Computational Problems Related to the Galton-Watson Process

by

Marcel F. Neuts (*)
Purdue University

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series No. 402

February 1975

(*) Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR-72-2350B. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

Abstract

An algorithm to compute the probability distributions of the successive generation sizes in a Galton-Watson process is presented. The distribution of the number of offspring of each individual is assumed to be of phase type. A probability distribution is of phase type if it can be identified as the distribution of the time till absorption in an absorbing finite Markov chain with appropriate initial conditions. A detailed analysis of the error due to truncation is given, as well as an application in a problem related to the M/G/1 queue.

A second algorithm deals with the distribution of the maximum generation size before extinction. Several theorems on probability distributions of phase type are proved.

1. The Probability Distributions of Phase Type

In computational probability problems it is frequently desirable to have available a versatile class of distributions which may be concisely represented, which satisfy simple recurrence relations, in addition to having algorithmically useful properties under the operations of convolution and mixing. The probability distributions of phase type, which were discussed in [8], have many such properties in addition to the ones proved here. In the present paper, the computational advantages of these distributions in the study of the Galton-Watson process will be examined.

Let $\underline{\alpha}$ denote an m-vector of probabilities and $\underline{e}$ an m-vector with $e_i=1$, for $i=1,\ldots,m$. We define α_{m+1} by $\alpha_{m+1}=1-\underline{\alpha}\,\underline{e}$, and assume that $\alpha_{m+1} \geq 0$. The matrix T of order m is substochastic and such that I-T is nonsingular. The vector $\underline{T}^o$ is defined by $\underline{T}^o = \underline{e}-T\underline{e}$. The matrix T^o of order m has m identical columns given by $\underline{T}^o$. The matrix A is $\mathrm{diag}(\alpha_1,\alpha_2,\ldots,\alpha_m)$.

A probability density $\{p_k\}$ on the nonnegative integers is <u>of phase type</u> and has representation $(\underline{\alpha},T)$ if and only if

(1) $$p_0 = \alpha_{m+1},$$
$$p_k = \underline{\alpha}\, T^{k-1}\, \underline{T}^o, \qquad \text{for } k \geq 1.$$

It is shown in [8] that all probability densities on a finite number of nonnegative integers and all generalized negative binomial densities are of phase type. It is further shown that all right shifts, all finite mixtures and all finite convolution products of distributions of phase type are themselves of phase type and appropriate representations for these are constructed. It is easy to see that $\{p_k\}$ is the probability density of the time till absorption in the (m+1)-state Markov chain with transition probability matrix

$$(2) \qquad P = \begin{pmatrix} T & \underline{T}^o \\ \underline{0} & 1 \end{pmatrix},$$

and initial probability vector $(\underline{\alpha}, \alpha_{m+1})$. The condition that I-T is nonsingular guarantees that absorption into the state m+1 is certain for any initial probability vector. For purposes of representation of the density $\{p_k\}$ we may further assume without loss of generality that the stochastic matrix $Q = T + (1-\alpha_{m+1})^{-1}T^oA$, is irreducible.

There is a completely parallel development for probability distributions of phase type on the nonnegative real line. In this case we consider an (m+1)-state Markov chain in continuous time with infinitesimal generator

$$(3) \qquad B = \begin{pmatrix} T & \underline{T}^o \\ \underline{0} & 0 \end{pmatrix}.$$

The matrix T now has negative diagonal elements and nonnegative off-diagonal elements. The matrix T is nonsingular and $T\,\underline{e} + \underline{T}^o = \underline{0}$. The probability distribution of the time till absorption into the state m+1 with initial probability vector $(\underline{\alpha}, \alpha_{m+1})$ has a jump of α_{m+1} at zero and a density component given by

$$(4) \qquad \varphi(u) = \underline{\alpha} \exp(u\,T)\underline{T}^o, \qquad \text{for } u > 0.$$

Phase distributions on the positive real line play only an incidental rôle in this paper and except where indicated, we shall consider only the discrete case.

