int_{\{l\leq t, \nu=k\}} s^Z\, dP. \tag{B.7}$$

Because $Z(t,\omega)=1$ if $l>t$, we have $\int_{\{l>t\}} s^Z\,dP = s\Pr\{l>t\} = s[1-G(t)]$.

Let us consider Ω as a product space $\Omega'\times\Omega_1\times\Omega_2\times\cdots$ of points $(l,v;\omega_1,\omega_2,\ldots)$. Let P' be the probability measure on the pair (l,v) and let P_i be the probability measure on Ω_i. Now, it is possible to use Eq. (B.3). If l is fixed, then $Z(t-l,\omega_i)$ is a function on Ω_i and hence, if k is any positive integer, we have

$$\int_{\{l\le t,v=k\}} s^Z\,dP = \int_{\{l\le t,v=k\}} dP'(l,v)\int_{\Omega_1} s^{Z(t-l,\omega_1)}\,dP_1\ldots\int_{\Omega_k} s^{Z(t-l,\omega_k)}\,dP_k.$$

Now, each of the integrals $\int_{\Omega_i} s^{Z(t-l,\omega_i)}\,dP_i$ is equal to $F(s,t-l)$, as the probability measure $dP_i(\omega_i)$ is exactly the same as $dP(\omega)$. Hence, the last equation is equal to $p_k\int_{0-}^{t+}[F(s,t-u)]^k\,dG(u)$. The same can be seen directly true if $k=0$. Substitution into the right-hand side of Eq. (B.7) yields the desired result.

B.2.3 Uniqueness of $F(s,t)$ and finiteness of $Z(t)$

Theorem 32 states that the pgf of $Z(t)$ satisfies Eq. (B.6), but it does not state that this solution is unique, nor that $\lim_{s\uparrow 1}F(s,t)=1$ [i.e., that $Z(t)<\infty$]. We will outline here the arguments proving both these properties.

Regarding uniqueness, let us assume that there exists another pgf solution $\tilde F(s,t)$ of Eq. (B.6). Then,

$$|F(s,t)-\tilde F(s,t)| \le \int_0^t |F(s,t-y)-\tilde F(s,t-y)|\,dG(y). \tag{B.8}$$

We see that because both F and $\tilde F$ are pgf's, $|F(s,t)-\tilde F(s,t)|\le 1$. Substituting into the right-hand side of Eq. (B.8), we obtain $|F-\tilde F|\le G(t)$. Substituting this and repeating the estimate, we obtain that $|F-\tilde F|\le G^{*i}(t)$ for any i. However, $\lim_{i\to\infty}G^{*i}(t)=0$ for any t (see Lemma 4), which yields $|F-\tilde F|=0$.

Finiteness of $Z(t,\omega)$ may be obtained by estimating another random variable $\bar Z(t,\omega)$ equal to the total number of objects in family $I(\omega)$ that are born up to and including time t (i.e., the counting function of births). Of course, $Z(t,\omega)\le\bar Z(t,\omega)$. We will consider the expected value of $\bar Z$. If it is finite, then $\bar Z$ is finite and so is Z [and, consequently, $F(1-,t)=1$].

For the argument, let us consider an object $<\iota>\neq<0>$, where $\iota=i_1i_2\cdots i_k$. Let u_ι be a random variable that is 1 if $<\iota>$ is in the family $I(\omega)$ (i.e., if it is ever born), and 0 otherwise, and let v_ι be a random variable that is 1 if $l+l_{i_1}+\cdots+l_{i_1i_2\cdots i_{k-1}}\le t$ and 0 otherwise. Then, $<\iota>$ is born at or before t if and only if $u_\iota v_\iota=1$, and

$$\bar Z(t)=1+\sum_{k=1}^{\infty}\sum_{i_1i_2\cdots i_k=1}^{\infty} u_{i_1i_2\cdots i_k}v_{i_1i_2\cdots i_k}.$$

The expected value $\mathrm{E}(v_\iota)$ is equal to $G^{*k}(t)$. The rv u_ι is the indicator function of the event that object $<\iota>$ is ever born and, therefore, its expectation is equal to the probability of this event; that is, to

$$\mathrm{E}(u_\iota)=\mathrm{P}\{u\ge i_1,u_{i_1}\ge i_2,\ldots,u_{i_1\cdots i_{k-1}}\ge i_k\}$$

$$= \mathrm{P}\{u \geq i_1\}\mathrm{P}\{u_{i_1} \geq i_2\} \ldots \mathrm{P}\{u_{i_1 \cdots i_{k-1}} \geq i_k\}.$$

The u_ι's and v_ι's are independent, so that

$$\begin{aligned}\mathrm{E}[\bar{Z}(t)] &= 1 + \sum_{k=1}^{\infty} G^{*k}(t) \sum_{i_1} \mathrm{P}\{u \geq i_1\} \sum_{i_2} \mathrm{P}\{u_{i_1} \geq i_2\} \ldots \sum_{i_k} \mathrm{P}\{u_{i_1 \ldots i_{k-1}} \geq i_k\} \\ &= 1 + \sum_{k=1}^{\infty} G^{*k}(t)[f'(1-)]^k.\end{aligned}$$

Lemma 4 states that this sum is $< \infty$ for all t and so $\mathrm{E}[\bar{Z}(t)] < \infty$.

APPENDIX C

General Processes

C.1 Introduction to the Jagers–Crump–Mode Process

This section is a useful reference, but it can be omitted at first reading. Its aim is to introduce the reader in an informal way to the basics of the general branching processes. In most part, the book is concerned with less general processes; therefore the subject can be postponed to a later reading. However, there are issues that are best expressed when phrased in terms of general processes. An example is an application of a general process to cell populations in Section C.2. Another recent example is an application of the general process in the genetics of aging (Olofsson et al. 2001). This latter work is also, to our knowledge, the only such application based on real-life data.

C.1.1 Definition of the general branching process

A basic source concerning general branching processes is the book by Jagers (1975). Our account is also based on Taïb (1992).

Individuals

We consider development in time of a population started by a single individual. The individuals can be considered elements of the set

$$I = \bigcup_{n=0}^{\infty} N^n,$$

called the Ulam–Harris space, where $N = \{1, 2, \ldots\}$ and $N^0 = \{0\}$. Individual 0 is the ancestor of the population. Each element of N^n is of the form $x = (x_1, \ldots, x_n)$. The meaning of this notation is that the individual belongs to the nth generation

and is the x_nth progeny of the x_{n-1}st progeny, ..., of the x_1st progeny of the ancestor. This description is redundant, as not all these individuals will come to existence in a given realization of the process. Each of the individuals evolves in a space Ω, which is large enough to allow for all possible life spans and progeny-bearing processes of this individual. An element $\omega \in \Omega$, is this individual's life. The probability measure on a σ-algebra $\mathcal{F}$ of Ω is called Q.

Lives

For each individual, $\tau(\omega, k),\ k = 1, 2, \ldots$, denotes successive ages at childbearing. In particular, $\tau(\omega, k)$ is the age at which the individual has its kth progeny. These ages are organized as epochs of a point process, a random collection of time moments or equivalently a random collection of non-negative integer-valued measures, denoted ξ. Mathematically,

$$\xi(\omega, [0, t]) = \xi(t) = \#\{k : \tau(\omega, k) \leq t\}$$

is the counting function of births (i.e., the number of progeny begotten before or at the age of t). In addition, λ, the duration of life ω of an individual, is a random variable $\lambda : \Omega \to R^+$.

The time evolution of the individuals is governed by the connections between their times of births. Let σ_x denote the moment of birth of individual x ($\sigma_0 = 0$, for the ancestor). Then, if we denote by xk the individual being the kth progeny of x, we set

$$\sigma_{xk} = \sigma_x + \tau_x(k).$$

In this latter expression, the argument ω is dropped, as it will be frequently done.