Countable mixtures of probability distributions of phase type are generally not of phase type, but the following positive result is of some interest in the Galton-Watson process.

Theorem 1

Let $\{p_k\}$ be a density of phase type with representation $(\underline{\alpha},T)$. Let $\{s_k\}$ be of phase type with representation $(\underline{\beta},S)$, where the order n of the matrix S may be different from that of the matrix T. If P(z) denotes the probability generating function of the density $\{p_k\}$, then the mixture $\{v_k\}$ of the successive convolutions of $\{p_k\}$, whose probability generating function V(z) is given by

$$(5) \qquad V(z) = s_0 + \sum_{\nu=1}^{\infty} s_\nu P^\nu(z),$$

is itself of phase type. A representation $(\underline{\gamma},K)$ with K of order mn, of the density $\{v_k\}$ is given by

$$(6) \qquad K = T \otimes I_n + T^o A \otimes (I_n - \alpha_{m+1}S)^{-1}S,$$

where $\otimes$ denotes the Kronecker product of matrices, and by the mn-vector $\underline{\gamma}$, given by

$$(7) \qquad \underline{\gamma} = \underline{\alpha} \otimes \underline{\beta}(I_n - \alpha_{m+1}S)^{-1}.$$

The corresponding vector $\underline{K}^o$ is given by

$$(8) \qquad \underline{K}^o = \underline{T}^o \otimes (I_n - \alpha_{m+1}S)^{-1}\underline{S}^o,$$

and γ_{mn+1} is given by

$$(9) \qquad \gamma_{mn+1} = \beta_{n+1} + \alpha_{m+1}\, \underline{\beta}(I_n - \alpha_{m+1}S)^{-1}\underline{S}^o.$$

Proof

The probability generating functions P(z) and S(z) of $\{p_k\}$ and $\{s_k\}$ are given respectively by

$$(10) \qquad P(z) = \alpha_{m+1} + z\, \underline{\alpha}(I_m - z\, T)^{-1}\underline{T}^o,$$

$$S(z) = \beta_{n+1} + z\, \underline{\beta}(I_n - z\, S)^{-1}\underline{S}^o,$$

and hence

$$(11)\qquad V(z) = \beta_{n+1} + \sum_{\nu=1}^{\infty} [\alpha_{m+1} + z\,\underline{\alpha}(I_m - z\,T)^{-1}\underline{T}^o]^\nu \underline{\beta}\, S^{\nu-1}\underline{S}^o$$

$$= \beta_{n+1} + \alpha_{m+1}\,\underline{\beta}(I_n - \alpha_{m+1}S)^{-1}\underline{S}^o + \sum_{\nu=0}^{\infty}\sum_{j=0}^{\nu} \binom{\nu+1}{j}\alpha_{m+1}^j\,[z\,\underline{\alpha}(I_m - z\,T)^{-1}\underline{T}^o]^{\nu-j+1}\,\underline{\beta}\,S^\nu\underline{S}^o.$$

The first two terms yield γ_{mn+1}. It therefore suffices to show that

$$(12)\qquad z\,\underline{\gamma}(I_{mn} - z\,K)^{-1}\underline{K}^o = \sum_{\nu=0}^{\infty}\sum_{j=0}^{\nu}\binom{\nu+1}{j}\alpha_{m+1}^j[z\,\underline{\alpha}(I_m - z\,T)^{-1}\underline{T}^o]^{\nu-j+1}\underline{\beta}\,S^\nu S^o.$$

Observing that

$$(13)\qquad [z\,\underline{\alpha}(I_m - z\,T)^{-1}\underline{T}^o]^{t+1} = z\,\underline{\alpha}\,[z(I_m - z\,T)^{-1}T^oA]^t(I_m - z\,T)^{-1}\,\underline{T}^o,$$

for $t \geq 0$, and interchanging the order of summation, we write the right hand sum as

$$(14)\qquad \sum_{j=0}^{\infty}\sum_{t=0}^{\infty}\binom{j+t+1}{j}\alpha_{m+1}^j\; z\,\underline{\alpha}[z(I_m - z\,T)^{-1}T^oA]^t(I_m - z\,T)^{-1}\,\underline{T}^o\cdot\,\underline{\beta}\,S^{j+t}\,\underline{S}^o.$$