Construction of the process

If the space Ω is a Polish space (i.e., it is metric, complete, and separable), then the σ-algebra $\mathcal{F}$ can be selected as the class of Borel sets of Ω. The triplet (Ω, $\mathcal{F}$, Q) is the probability space of a single individual. If we assume that the lives of individuals are independent, then the space of the process can be constructed as a product space of the form (Ω^I, $\mathcal{F}^I$, Q^I), where I is the collection of all individuals. From now on, we will write P instead of Q^I and ω instead of $\{\omega_x, x \in I\}$.

The model presented can be specialized to include the classical branching processes, by assuming that all $\tau(\omega, k),\ k = 1, 2, \ldots$, are concentrated at $\lambda(\omega)$ (i.e., all progeny are born at the same time). Then, if $\lambda(\omega) = 1$, we obtain the Galton–Watson process. If $\lambda(\omega)$ is a non-negative rv, we obtain the Bellman–Harris process and so forth.

C.1.2 Random characteristics and basic decomposition

The method of random characteristics makes it possible to account for individuals existing in the process, individuals being born during a given time interval, individuals with ages from a given interval, individuals with a given number of progeny

and so forth. The random characteristic is a random function $\chi_x(a)$ defined on an individual's life. It defines the contribution, of a desired type, of individual x, from its birth until it reaches age a. The summary contribution of all individuals at a given time t, is equal to

$$Z_t^\chi = \sum_{x \in I} \chi_x(t - \sigma_x).$$

Z_t^χ is called the process counted by random characteristic $\chi_x(a)$. For example, if

$$\chi_x(a) = \begin{cases} 1 & \text{if } a \geq 0 \\ 0 & \text{otherwise,} \end{cases}$$

then Z_t^χ counts all individuals born until time t. If

$$\chi_x(a) = \begin{cases} 1 & \text{if } a \in [0, \lambda_x) \\ 0 & \text{otherwise,} \end{cases} \tag{C.1}$$

then Z_t^χ counts all individuals alive at time t. If

$$\chi_x(a) = \begin{cases} 1 & \text{if } a \in [0, \lambda_x) \cap [\tau_x(k), \infty) \\ 0 & \text{otherwise,} \end{cases}$$

then Z_t^χ counts all individuals alive at time t, with at least k progeny born before t.

For the process counted by random characteristics, it possible to write a backward decomposition, analogous to Eq. (1.1):

$$Z_t^\chi = \chi_0(t) + \sum_{i=1}^{X} Z_{t-\tau_0(i)}^{\chi\,(i)},$$

where X is the number of progeny effectively begotten by the ancestor and superscript (i) denotes the ith iid copy of the process.

C.1.3 Expectations, Malthusian parameter, and exponential growth

Reproductive measure is the expectation of the point process of progeny births,

$$\mu(A) = \mathrm{E}[\xi(\omega, A)].$$

It is characterized by the reproductive counting function $\mu(a) = \mu([0, a])$. The expectation of the process, $m_t = \mathrm{E}(Z_t^\chi)$ counted by characteristic $\chi(a)$ with expectation $g(a) = \mathrm{E}[\chi(a)]$, can be represented by the expression

$$m_t = \sum_{n=0}^{\infty} \int_0^t g(t-u)\, d\mu^{*n}(u) = \int_0^t g(t-u)\, d\nu(u),$$

where $\nu(u) = \sum_{n=0}^{\infty} \mu^{*n}(u)$. The nth convolution power of the reproductive measure, μ^{*n}, counts the expected number of progeny born to the nth-generation

individuals in the process. Then, each of μ^{*n} has to be convolved with the expectation of the random characteristic, to account for proper bookkeeping, and the result summed over all generations of the process. Under mild conditions (e.g., no concentration of births at age 0 and expected total progeny of an individual finite), this sum is finite. Expectation m_t satisfies a renewal-type integral equation:

$$m_t = \int_0^t m_{t-u}\, d\mu(u) + g(t). \tag{C.2}$$

A major role in the theory is played by the Malthusian parameter, which determines (if it exists) the asymptotic rate of growth of m_t. The Malthusian parameter is the real solution of the equation

$$\hat{\mu}(\alpha) \equiv \int_0^\infty e^{-\alpha u}\, d\mu(u) = 1.$$

This solution, if it exists, is unique. In what follows, we will limit ourselves to the supercritical case {i.e., when $\mu([0,\infty)) > 1$} [see the classification (1.5)]. In this case the Malthusian parameter exists and is positive. The renewal theorem demonstrates, in the same way as was explained in Section 5.2 for the Bellman–Harris process, that m_t behaves asymptotically like $e^{\alpha t}$;

$$e^{-\alpha t} m_t \longrightarrow \frac{\int_0^\infty g(u) e^{-\alpha u}\, du}{\underbrace{\int_0^\infty u e^{-\alpha u}\, d\mu(u)}_{\beta}} \equiv c(\chi) \quad \text{as } t \to \infty. \tag{C.3}$$

If we assume that all progeny are born at the same time τ in the life of the individual, so that $\mu(u) = mG(u)$, where m is the mean count of progeny and $G(\cdot)$ is the cumulative distribution of τ, and that this is exactly the moment of individual's death (i.e., that $\lambda = \tau$), we obtain the Bellman–Harris process of Chapter 5. If we wish to account for individuals alive at time t, then we use the random characteristics of the form $\chi_x(a) = 1$ if $a \in [0, \tau)$, and $\chi_x(a) = 0$, otherwise, as in Eq. (C.1). This means that $g(u) = 1 - G(u)$. Substituting into expression (C.3), we obtain the expression derived for the Bellman–Harris process [Eq. (5.13)].

Without getting into more detail, we state that in the supercritical case, the entire process counted by a random characteristic behaves very much the same way as its expectation. Indeed, there exists a random variable W, with $\mathrm{E}(W) = 1$, such that

$$Z_t^\chi e^{-\alpha t} \longrightarrow c(\chi) W$$

as $t \to \infty$, with probability 1.

C.1.4 Abstract type spaces and composition of the process

Let us suppose that each newborn individual is endowed, at birth, with a type selected from a measurable space $(\Gamma, \mathcal{G})$, where $\mathcal{G}$ is a σ-algebra of subsets of Γ. In

other words, there exist measurable mappings $\rho(j) : \Omega \to \Gamma$, which determine the types of newborn individuals. The point process ξ, which describes reproduction, is now defined by

$$\xi(A \times B) = \#\{i \in N; \rho(i) \in A, \tau(i) \in B\}.$$

Intuitively, $\xi(A \times B)$ is the number of progeny of an individual, born in time set B, with types in set A. The population of individuals can be defined on $(\Gamma \times \Omega^I)$, where Γ describes the type of the ancestor. The theorem of Ionesco–Tulcea enables one to construct a unique probability measure P_γ on $(\Gamma \times \Omega^I, \mathcal{G} \times \mathcal{A}^I)$ for the process with a type-γ ancestor. Similarly as before, a major role is played by the reproduction kernel $\mu(\gamma, A \times B) = \mathrm{E}_\gamma[\xi(A \times B)]$. For each real λ, we define

$$\mu_\lambda(\gamma,\ d\gamma' \times du) = e^{-\lambda u}\mu(\gamma, d\gamma' \times du)$$

and

$$\hat{\mu}_\lambda(\gamma,\ d\gamma') = \int_0^\infty \mu_\lambda(\gamma, d\gamma' \times du).$$