Since

$$(15)\qquad \sum_{j=0}^{\infty}\binom{j+t+1}{j}\alpha_{m+1}^j\,S^j = [I_n - \alpha_{m+1}S]^{-t-2},$$

the sum reduces to

$$(16)\qquad \sum_{t=0}^{\infty} z\,\underline{\alpha}\,[z\,(I_m - z\,T)^{-1}T^oA]^t(I_m - z\,T)^{-1}\underline{T}^o\cdot$$
$$\underline{\beta}[I_n - \alpha_{m+1}S]^{-1}\{[I_n - \alpha_{m+1}S]^{-1}S\}^t\,[I_n - \alpha_{m+1}S]^{-1}\,\underline{S}^o$$
$$= z\{\underline{\alpha}\otimes\underline{\beta}[I_n - \alpha_{m+1}S]^{-1}\}\cdot\sum_{t=0}^{\infty}\{z(I_m - z\,T)^{-1}\,T^oA\otimes(I_n - \alpha_{m+1}S)^{-1}S\}^t\,\cdot$$
$$\{(I_m - z\,T)^{-1}\otimes I_n\}\{\underline{T}^o\otimes[I_n - \alpha_{m+1}S]^{-1}\underline{S}^o\},$$

by repeated application of the property

$$(17)\qquad (L\otimes M)(K\otimes N) = LK\otimes MN,$$

of the Kronecker product.

Finally

$$(18) \quad \sum_{t=0}^{\infty} \{z(I_m - z\,T)^{-1}T^oA \otimes (I_n - \alpha_{m+1}S)^{-1}S\}^t\{(I_m - z\,T)^{-1} \otimes I_n\}$$

$$= \{I_{mn} - z\,(I_m - z\,T)^{-1}T^oA \otimes (I_n - \alpha_{m+1}S)^{-1}S\}^{-1}\{(I_m - z\,T)^{-1} \otimes I_n\}$$

$$= \{(I_m - z\,T) \otimes I_n - z\,T^oA \otimes (I_n - \alpha_{m+1}S)^{-1}S\}^{-1}$$

$$= [I_{mn} - z\,T \otimes I_n - z\,T^oA \otimes (I_n - \alpha_{m+1}S)^{-1}S]^{-1} = [I_{mn} - z\,K]^{-1}.$$

We now turn to the discussion of the existence of the matrix inverses, used in the preceding manipulations. Since I - T and I - S are nonsingular, and since the matrices T and S are nonnegative, the inverses $[I_n - \alpha_{m+1}S]^{-1}$ and $[I_m - z\,T]^{-1}$, $|z| \leq 1$, exist. This implies the validity of the series representation in formula (15). The validity of the matrix series in formula (18) is shown if the spectral radius of the nonnegative matrix $(I_m - T)^{-1}T^oA \otimes (I_n - \alpha_{m+1}S)^{-1}S$ is less than one. Since

$$(19) \qquad (I_m - T)^{-1}T^oA\,\underline{e} = (1-\alpha_{m+1})\underline{e},$$

the matrix $(1-\alpha_{m+1})^{-1}(I_m-T)^{-1}T^oA$ is substochastic. The matrix $(1-\alpha_{m+1})(I_n-\alpha_{m+1}S)^{-1}$ is clearly also substochastic, so that the spectral radius of $(1-\alpha_{m+1})(I_n-\alpha_{m+1}S)^{-1}S$ is strictly less than one. The mn eigenvalues of the Kronecker product $(I_m-T)^{-1}T^oA \otimes (I_n-\alpha_{m+1}S)^{-1}S$ are the products of the eigenvalues of the matrices $(1-\alpha_{m+1})^{-1}(I_m-T)^{-1}T^oA$ and $(1-\alpha_{m+1})(I_n-\alpha_{m+1}S)^{-1}S$. All these products are less than one in modulus, which proves the desired result. The properties of Kronecker products used here, may be found in Marcus and Minc [4].

Theorem 1 may be used to compute the probability density $\{v_k\}$ recursively for given representations $(\underline{\alpha},T)$ and $(\underline{\beta},S)$. It also has the following consequence of interest to the Galton-Watson process.

Theorem 2

Consider a Galton-Watson process in which the number of offspring of an individual has the probability density $\{p_k\}$ of phase type with representation $(\underline{\alpha},T)$, with generating function $P(z)$, given by formula (10). The probability generating function $P_n(z)$ of the number of descendants in the n-th generation of a single progenitor is then given by the n-fold functional iterate of $P(z)$. It follows that the corresponding probability density $\{p_k(n)\}$ is of phase type and has the representation $[\underline{\alpha}(n),T(n)]$, where the order of the matrix $T(n)$ is m^n. The vector $\underline{\alpha}(n)$ and the matrix $T(n)$ are recursively defined by

$$\begin{aligned}
(20)\qquad & T(1) = T, \qquad \underline{\alpha}(1) = \underline{\alpha}, \\
& \underline{T}^o(1) = \underline{T}^o, \qquad \alpha^o(1) = \alpha_{m+1}, \\
& T(n+1) = T(n) \otimes I_m + T^o(n) \quad A(n) \otimes [I_m - \alpha^o(n)T]^{-1}T, \\
& \underline{T}^o(n+1) = \underline{T}^o(n) \otimes [I_m - \alpha^o(n)T]^{-1}\underline{T}^o, \\
& \underline{\alpha}(n+1) = \underline{\alpha}(n) \otimes \underline{\alpha}[I_m - \alpha^o(n)T]^{-1}, \\
& \alpha^o(n+1) = \alpha_{m+1} + \alpha^o(n) \ \underline{\alpha} \ [I_m - \alpha^o(n)T]^{-1}\underline{T}^o,
\end{aligned}$$

for $n \geq 1$. $A(n)$ is a diagonal matrix of order m^n with the components of $\underline{\alpha}(n)$ on the diagonal.

Proof

Immediate by repeated application of Theorem 1.

Theorem 2 generalizes the classical result [2], that in a Galton-Watson process in which the density $\{p_k\}$ is geometric, or more generally where $P(z) = p_0 + (1-p_0)pz(1-qz)^{-1}$, all the successive generation sizes have a geometric distribution. If $m \geq 2$, the algorithmic utility of the recursive formulas (20) is limited due to the rapid growth in the order of the matrices appearing in the representation. A feasible algorithm to compute the densities of the successive generation sizes is given below.

The last of the formulas (20) is computationally useful, since $\alpha^o(n)$ is the probability that the population is extinct by the n-th generation. The sequence $\{\alpha^o(n),\ n \geq 1\}$ is the probability distribution sequence of the time till extinction.

The following result is useful in certain computations related to the M/G/1 queue. Let X be the length of a random time interval with probability distribution $F(\cdot)$ with Laplace-Stieltjes transform $f(s)$. Consider a Poisson process of rate λ, independent of X, and let N be the number of arrivals in the Poisson process in $[0,X]$.

Theorem 3

If the random variable X has a (continuous parameter) distribution of phase type on $[0,\infty)$, with representation $(\underline{\alpha},T)$, then the random variable N is of phase type on the nonnegative integers with the representation $(\underline{\beta},S)$, given by

$$(21) \qquad S = \lambda(\lambda I - T)^{-1}, \qquad \underline{\beta} = \underline{\alpha}\ \lambda(\lambda I - T)^{-1},$$

and correspondingly

$$(22) \qquad \underline{S}^o = (\lambda I - T)^{-1}\underline{T}^o, \qquad \beta_{m+1} = \alpha_{m+1} + \underline{\alpha}\ (\lambda I - T)^{-1}\underline{T}^o.$$

Proof

The probability generating function $K(z)$ of N is given by

$$(23) \qquad K(z) = f(\lambda - \lambda z) = \alpha_{m+1} + \underline{\alpha}\ (\lambda I - \lambda z I - T)^{-1}\underline{T}^o,$$

for $|z| \leq 1$.