The Malthusian parameter α is selected so that the kernel $\hat{\mu}_\alpha(\gamma,\ d\gamma')$ has a Perron–Frobenius eigenvalue equal to 1 (assuming this latter exists). The Perron–Frobenius eigenvalue is the real eigenvalue strictly dominating absolute values of all remaining eigenvalues. If we set $\nu_\alpha(\gamma,\ d\gamma' \times du) = \sum_{n\geq 0} \mu_\alpha^n(\gamma,\ d\gamma' \times du)$, where $\mu_\alpha^n(\gamma,\ d\gamma' \times du)$ is the n-fold convolution of measure $\mu_\alpha(\gamma,\ d\gamma' \times du)$ with respect to elements $d\gamma' \times du$, we can write

$$\mathrm{E}_\gamma[e^{-\alpha t} Z_t^\chi] = \int_{\Gamma \times R_+} \mathrm{E}_\gamma[e^{-\alpha(t-u)}\chi(t-u)]\nu_\alpha(\gamma,\ d\gamma' \times du).$$

So, we see that $\mathrm{E}_\gamma[Z_t^\chi]$ is of the form $R * g(\gamma, t)$, where $R = \nu_\alpha$ and

$$g(\gamma, t) = \mathrm{E}_\gamma[e^{-\alpha t}\chi(t)].$$

Asymptotic behavior of the expectation of the process and of the process itself in the supercritical case ($\alpha > 0$) depends on the conservativeness of the kernel $\hat{\mu}_\alpha(\gamma, d\gamma')$. For countably generated $\mathcal{G}$, the kernel is conservative if its potential $\hat{\nu}_\alpha(\gamma, d\gamma') = \sum_{n\geq 0} \hat{\mu}_\alpha^n(\gamma, d\gamma')$ has the property that there exists a σ-finite measure m on $(\Gamma, \mathcal{G})$ such that

$$m(A) > 0 \implies \hat{\nu}_\alpha(\gamma, A) = \infty \tag{C.4}$$

for all $\gamma \in \Gamma$. This property is a generalization of positive regularity of matrices.

If the kernel $\hat{\mu}_\alpha$ is conservative, there exists an eigenfunction h satisfying

$$\begin{aligned} h(\gamma) &= \int_{R_+}\int_\Gamma e^{-\alpha u} h(\gamma')\mu(\gamma,\ d\gamma' \times du) \\ &= \int_\Gamma h(\gamma')\mu_\alpha(\gamma,\ d\gamma'). \end{aligned} \tag{C.5}$$

So, $e^{-\alpha u}[h(\gamma')/h(\gamma)]\mu(\gamma,\ d\gamma' \times du)$ has total mass on $\Gamma \times R_+$ equal to 1 and it is a probability measure. $h(\gamma)$ is the reproductive value of individuals of type

γ. It indicates the relative long-term contribution of individuals of this type to the population.

If the kernel $\hat{\mu}_\alpha$ is conservative, there also exists a probability measure π, which satisfies

$$\pi(d\gamma') = \int_\Gamma \hat{\mu}_\alpha(\gamma,\ d\gamma')\pi(d\gamma). \tag{C.6}$$

This equation can also be written in the following manner:

$$h(\gamma')\pi(d\gamma') = \int_\Gamma \frac{h(\gamma')}{h(\gamma)}\hat{\mu}_\alpha(\gamma,\ d\gamma')h(\gamma)\pi(d\gamma)$$

if $\inf h(\gamma) > 0$. We can then normalize the equation so that we obtain $\int_\Gamma h(\gamma)\pi(d\gamma) = 1$. The measure π, defined above, can be interpreted as a stable distribution of the types of the newborn. Consequently, an individual drawn at random from a very old population is of a random type decided by π, independently of the initial conditions.

Another interesting expression,

$$\beta = \int_\Gamma \int_\Gamma \int_{R_+} te^{-\alpha t}h(\gamma')\mu(\gamma,\ d\gamma' \times dt)\pi(d\gamma),$$

can be considered the expected age at reproduction.

Similarly as in the single-type case, in the supercritical case ($\alpha > 0$) a generalization of the key renewal theorem makes it possible to calculate the limit of $\mathrm{E}[e^{-\alpha t}Z_t^\chi]$. We will denote $\mathrm{E}_\pi(X) = \int_\Gamma \mathrm{E}[X]\pi(d\gamma)$, the expectation in the process with the type of ancestor being randomly drawn according to measure π. Then, we have

$$\mathrm{E}[e^{-\alpha t}Z_t^\chi] \longrightarrow \frac{\mathrm{E}[\hat{\chi}(\alpha)]}{\alpha\beta}h(\gamma)$$

as $t \to \infty$, for all γ except sets of π-measure 0. The process behaves in the supercritical case very much like its expectation (also, see Berndtsson and Jagers 1979).

The multitype formulation provides a great generality and was used in applications, particularly concerning evolution theory (Taïb 1992).

C.2 *Application*: Alexandersson's Cell Population Model Using a General Branching Process

An elegant example of modeling using general processes and counting characteristics (Section C.1) is a part of Alexandersson's (1999) thesis. This application demonstrates how a branching process approach complements existing deterministic approaches while the construction of the process is very straightforward.

C.2.1 The model

Let us consider a cell population, where each cell inherits a type at birth, grows during a stochastic time span, and when its cell cycle is completed, it divides into two not necessarily equal daughter cells. The type of the individual is the birth size of the cell expressed as mass, volume, DNA content, and so forth. Because cells have only two progeny, the Ulam–Harris space of all possible cells reduces to

$$I = \bigcup_{n=0}^{\infty} \{1, 2\}^n,$$

where $\{1, 2\}^0 = \{0\}$.

The type space is an interval $S = (0, M]$ of the real line, where $M < \infty$ is the largest possible birth size of a cell and $\mathcal{S}$ is the Borel-σ-algebra on S. A cell with birth size $r \in S$ chooses a life ω from $(\Omega, \mathcal{A})$ using $P(r, \cdot)$, the life law of cells of type r.

We construct the population space $(S \times \Omega^{\mathcal{I}},\ \mathcal{S} \times \mathcal{A}^{\mathcal{I}})$ as in Section C.1. Under the assumption that the daughter processes of different cells are conditionally independent, there exists a unique probability measure P_r on the entire population process, where $r \in S$ is the type of the ancestor.

The size of a cell with initial size r increases with time according to a deterministic growth function g. We let $m(r, t)$ denote the size of an r-type cell at age t. The functions m and g are related by the initial value problem

$$\frac{dm}{dt} = g(m), \qquad m(r, 0) = r.$$

The cell grows and, after division, the daughter cells do not necessarily have the same size (type) at birth. Note that we do not allow cell death in this model, so our branching population is supercritical. Let λ denote the age of the cell at division (the cell cycle time) and let the distribution of λ be defined by its hazard rate function $b(s)$, $s \in (0, 2M]$ {i.e., $\mathrm{P}[\lambda > s] = \exp[-\int_0^s b(u)\, du]$}.

A cell of type r divides into fractions δ and $1 - \delta$, where δ is a random variable on $(0, 1)$ with density function $f_\delta(m, p)$, $p_1 \le p \le p_2$, where $p_1 = 1 - p_2 \in (0, 1)$ depends on $m = m(r, \lambda)$, the cell size at division. We will assume that f_δ is unimodal and that δ is symmetrically distributed around $1/2$ {i.e., for all $r \in S$, $f_\delta(m, p) = f_\delta(m, 1 - p)$, and $\mathrm{E}_r[\delta] = 1/2$}.