The function $K(z)$ may be rewritten as

$$(24) \quad K(z) = \alpha_{m+1} + \underline{\alpha}(\lambda I - T)^{-1}\underline{T}^o + z\ \underline{\alpha}\ \lambda(\lambda I - T)^{-1}[I - \lambda z(\lambda I - T)^{-1}]^{-1}(\lambda I - T)^{-1}\underline{T}^o.$$

It suffices to show that formula (24) is a valid representation of the generating function of a density of phase type. Since T is a stable matrix, whose inverse exists, the matrix $\lambda I - T$ is nonsingular. In order to show that $\lambda(\lambda I - T)^{-1}$ is substochastic, consider the probability A_{ij} that at the first event in the Poisson process, the Markov chain of phases is in the state $j \neq m+1$, given that at time 0 it was in the state $i \neq m+1$. The matrix $A = \{A_{ij}\}$ is then clearly substochastic. It is explicitly given by

$$(25) \qquad A = \int_0^\infty \lambda e^{-\lambda u} e^{Tu}\, du = \lambda(\lambda I - T)^{-1}.$$

Furthermore

$$(26) \qquad \lambda(\lambda I - T)^{-1}\underline{e} + (\lambda I - T)^{-1}\underline{T}^o = (\lambda I - T)^{-1}(\lambda\underline{e} + \underline{T}^o) = \underline{e},$$

since $\underline{T}^o = -T\,\underline{e}$.

Finally the matrix $I - \lambda(\lambda I - T)^{-1}$ is nonsingular, since

$$(27) \quad I - \lambda(\lambda I - T)^{-1} = (\lambda I - T)^{-1}(\lambda I - T - \lambda I) = -(\lambda I - T)^{-1}T,$$

and T is nonsingular.

The probability density $\{h_k\}$ of N is given explicitly by

$$(28) \qquad h_0 = \alpha_{m+1} + \underline{\alpha}\,(\lambda I - T)^{-1}\underline{T}^o,$$

$$h_k = \lambda^k\, \underline{\alpha}(\lambda I - T)^{-k-1}\underline{T}^o, \qquad \text{for } k \geq 1.$$

In the context of the $M^X/G/1$ queue with group arrivals, we consider the case where at each event in the Poisson process, a random group of customers arrives in the queue. If the probability density of the group sizes is also of phase type, it follows from Theorem 1 that the corresponding random variable N again has a density of phase type, since its generating function is given by $K[\varphi(z)]$, where $\varphi(z)$ denotes the probability generating function of the group sizes. The representation of the density of N can readily be constructed from that of $K(z)$ and $\varphi(z)$ by application of Theorem 1.

Formula (28) permits a particularly simple recursive computation of the density $\{h_k\}$ and therefore of the stationary queue length distribution in an M/G/1 queue in which the service time distribution is of phase type. In [8] it is further shown that the stationary distribution of the FIFO waiting time is itself of phase type and may be computed by solving a linear system of differential equations with constant coefficients.

2. The Successive Generation Sizes in a Galton-Watson Process

Except for particularly simple densities of the number of offspring per individual, the probability densities of the successive generation sizes are not easily computed. If $\{p_k(n)\}$ is the probability density of the size of the n-th generation, then we have

$$(29) \qquad \begin{aligned} p_k(1) &= p_k, \\ p_k(n+1) &= \sum_{\nu=0}^{\infty} p_\nu(n)\, p_k^{(\nu)}, \qquad \text{for } k \geq 0. \end{aligned}$$

Even if the original density $\{p_k\}$ is concentrated on the integers $0,\ldots,M$, the density $\{p_k(n)\}$ will be concentrated on the integers $0,\ldots,M^n$. For larger values of n, this will require a truncation of the density $\{p_k(n)\}$ with a resulting error which will be propagated in the computation of $\{p_k(n+1)\}$. This is a fortiori the case when the density $\{p_k\}$ does not have a bounded support. The rapid growth of the number of points in the support of $\{p_k(n)\}$ also appears to exclude the use of the fast Fourier transform as a feasible computational technique.

If the density $\{p_k\}$ is of phase type, it is possible to construct an algorithm which, at least for subcritical processes, is highly efficient and accurate. The main ingredient of the algorithm is a recursive procedure for the computation of the density of finite mixtures of the general type

$$(30) \qquad \underline{r} = a_0 + a_1\, \underline{p} + a_2\, \underline{p}^{(2)} + \ldots + a_N\, \underline{p}^{(N)},$$

where $\underline{p}^{(k)}$ denotes the k-fold convolution of the probability density $\underline{p}$, and a_j, $0 \leq j \leq N$, are probabilities whose sum is one.

The probability density $\underline{r}$ is itself of phase type and we now construct its representation. Let the density $\underline{p}$ have the representation $(\underline{\alpha}, T)$, where T is a matrix of order m. The density $\underline{r}$ will be represented as the probability density of the time till absorption in a Markov chain with Nm+1 states. For notational convenience we label the absorbing state 0 and the transient states 1,..., Nm. It will be further convenient to think in terms of an urn model, in which at time n=0, a random number J with $P\{J=j\} = a_j$, $0 \leq j \leq N$, of particles are placed. At time 0, we also "start" a Markov chain with m+1 states and transition probability matrix B, given by formula (3). The initial state is chosen with probabilities $\alpha_1,\ldots,\alpha_m$, α_{m+1}. If the state m+1 is drawn, a particle is removed from the urn and a new independent multinomial trial is performed. This is continued until either the urn is empty or until an initial state other than m+1 is selected. This procedure determines the content of the urn at time n=0+. If at time n=0+, the urn is not empty, we consider the Markov chain B at successive time points n=1,2,... As long as states other than m+1 are visited, no particles are removed from the urn. Whenever the state m+1 is reached, a particle is removed from the urn and an "instantaneous" sequence of multinomial trials is performed with probabilities $\alpha_1,\ldots,\alpha_m$, α_{m+1}. Whenever the state m+1 appears a particle is removed from the urn. The absorbing chain is restarted in this manner until the urn becomes empty. In order to construct the representation for $\underline{r}$, we consider the number J_n of particles in the urn at time n+, and the state I_n of the Markov chain B at time n+.

Clearly $0 \leq J_n \leq N$, and $1 \leq I_n \leq m$. If $J_n = 0$, I_n is not defined.

We shall say that the Markov chain, associated with $\underline{r}$ (the large chain) is in the state $(i-1)m+j$, $1 \leq i \leq N$, $1 \leq j \leq m$, at time n, if and only if $J_n = i$, $I_n = j$. It is clear that $\underline{r}$ is the probability density of the time until the urn becomes empty.

If we denote the representation of $\underline{r}$ by $(\underline{\gamma}, T^*)$, then $\underline{\gamma}$ is given by

$$(30) \qquad \gamma_{(i-1)m+j} = \gamma_{ij} = \sum_{\nu=i}^{N} a_\nu \alpha_{m+1}^{\nu-i} \alpha_j, \qquad \text{for } 1 \leq i \leq N,\ 1 \leq j \leq m.$$

$$\gamma_0 = \sum_{\nu=0}^{N} a_\nu \alpha_{m+1}^{\nu}.$$

The matrix T^*, which is of order m N, is of a block lower triangular form and may be written as

$$(31) \qquad T^* = \begin{bmatrix} T & 0 & 0 & \cdots & 0 \\ T^oA & T & 0 & \cdots & 0 \\ \alpha_{m+1}T^oA & T^oA & T & \cdots & 0 \\ \alpha_{m+1}^2 T^oA & \alpha_{m+1}T^oA & T^oA & \cdots & 0 \\ \vdots & \vdots & \vdots & & \\ \alpha_{m+1}^{N-2}T^oA & \alpha_{m+1}^{N-3}T^oA & \alpha_{m+1}^{N-4}T^oA & \cdots & T \end{bmatrix}$$