Let $T(x) = \int_0^x [1/g(y)]\, dy$, $x \in S$. To see how to interpret this function, consider

$$T(x) - T(r) = \int_r^x \frac{1}{g(y)}\, dy. \tag{C.7}$$

Making a change of variable $y = m(r, t)$ yields $dy = dm(r, t) = g(m(r, t))\, dt$ and Eq. (C.7) becomes

$$\int_0^u \frac{g(m(r, t))}{g(m(r, t))}\, dt = \int_0^u dt = u,$$

where u is the time it takes for a cell to grow from size r to size x. Consequently, $T(x) - T(r)$ is precisely this time. Because $T(m(r, t)) - T(r) = t$, we have $m(r, t) =$

$T^{-1}(T(r)+t)$. Further, let $C(x)=\int_0^x [b(y)/g(y)]\,dy$ and $Q(x)=b(x)/[xg(x)]$, and assume that each cell has to divide before it reaches size $2M$; that is, b is such that for $\epsilon>0$,

$$\int_0^{2M} \frac{b(y)}{g(y)}\,dy=\infty \quad \text{and} \quad \int_0^{2M-\epsilon} \frac{b(y)}{g(y)}\,dy<\infty.$$

The reproduction kernel $\mu(r,\ ds\times dt)$, which is the expected number of children with birth sizes in ds to a cell of type r with age in dt, takes the form

$$\begin{aligned}\mu(r, ds\times dt) &= \mathrm{E}_r[\xi(ds\times dt)]\\ &= \mathrm{E}_r[\mathbf{1}(\lambda\in\ dt)(\mathbf{1}(\delta m(r,\lambda)\in ds)+\mathbf{1}((1-\delta)m(r,\lambda)\in ds))]\\ &= 2\int_0^\infty \mathbf{1}(u\in dt)\int_0^1 \mathbf{1}(pm(r,u)\in ds) f_\delta(m(r,u),p)\,dp\\ &\qquad \times b(m(r,u))\exp\left[-\int_0^u b(m(r,v))\,dv\right]du,\end{aligned}$$

where the factor 2 comes from the fact that δ and $(1-\delta)$ are identically distributed.

The inner integral is zero everywhere except when $p=s/m(r,u)$ and $dp=ds/m(r,u)$, so we have

$$\begin{aligned}\mu(r, ds\times dt) &= 2\int_0^\infty \mathbf{1}(u\in dt) f_\delta(m(r,u), s/m(r,u))\frac{b(m(r,u))}{m(r,u)}\\ &\qquad \times \exp\left[-\int_0^u b(m(r,v))\,dv\right]du\,ds. \qquad \text{(C.8)}\end{aligned}$$

Making a change of variable in the same manner as above, with $x=m(r,u)$, we get that $du=[dx/g(x)]$ and the kernel becomes

$$\mu(r, ds\times dt)=2\int_r^{2M} \mathbf{1}(T(x)-T(r)\in dt) f_\delta(x, s/x)Q(x)e^{-(C(x)-C(r))}\,dx\,ds.$$

C.2.2 Existence of the stable birth size distribution

If the Malthusian parameter α exists such that $\hat{\mu}_\alpha$ is conservative, then the Perron–Frobenius theorem gives the existence of a function h [see Eq. (C.5)] and a measure π [see Eq. (C.6)]. By requiring a strong or positive α-recurrence (Jagers and Nerman, 1996) and $\inf h>0$ we can norm to

$$\int_S h(s)\pi(ds)=1, \qquad \int_S \pi(ds)=1.$$

The measure π is then called the stable-birth-type-distribution. Hence, we want to prove the existence of the Malthusian parameter [i.e., prove the existence of a number $\alpha>0$ such that the Perron root $\rho(\hat{\mu}_\alpha)=1$], where

$$\hat{\mu}_\alpha(r, A)=\int_{\mathrm{R}_+} e^{-\alpha t}\mu(r, A\times dt)$$

and also that $\hat{\mu}$ is conservative.

Theorem 33. *Under the assumptions stated in Section C.2.1 on the reproduction kernel μ, the Malthusian parameter α exists and the kernel $\hat{\mu}_\alpha$ is conservative.*

C.2.3 Asymptotics of the cell model

We discuss the asymptotics of our cell model. When looking at a population, one can either consider all cells alive at the moment or all cells born into the population up until now. Even if it seems more natural to look at all cells alive, it is mathematically more convenient to consider all born. In this chapter, we will concentrate on all born cells, but we will also show that all the results presented can easily be obtained for all cells alive as well. When calculating the asymptotics of our model, we construct random characteristics used to count the population with respect to some property. An alternative way, described in Jagers and Nerman (1996) is to sample an individual at random in an already stabilized population and consider the population with time centered around this individual. The individual sampled at random is called ego.

The α curve is the graph of the function $\alpha(a)$ describing the proportion of cells still undivided at age a. An alternative interpretation is that $\alpha(a)$ is the probability that the age at division of a cell sampled at random, ego is larger than a. In order to find an expression for $\alpha(a)$, we define a random characteristic χ (cf. Section C.1) such that z_t^χ counts the number of cells born up to time t with respect to χ. Then, if y_t denotes the number of all cells born up to time t, we can use the result that under suitable conditions

$$\frac{z_t^\chi}{y_t} \to \mathrm{E}_\pi[\hat{\chi}(\alpha)] \quad \text{as } t \to \infty$$

in probability (on the set of nonextinction), where $\mathrm{E}_\pi[X] = \int_S \mathrm{E}_s[X]\pi(ds)$, $\hat{\chi}(\alpha) = \int_{\mathrm{R}_+} \alpha e^{-\alpha t}\chi(t)\,dt$, and π is the stable-birth-type distribution.

The random characteristic that gives score one for each cell x born up to time t and with life length λ_x longer than a can be written as

$$\chi_x(t) = \mathbf{1}_{\mathrm{R}_+}(t - \tau_x)\mathbf{1}(\lambda_x > a),$$

where τ_x is the birth time for cell x. Making a change of variable $u = t - \tau_x$ gives

$$\chi(u) = \mathbf{1}_{\mathrm{R}_+}(u)\mathbf{1}(\lambda > a).$$

This yields

$$\begin{aligned}
\alpha(a) = \mathrm{E}_\pi[\hat{\chi}(\alpha)] &= \int_S \mathrm{E}_r[\hat{\chi}(\alpha)]\pi(dr) \\
&= \int_S \mathrm{E}_r\left[\int_{\mathrm{R}_+} \alpha e^{-\alpha u}\chi(u)du\right]\pi(dr) \\
&= \int_S \int_{\mathrm{R}_+} \alpha e^{-\alpha u}\,du\;\mathrm{E}_r[\mathbf{1}(\lambda > a)]\pi(dr) \\
&= \int_S \mathrm{P}_r(\lambda > a)\pi(dr)
\end{aligned}$$

$$= \int_S \exp\left[- \int_0^a b(m(r, v))\, dv \right] \pi(ds).$$

The β curves are used to describe the proportions of sister cells, cousin cells, and so on with life lengths that differ by more than a time units. The β_1 curve describes this proportion for sister cells, β_2 for cousin cells, and so on. Alexandersson's (1999) thesis includes further asymptotic results for the β curves and numerical computations for the model we outlined. Furthermore, it also deals with a much more complicated example of cell proliferation, which we consider, using different methods, in Section 7.7.2.

APPENDIX D

Glossaries

D.1 Biological Glossary for Mathematicians

Cross-references to other glossary terms are *italicized*.

Amino acids The 20 different basic units of *proteins*.

Amplification (Gene Amplification) The increase in the number of copies of a *gene*. May result from errors in DNA *replication* or *recombination*.

Antibody A *protein* produced by the immune system in response to a foreign molecule (*antigen*) that interacts specifically with the foreign molecule.

Antigen A molecule that induces an *antibody*.

bp Base pair(s), usually used as a unit of length of a *DNA* strand, spanning one pair of complementary nucleotides.

Bacteria *Cells* of a lower form of life without a *nuclear* membrane.

Cancer A population of *cells* that continue to divide and survive under conditions in which normal cells would stop dividing or die. The cancer cell population usually is initiated from a single cell (clonal origin). As the progeny of the single cell multiply they accumulate *mutations* and acquire new characteristics (tumor progression). They may invade adjacent tissues and travel to distant sites to form secondary tumors (metastases).

Cell The basic unit of life. Cells of higher forms of life have an outer membrane surrounding the cytoplasm and the *nucleus*. In the cytoplasm there are *proteins* (enzymes) that carry out biochemical functions, machinery (ribosomes) for making proteins, and compartments (organelles) such as *mitochondria*. Higher forms of life, such as mammalian cells, which have a membrane surrounding their nucleus, are referred to as eukaryotes. Lower forms of life, such as *bacteria*, which do not have a membrane surrounding their nucleus, are referred to as prokaryotes.

Cell cycle The stages of *cell* growth and division. Includes the following stages (phases): division of one cell to produce two cells (cytokinesis), a gap of time (G_1

phase) between cytokinesis and the initiation of *DNA* synthesis (S phase), a gap of time (G_2 phase) between the end of DNA synthesis and the formation of visible chromosomes, and *mitosis* (M phase). In mitosis, the duplicated *chromosomes* (chromatids) containing replicated *DNA* are partitioned to new *cells* at cell division. The time between one cell division and another is referred to as the cell lifetime.

Centromere A part of the *chromosome* required for proper movement of the daughter *chromosomes* (chromatids) to daughter *cells*. A piece of *DNA* that is not part of a chromosome and does not contain a centromere DNA sequence is referred to as an acentric extrachromosomal element or double minute chromosome. Such acentric extrachromosomal elements do not segregate properly into daughter cells.

Chemotherapy The treatment of *cancer* cells with chemicals that kill them. In combination drug therapy, two or more chemicals with different modes of action are used to increase the efficiency of killing cancer cells.

Chromosome The linear structure containing *DNA* and *protein* that is visible under a microscope at *mitosis*. Chromosomes contain DNA sequences (*genes*) that code for proteins and DNA sequences that do not code for proteins. Among the noncoding DNA sequences, there are *centromeres* necessary for the separation of daughter chromosomes (chromatids) during *mitosis* and *telomeres*, which function to maintain the integrity of the ends of chromosomes.

Colony A population of *cells* that are the progeny of a single cell.

DNA Deoxyribonuleic acid; the genetic material. A long double helix with a structure similar to a twisted ladder. The backbones of the ladder are strands composed of alternating sugar (deoxyribose) and phosphate groups. The rungs of the ladder are pairs of nucleotide subunits. The nucleotide subunits are abbreviated A, T, G, and C. A is paired with T, and G is paired with C. The genetic information in DNA is stored in the sequence of nucleotides. The information is transcribed into complementary copies of a sequence of nucleotides in messenger *RNA* and is then translated into a sequence of *amino acids* in *protein*. During DNA replication, the two strands of a double helix separate and each acts as a template to synthesize a new complementary strand. Each of the two double helices (one new strand and one old strand) is contained in each one of a pair of sister chromatids (the daughters of *chromosomes*). The sister chromatids segregate into daughter cells at *mitosis*.

Drug resistance The continued survival of *cells* in the presence of chemicals (drugs) intended to kill them. Resistance to two or more drugs is referred to as double resistance or cross-resistance.

Eve The hypothetical common human female ancestor of all extant humans. Suggested by some common genetic features of individuals in current human populations.

Flow cytometry A method for the analysis of the distribution of the amount of a molecule (such as *DNA* or *protein*) in a population of *cells*. Cells are stained and pumped through a thin tube between a light source and a detector. Measurements of the amount of DNA per cell are used to indicate the number of cells in each phase of the cell cycle. Measurements of the amount of a specific protein per cell are used to indicate overproduction of the protein as a result of, for instance, *gene amplification*.

Fluctuation analysis Also, Luria and Delbrück fluctuation analysis. A method to determine *mutation* rates of *bacteria* or mammalian cells. Parallel cultures of cells are grown for a number of generations and then the number of mutants in each culture, the average number of mutants per culture, and the number of cultures containing no mutants are determined. This information can be used to calculate the number of mutations per cell per generation.

Gene A sequence of bases in *DNA* that codes for a *protein* and influences the inherited characteristics of a *cell* or organism.

Genome All of the *DNA* in an organism, including the DNA that codes for *proteins* and the DNA that does not code for proteins.

Heterogeneity (Tumor Heterogeneity) Populations of *cancer* cells that contain subpopulations with different characteristics, such as relative resistance to drugs.

Meiosis The formation of gametes (sex *cells*) by two successive cell divisions and only one round of *DNA* synthesis. This results in the segregation of nonidentical forms of genes (alleles) into different gametes. The gametes are haploid, containing half as much DNA as diploid body cells.

Mitochondria *Organelles* in the cytoplasm of *cells* of higher organisms needed for generating energy. Mitochondria contain *DNA*. They are inherited only from the mother, hence the term "maternal inheritance."

Mitosis The stage of the cell cycle of somatic (body) *cells* in which replicated *chromosomes* (chromatids) are separated into daughter cells. The result of mitosis is two daughter cells that have identical sets of *genes*. Daughter cells may be different in size as a result of asymmetric division of the cytoplasm at cell division.

Molecular clock hypothesis The assumption that *mutations* in a *gene* occur randomly and at an approximately equal rate over long time intervals during evolution.

Mutant An organism or *cell* that has a different inherited characteristic than the remainder of the cells in a population. Usually the result of a change in *DNA* sequence.

Mutation A change in *DNA* sequence. Usually detected by a sudden and inherited change in an observed characteristic (*phenotype*) of a *cell* or of an organism. However, a mutation may be detected directly by determining a change in the DNA sequence, even though there is no visible characteristic change in the cell or organism. The progeny of the mutant may revert to the previous phenotype, in which case the new mutation is referred to as a reverse mutation or back mutation. A phenotype resulting from a series of two mutations is referred to as a two-stage mutation. The rate of mutation may be determined by *fluctuation analysis*.

Nucleus The part of a *cell* containing *DNA*. The part of the cell outside of the nucleus is referred to as the cytoplasm.

Oncogene A *gene* (*DNA* sequence) associated with *cancer*. An oncogene can be detected and mapped by its pattern of inheritance in cancer-prone families. A piece of DNA containing an oncogene can be detected by the ability of the DNA to induce cancer-like changes when transferred into *cells* growing in culture.

Organelle A part of a *cell* which carries out a specialized function. An example is a mitochondrion (plural: *mitochondria*). A mitochondrion is a *DNA*-containing,

membrane-enclosed structure located in the cytoplasm. It functions to produce high-energy molecules for cell metabolism. During cell division, mitochondria may or may not be distributed to daughter cells in equal numbers.

Phenotype The visible characteristics of a *cell* or organism; as opposed to genotype, the genetic information of a cell.

Plasmid In *bacteria*, a circular piece of *DNA* that is separate from the major (chromosomal) piece of DNA. Plasmids replicate and segregate at *cell* division independently of the chromosomal DNA. Each bacterial cell may contain multiple numbers of plasmids which may be randomly distributed at cell division.

Polymerase chain reaction An experimental procedure for obtaining a large number of copies of a piece of *DNA*. The procedure employs short pieces of DNA complementary to the ends of the desired sequence and the enzyme DNA polymerase to exponentially increase the number of copies of the desired DNA sequence.

Protein A polymer molecule consisting of monomer subunits of *amino acids*. The linear sequence of amino acids in a protein is determined by the corresponding sequence of nucleotides in *DNA* (*gene*). Some proteins (enzymes) function to encourage chemical reactions; other proteins have a structural function.

Quiescence A phase when *cells* are pausing before the initiation of *DNA* synthesis rather than actively progressing through the *cell cycle*. Most cells of higher organisms are quiescent rather than actively dividing.

Recombination The formation of new combinations of *genes* by the exchange of genetic information between *chromosomes*.

Repeat DNA Sequences of *DNA* nucleotides that are tandemly iterated. In some diseases, the number of repeats may vary between individuals, and the number may change from parents to progeny.