The corresponding column vector $\underline{T}^{*o}$ is given by $[\underline{T}^o, \alpha_{m+1}\underline{T}^o, \ldots \alpha_{m+1}^{N-1}\underline{T}^o]$.

The computation of the distribution of the density $\underline{r}$ is now equivalent to that of the probability distribution of the time till absorption in the (m N+1)-state Markov chain

$$(32) \qquad P^* = \begin{pmatrix} 1 & \underline{0} \\ \underline{T}^{*o} & T^* \end{pmatrix},$$

with the initial probability vector $(\gamma_0, \underline{\gamma})$. Consider a sequence of vectors $\underline{v}(n) = \{v_0(n), \underline{v}^{(1)}(n), \ldots, \underline{v}^{(N)}(n)\}$, of dimension m N+1, defined by

$$(33)\qquad \begin{aligned} v_0(0) &= \gamma_0, \qquad \{\underline{v}^{(1)}(0),\ldots,\underline{v}^{(N)}(0)\} = \underline{\gamma}, \\ v_0(n+1) &= v_0(n) + \sum_{\nu=1}^{N} \alpha_{m+1}^{\nu-1}\, \underline{v}^{(\nu)}(n)\underline{T}^{o}, \\ \underline{v}^{(i)}(n+1) &= \underline{v}^{(i)}(n)T + \sum_{\nu=i+1}^{N} \alpha_{m+1}^{\nu-i-1}\, \underline{\alpha}\, [\underline{v}^{(\nu)}(n)\underline{T}^{o}], \end{aligned}$$

for $1 \leq i \leq N$, and $n \geq 0$, then it is clear that the sequence $\{v_0(n),\ n \geq 0\}$ is the probability distribution corresponding to the required density. The recurrence relations (33) may be easily programmed for numerical computation. Since $v_0(n)$ tends to one, the sum of the other components of the vector $\underline{v}(n)$ tends to zero. In particular we have $\underline{v}^{(N)}(n+1) = \underline{v}^{(N)}(n)\ T = a_N\ \underline{\alpha}\ T^{n+1}$. For larger values of n, the recurrence relations (33) involve many computational steps which contribute only negligeable amounts to the terms $v_0(n)$. Many of these steps can be eliminated at the expense of a small error but with a significant reduction in computation time. This aspect is discussed below, but for use in the sequel, we first compute the inverse of the matrix $I - T^*$.

Theorem 4

The inverse J of the matrix $I - T^*$ is a block lower triangular matrix whose entries are given by

$$(34)\qquad \begin{aligned} J(i,i) &= (I-T)^{-1}, && \text{for } 1 \leq i \leq N, \\ J(i,j) &= D(I-T)^{-1}, && \text{for } i > j > 1. \end{aligned}$$

The matrix $D = (I-T)^{-1}T^{o}A$, has identical rows, all equal to $\underline{\alpha}$.

Proof

Since $(I-T)^{-1}\underline{T}^{o} = \underline{e}$, it is clear that D has the stated property. The diagonal blocks of $(I-T^*)J$ are all equal to I. Computing the off-diagonal blocks we obtain for $i > j > 1$, that

(35) $$[(I-T^*)J](i,j) = -\alpha_{m+1}^{i-j-1} T^o A(I-T)^{-1} - \sum_{\nu=0}^{i-j-2} \alpha_{m+1}^{i-j-\nu} T^o AD(I-T)^{-1} + (I-T)D\,(I-T)^{-1}.$$

Since $T^o AD = (1-\alpha_{m+1})T^o A$, it readily follows that all the off-diagonal blocks are zero.