Replication The duplication of *DNA*. Two strands of DNA separate, like a zipper, at a moving replication fork. Each strand acts as a template to code for a complementary sequence of nucleotides in a new strand. The result is two new pieces of DNA, each double stranded, each piece containing one new strand and one old strand. This is referred to as semiconservative replication. Errors may occur during DNA replication, slippage at the replication fork or redundant replication forks, resulting in sequences that are added or deleted (*amplification* or deamplification).

RNA Ribonucleic acid. A molecule similar to *DNA*, but with a different sugar (ribose rather than deoxyribose), one different nucleotide (U instead of T), and mostly single stranded (rather than double stranded). There are several kinds of RNA. One of these, messenger RNA (mRNA), is transcribed as a complementary copy of the sequence of nucleotides in DNA and functions to determine the sequence of *amino acids* in *protein*.

Segregation The separation of different forms of *genes* (alleles) into sex *cells* (gametes) at *meiosis*. Also, the distribution of double minute *chromosomes* to daughter cells during *mitosis*.

Senescence The inability of some normal *cell* populations to continue to divide indefinitely when grown in culture. Some *cancer* cell populations can continue

to divide indefinitely in culture and are therefore referred to as immortal. Senescence has been related to the continued activity of molecules that control *cell cycle* progression and to the maintenance of the length of *telomeres* at the ends of *chromosomes*.

Telomeres The ends of *chromosomes*. The *DNA* at the ends of chromosomes contains repeated sequences (terminal restriction fragments, TRF) that are necessary for replicating DNA at the ends of chromosomes and for maintaining the structural integrity of chromosomes.

Virus An intracellular parasite of *cells*. There are viruses of bacteria and of higher cells, including mammalian cells. They replicate within cells and can be transferred between cells. The extracellular forms contain genetic material (*DNA* or *RNA*), *proteins*, and some contain membranes. Within cells, the viral genetic material may subvert the machinery of the host cells and alter the host cell's properties. The genetic material of some viruses will actively replicate within a cell and produce new viruses. The genetic material of other viruses will integrate a DNA copy into the DNA of the host cell and replicate the viral genetic material along with the DNA of the host cell once per *cell cycle*.

D.2 Mathematical Glossary for Biologists

Cross-references to other glossary terms are *italicized*.

Abel's equation One of the classical functional equations of calculus. For a supercritical *branching process*, the characteristic function of the limit *random variable W* equal to the standardized particle count satisfies Abel's equation (3.19).

Age-dependent branching process A *branching process* in which the lifetimes of particles are non-negative *random variables*. In the special case when the lifetimes are exponentially distributed, the number of particles existing in the process, as a function of time, is a time-continuous *Markov chain*.

Asymptotic behavior Behavior of a time-dependent process (or a biological or physical phenomenon) after a sufficiently long time.

Backward approach Decomposition of the *branching process* into subprocesses started by direct progeny of the ancestor. By the branching property (a form of *self-recurrence*), these latter are distributed identically as the whole process. This decomposition provides the means to derive recurrent relationships or equations for the distributions of the process.

Bellman–Harris branching process A *branching process* in which the lifetimes of particles are non-negative *random variables* (age-dependent process) and the progeny is born exactly at the moment of the death of the parent.

Branching diffusion process A *branching process*, with a continuum *type space*, in which the type of the particle is defined as its position in a subset of real numbers (or points in higher-dimensional space) and the transitions in the type space are translations by a real-valued *random variable* (or a vector), with

special rules on the boundary. The type may be understood as a spatial coordinate of the particle.

Branching process A random collection of individuals (particles, objects, cells), proliferating according to rules involving various degrees of randomness of the life length and the number of progeny of an individual. The unifying principle is the so-called branching property, which states that the life length and type of progeny of a newborn particle, conditional on the current state of the process, are independent of any characteristics of other particles present at this time or in the future. The branching property is a form of *self-recurrence*.

Branching random walk A *branching process*, with a *denumerable*-type space, in which the type of the particle is defined as its position in the set of integers (or non-negative integers) and the transitions in the type space are translations by an integer *random variable*, with special rules on the boundary. An example is the process of *gene amplification* in proliferating *cells*. In this process, the type of cell is the number of copies of a *gene* present in the cell's *DNA*. Progeny cells may gain or lose copies of this gene, inherited from the parent cell. So, if the number of gene copies in the parent is equal to i; then in the progeny, it may be equal to $i-1$, i, or $i+1$.

Càdlàg path Function of time continuous from the right and bounded from the left of each point (French: **c**ontinue **à d**roite, **l**imitée **à g**auche).

Chapman–Kolmogorov equation Fundamental relationship governing the time evolution of *Markov chains*. It is represented in various forms {e.g., $P(s+t) = P(s)P(t)$ or $P_{ij}(s+t) = \sum_k P_{ik}(s)P_{kj}(t)$, where $P(s) = (P_{ij}(s))$ is the matrix (finite or infinite) of transition probabilities between states, $P_{ij}(s) = \mathrm{P}[X_{t+s} = j | X_t = i]$}. Intuitively, to calculate the probability of the chain moving from i to j in time $t+s$, it is necessary to add the probabilities of moving from i to k in time t and from k to j in time s, over all states k.

Criticality *Branching process* is critical if the expected (mean) count of progeny of a particle is equal to 1. It is supercritical if the mean count of progeny of a particle is greater than 1 and subcritical if it is less than 1. This classification leads to profound differences in *asymptotic properties* of the process. In particular, critical processes behave in a counterintuitive way because they become *extinct* with probability 1 while the expected number of particles stays constant.

Denumerable A set is called denumerable (or countable) if it is infinite but its elements can be indexed by non-negative integers. Other categories of infinite sets include continuum (i.e., a set the elements of which can be indexed by real numbers from an interval). The set of all rational numbers (ratios of integers) is countable; the set of all infinite sequences of zeros and ones is a continuum (because such sequences are just binary expansions of real numbers from the [0, 1] interval).

Exponential Steady State For idealized populations growing without spatial or nutritional constraints, the condition in which the number of individuals increases or decreases exponentially while the proportions of individuals in distinct age classes and any other identifiable categories remain constant. Usually attained *asymptotically*.

Extinction The event of all particles (individuals) of the *branching process* dying out.

Forward approach An approach dual to the backward approach, easiest to explain for the *Galton–Watson branching process*. Particles existing in generation t of the process are traced to their parents in generation $t-1$. Therefore, if the number Z_{t-1} of particles in generation $t-1$ is known, the number Z_t of particles in generation t is equal to the sum $Z_t = X_1 + X_2 + \cdots + X_{Z_{t-1}}$, where X_k is the number of progeny of the kth out of Z_{t-1} particles of generation $t-1$. This leads to a recurrence for the *pgf*'s of the particle counts.

Galton–Watson branching process Arguably, the simplest *branching process*. It evolves in discrete time measured by non-negative integers. At time 0, an ancestor individual (particle, cell, object) is born. At time 1, the ancestor dies, producing a random number of progeny. Each of these becomes an ancestor of an independent subprocess, distributed identically as the whole process. This definition implies that the numbers of progeny produced by each particle ever existing in the process are independent identically distributed *random variables* and that all particles live for one time unit. Discrete-time moments coincide with generations of particles. The number of particles existing in the Galton–Watson branching process, as a function of time, constitutes a time-discrete *Markov chain*.

Gelation In a model of aggregation of chemical molecules, the idealized process of infinite aggregation, resulting in disappearance of finite aggregates of molecules. In Macken and Perelson's branching model of aggregation, gelation is represented by escape of the *branching process* to infinity (possible only in the supercritical case).

Genealogies Branching (tree-like) graphs, usually random with respect to structure and branch lengths, representing ancestry of a sample of individuals from a *branching process* or, more generally, from an abstract or real-life population of molecules, *genes*, *cells*, or other objects. The process of reducing the number of distinct ancestors of the sample, followed in the reverse time, is called coalescence.

Genetic distance Distance between biological organisms, computed based on genetic characteristics. An example is the distance between relevant subsequences of *DNA* of the two individuals, computed as the number of nucleotides different in these two individuals (number of mismatches). For example, if in individual 1 the DNA sequence is ATGGACGA and in individual 2 it is ATcGgCGt, then the genetic distance is equal to 3.

iid Independent, identically distributed (*random variables*). The most frequently encountered assumption concerning a family of random variables. Makes proofs of theorems easier, when it can be assumed. In statistics, the so-called random samples are assumed to be iid.

Instability of branching processes The fact that, as time tends to infinity, the *branching process* either becomes *extinct* or infinitely large. Instability is due to the independence assumptions inherent in the definition of a branching process (i.e., that the number of progeny and life length of a newborn particle, conditional on the current state of the process are independent of any characteristics of other particles present at this time or in the future).

Jagers–Crump–Mode process The general *branching process*. The difference with respect to the classical branching processes, such as the *Galton–Watson branching process* or the *Bellman–Harris branching process* is that in the general process, the progeny may be produced before the death of the individual. The ages at which the individual begets progeny are random. Also, the *type space* may be of a very general form. The theory, developed for general processes, allows finding distributions of the process counted by random characteristics [i.e. of the weighted counts of events associated with a desired subclass of individuals (e.g., the number of first-born progeny of all individuals born after January 1, 1980, etc.)].

Kolmogorov theorem In the theory of *stochastic processes*, a fundamental result ensuring the existence of the stochastic process, given that for all finite collections of times, there exist joint distributions of *random variables*, being the values of the process at these times. These finite-dimensional distributions have to satisfy consistency conditions.

Linear-fractional case An important case of the *Galton–Watson branching process*, in which the number of progeny of an individual is a random variable with modified geometric distribution {i.e., $\mathrm{P}[X = 0] = 1 - bp/(1 - p)$ and $\mathrm{P}[X = k] = bp^k$, for $k = 1, 2, \ldots$. The name is derived from the fact that the *pgf* of such random variable is a ratio of two linear functions. In the linear-fractional case, the number of particles existing at any time has a modified geometric distribution, with parameters, which can be explicitly computed.

Malthusian parameter For a *branching process*, a parameter α such that the number $Z(t)$ of particles present in the process, normalized by dividing it by $\exp(\alpha t)$, converges to a limit *random variable*, as time tends to infinity. The Malthusian parameter always exists for the supercritical processes and is positive in this case.

Markov branching process A type of time-continuous *branching process*. At time 0, an ancestor individual (particle, cell, object) is born. The ancestor lives for time τ, which is an exponentially distributed *random variable*, and then the ancestor dies, producing a random number of progeny. Each of these becomes an ancestor of an independent subprocess, distributed identically as the whole process. The number of particles existing in the Markov branching process, as a function of time, is a time-discrete Markov chain (hence the name). Interestingly, if the Markov branching process is observed at times equal to multiples of a constant interval Δt, the numbers of particles at these observation times are distributed identically as in a *Galton–Watson branching process*.

Markov process A *stochastic process* with a limited memory (the Markov property). Intuitively, given the state of the process at time t, the future of the process depends only on this state and not on its states at times before t (time can be discrete or continuous). Mathematically,

$$\mathrm{P}[X_{t+s} \in A | X_s = x_s,\ 0 \le s \le t] = \mathrm{P}[X_{t+s} \in A | X_t = x_t],$$

where A is a subset of the state space of the process (space of values assumed by the process). The probability $P(s; x \to A) = \mathrm{P}[X_{t+s} \in A | X_t = x]$ is the transition

probability from state x to set of states A, in time s. If the states of the process form a finite or *denumerable* set, then the process is called a Markov chain. In this case, it is possible to define a matrix (finite or infinite) of transition probabilities between states $P(s) = (P_{ij}(s))$, where $P_{ij}(s) = \mathrm{P}[X_{t+s} = j | X_t = i]$. For discrete time, $P(s) = P(1)^s$, where $P(1)$ is the single-step transition probability matrix. For continuous time (under some additional assumptions if the number of states is infinite), $P(s) = \exp(Qs)$, where Q is called the transition intensity matrix.

Martingale In the discrete-time case, a *stochastic process*, having the property that its expected value at time $t+1$, conditional on its values at all times before $t+1$, is equal to the process value at time t. Mathematically, $\mathrm{E}(X_{t+1}|X_1, X_2, \ldots, X_t) = X_t$. Martingales, under some additional conditions, converge to limits (which are *random variables*). For this reason, proving that a process is a martingale allows an insight into its *asymptotic behavior*. Continuous-time martingales behave in a similar way, but they are technically more involved.

Maximum likelihood Statistical methodology of estimating parameters of models, based on observations. It consists of expressing the probability of observations as a function of parameters. This function is known as the likelihood function, $L(\theta) = f_X(x;\, \theta)$, where $f_X(\cdot)$ is the density of the distribution of *random variable* X, x is the vector of observations of random variable (known), and θ is the vector of parameters of the distribution (unknown). The values of parameters, which maximize $L(\theta)$, are called the maximum likelihood estimates of the parameters and are denoted $\hat{\theta}$.

Moments Expected values of powers of a *random variable* X. Absolute moments of order k (or kth absolute moments) are defined as $\mathrm{E}(X^k)$, central moments as $\mathrm{E}\{[X - \mathrm{E}(X)]^k\}$, and factorial moments as $\mathrm{E}[X(X-1)(X-2)\cdots(X-k)]$. The first absolute moment, $\mathrm{E}(X)$, represents the central tendency of the random variable, the second central moment, $\mathrm{Var}(X) = \mathrm{E}\{[X - \mathrm{E}(X)]^2\}$, represents the dispersion of the random variable around the expected value.

Multitype Galton–Watson process (positive regular) Generalization of the usual (single-type) *Galton–Watson branching process*. It evolves in discrete time measured by non-negative integers. Each individual belongs to one of a finite number of types. At time 0, an ancestor individual (particle, cell, object), of some type, is born. Processes started by individuals of different types are generally different. At time 1, the ancestor dies, producing a random number of progeny of various types. The distribution of progeny counts depends on the type of parent. Each of the first-generation progeny becomes an ancestor of an independent subprocess, distributed identically as the whole process (modulo ancestor's type). In the multitype process, asymptotic behavior depends on the matrix of expected progeny count. Rows of this matrix correspond to the parent types and columns correspond to the progeny types. The largest positive eigenvalue of this matrix (the *Perron–Frobenius* eigenvalue) is the *Malthusian parameter* of the process, provided the process is supercritical (the Perron–Frobenius eigenvalue larger than 1) and positive regular. This latter means that parent of any given type will have among its (not necessarily direct) descendents, individuals of all possible types, with nonzero probability.

Parsimony method in phylogenetics A method of inferring the *phylogenetic tree*. In this method, taxonomic units are represented by their *DNA* sequences (most commonly, from the *mitochondrial genome*). The method looks for the tree that requires the minimum number of changes between the extant and inferred ancestral sequences. The outcome may be equivocal and, also, because the number of possible tree structures is extremely large, the optimal tree is frequently not found.

Perron–Frobenius theory Collection of results concerning eigenvalues and eigenvectors of positive (or non-negative) matrices and operators. Important assumptions include irreducibility (positive regularity) (i.e., a strict positivity of iterates of the matrix or operator). A generic result states the existence of a strictly positive simple eigenvalue dominating all other eigenvalues and of a corresponding strictly positive eigenvector. The importance of these results is that they lead to characterizations of the *asymptotic behavior* of iterates of positive matrices or operators, in terms of dominant eigenvalues and eigenvectors. Mathematically, $m_0 M^i \sim \lambda^i \nu$, as $i \to \infty$, where M^i is the ith iterate of the positive matrix M, m_0 is the initial vector of states, λ is the dominant positive eigenvalue, and ν is the corresponding eigenvector. Results of this type are important in mathematical population dynamics, including the theory of *branching processes*.

pgf *Probability generating function*.

Phylogenetic tree The set of ancestry relationships between extant (contemporary) taxonomic or demographic units (species, populations, haplotypes, and others), usually in the form of a binary tree graph (at most three branches out of each node). The nodes of the phylogenetic tree represent extant and ancestral units, whereas the branches represent the intervals of evolutionary time separating them. Depending on the method of reconstruction, the graph may be rooted [i.e. having a uniquely defined common ancestor (and consequently, the direction of time specified in all branches), or unrooted (it is then sometimes called a network)]. The most commonly used methods of reconstruction are *parsimony*, distance matrix, and *maximum likelihood*.

Poisson process One of the most important *stochastic processes*. Random collection of time points (epochs) having the properties of complete randomness (the counts of events in any two disjoint time intervals are independent) and stationarity [the probability of an event occurring in a short time interval $(t, t + dt)$ is equal to $\lambda dt + o(dt)$, where, $o(dt)$ is small with respect to dt, i.e., $o(dt)/dt \to 0$ as $dt \to 0$]. The constant λ is called the intensity of the process. The number N of epochs of the Poisson process in an interval of length t has Poisson distribution with parameter λt [i.e., $\mathrm{P}[N = n] = \exp(-\lambda t)(\lambda t)^n/n!$, for $n = 0, 1, 2, \ldots$], and the time intervals T between any two epochs have exponential distribution with parameter λ (i.e., the density of distribution of T is equal to $f_T(t) = \lambda \exp(-\lambda t)$, for $t \geq 0$).

Population genetic models Models of inheritance, *mutation*, and selection of genetic material in populations of individuals. Classically, these models assume a constant number of individuals related to each other through common ancestry (Fisher–Wright model). Although very different from the *branching processes*

some of these models can be approximated by branching processes (e.g., when an expanding subpopulation of mutants arises within the large population). Such a situation arises when some genetic diseases are studied.

Positivity In general, the property of being positive. A matrix is positive if all elements of the matrix are positive; it is positive regular if all elements are non-negative and some power of the matrix is positive. If the matrix is a transition probability matrix of a *Markov process*, positive regularity means that there exist paths between all pairs of states of the process. Similarly, if the matrix is the mean progeny matrix of a multitype *branching process*, then positive regularity means that any particle has, among its descendants, particles of all types.

Probability generating function (*pgf*) The function $f_X(s)$ of a symbolic argument s, which is an equivalent of the distribution of a non-negative-integer-valued *random variable* X. If numbers $p_0, p_1, p_2, \ldots$ constitute the distribution of random variable X (i.e., $\mathrm{P}[X = k] = p_k$), then the pgf of random variable X is defined as $f_X(s) = \mathrm{E}(s^X) = \sum_{i=0}^{\infty} p_i s^i$, for $s \in [0, 1]$. Use of the pgf simplifies mathematical derivations involving non-negative integer-valued random variables.

Quasistationarity State i_a of a *Markov chain* $X(t)$ is called absorbing if the process cannot exit i_a once i_a has been visited (i.e., $\mathrm{P}[X(t+s) \neq i_a | X(t) = i_a] = 0$). Under certain additional conditions, the probability of eventual absorption in state i_a is equal to 1 (i.e., $\mathrm{P}[\lim_{t\to\infty} X(t) = i_a] = 1$). Then, the only *stationary* distribution is the one that assigns probability 1 to state i_a. Because such a distribution is not informative, it is usual to consider a distribution, which is stationary conditional on non-absorption. Such a distribution, if it exists, is called the quasistationary distribution. Mathematically, $\tilde{\pi} = (\tilde{\pi}_0, \tilde{\pi}_1, \tilde{\pi}_2, \ldots)$ is the quasistationary distribution, if $\mathrm{P}[X(t+s) = j | X(t+s) \neq i_a] = \tilde{\pi}_j$ (all j) provided $\mathrm{P}[X(t) = j | X(t) \neq i_a] = \tilde{\pi}_j$ (all j). An example of a quasistationary distribution is the limit distribution of the subcritical *branching process* conditional on nonextinction.

Random variable (*rv*) Intuitively, a numerical result of observation which displays random variation. Mathematically, a random variable $X(\omega)$ is a function mapping the elements ω of a probability space Ω (space of outcomes of a random experiment) into the set of real numbers. For technical reasons, this function has to be measurable (i.e., the counter image of an interval through X has to be a measurable set of elements of Ω).

Random walk A time-discrete *Markov chain* $X(t)$, such that $X(t+1) = X(t) + U(t)$, where the integer *random variables* $U(t)$ are independent and identically distributed.

Recurrent state See *transient state*.

Renewal theory A branch of probability concerned with renewal processes. The renewal process is a collection of random time points (called renewals) such that the intervals between these points are independent identically distributed *random variables*. A special case in which the intervals between renewals are exponentially distributed is the *Poisson process*.

rv *Random variable*

Self-recurrence Consider a random (*stochastic) process* $X(t)$ evolving from an initial value $X(0) = x_0$ on time interval $[0, \infty)$. Suppose that at some time

t_0, the process is stopped and then restarted. Then, suppose that given the value $X(t_0) = x_0$, the continuation process on the interval $[t_0, \infty)$, which is a subprocess of the original process, is identical (it has the same distributions), as the original process shifted by t_0. A process with such property is called self-recurrent. Self-recurrence may be considered a rephrasing of a causality principle. It leads to recurrent relationships for a wide class of processes, including *Markov processes, renewal processes*, and *branching processes*.

Stathmokinesis An experimental technique in which *cell* divisions are blocked, ideally without damage to cells. Cells traversing successive phases of their lives are accumulating in the predivision state (mitosis). The time pattern of accumulation depends on the demography of the cell population and kinetic parameters of the cell cycle. Therefore, it is possible to estimate some of these parameters based on observed accumulation patterns.

Stationarity The *Markov chain* $X(t)$ is said to be stationary if its distribution over the state space is invariant in time (this distribution is called the stationary distribution). Mathematically, $\pi = (\pi_0, \pi_1, \pi_2, \ldots)$ is the stationary distribution if $\mathrm{P}[X(t+s) = j] = \pi_j$ (all j) provided $\mathrm{P}[X(t) = j] = \pi_j$ (all j).

Stochastic process Intuitively, a function of time with a random component. Mathematically, a family of *random variables* parameterized by time. It has to satisfy so-called measurability conditions, which prevent certain mathematical problems from occurring.

Transient state States of a *Markov chain* can be classified into transient and *recurrent*. For a recurrent state, the probability of eventually returning to this state is equal to 1, whereas for a transient state, there is a nonzero probability of never returning.

Type space A collection of possible particle types existing in a *branching process*. If there is more than one but finitely many types, the process is called multitype. If the type space is *denumerable* or continuous, the behavior of the branching process can differ considerably from the multitype case. An example is a *branching random walk*, in which the asymptotic behavior can be, for example, exponential multiplied by a fractional power function, which does not occur in the finite case.

wp *With probability* (common abbreviation)

Yaglom's theorem Result stating that for subcritical *branching processes*, there exists a *quasistationary* distribution, conditional on non*extinction*.

Yule process *Markov age-dependent branching process* in which a particle can have at most two progeny (the binary-fission process). An important class of processes because the *pgf* of the distribution of particle count can be explicitly found. Also, the Yule process frequently serves as a model for populations of proliferating *cells*, although by its definition it is limited to exponentially distributed cell lifetimes.

Index