Remark

We further note that since $\underline{\alpha}(I-T)^{-1}\underline{e} = \mu$, the mean of the density $\underline{p}$, it follows that $D\,\underline{e} = \mu\,\underline{e}$.

2.1. Adaptive Trimming

In the recursive computation of the vectors $\underline{v}(n)$ by means of the recurrence relations (33), the sums of the upper components of the nonnegative vectors $\underline{v}^{(i)}(n)$ for n large tend rapidly to zero. It is to our advantage to reduce the value of N for appropriate values of n, thereby saving a significant number of arithmetic operations in the computation of $\underline{r}$.

In a first method, we determine for each $n > 0$, the index $N_1(n) = \max_i \{i: \underline{v}^{(i)}(n)\cdot\underline{e} \geq \varepsilon\}$ and implement the recurrence relations for n+1 with N replaced by $N_1(n)$. If we denote the reduced matrix following the first trimming by $T^*(N_1)$, then the amount of probability neglected in the tail of the distribution after the first trimming is given by

(36) $$\underline{\gamma}T^{*n-1}(N)\underline{e} - \underline{\gamma}(N_1)T^{*n-1}(N_1)\underline{e}(N_1) = \sum_{i=N_1+1}^{N} \underline{v}^{(i)}(n)\underline{e} < (N-N_1)\varepsilon.$$

Continuing this procedure until $N_1(n)$ reaches zero, we obtain the computed sequence $\{\hat{v}_0(n),\ n \geq 0\}$, which is nondecreasing and satisfies $\hat{v}_0(n) \leq v_0(n)$ for all $n \geq 0$ and $1 - \hat{v}_0(\infty) < \varepsilon N$. An upper bound on the error in each of the density terms is $2\,\varepsilon\,N$, but this bound is very conservative.

The first trimming method is undesirable in repeated applications of the algorithm, as needed in computations for the Galton-Watson process, since the computed density corresponding to $\{\hat{v}_0(n)\}$ plays the role of the mixing density $\underline{a}$ for the next generation. The number of terms computed varies from one generation to the next and the accumulated error grows in a generally unpredictable manner. Moreover the computed distributions are defective.

A more conservative trimming procedure is the following: Let $N_2(0) = N$, and determine for each $n \geq 1$, the index $N_2(n)$ by

$$(37) \qquad N_2(n) = \max_{1 \leq i \leq N_2(n-1)} \{i: \sum_{j \geq i} (j-i+1)\ \hat{\underline{v}}^{(j)}(n) \cdot \underline{e} \geq \varepsilon\},$$

where $\hat{\underline{v}}^{(i)}(n)$ is the computed value of $\underline{v}^{(i)}(n)$. If the set in braces is empty, set $\hat{v}_0(n+1) = 1$, and stop. If $N_2(n) = N_2(n-1) \geq 1$, implement the recurrence relations (33) with $N = N_2(n)$. If $N_2(n-1) > N_2(n) \geq 1$, replace the vector $\hat{\underline{v}}(n)$ by the vector

$$(38) \qquad \hat{\underline{v}}^{(1)}(n), \ldots, \hat{\underline{v}}^{(N_2(n)-1)}(n), \sum_{i=N_2(n)}^{N_2(n-1)} \hat{v}^{(i)}(n), \underline{0}, \ldots, \underline{0}\}$$

and implement (33) with $N = N_2(n)$. This method has the advantage that no probability is "lost" in the recursive computation in the sense that for all $n \geq 0$, we have that

$$(39) \qquad \hat{v}_0(n) + \sum_{i=1}^{N} \hat{\underline{v}}^{(i)}(n)\ \underline{e} = 1.$$

The computed values $\hat{v}_0(n)$ now satisfy $\hat{v}_0(n) \geq v_0(n)$, for all $n \geq 0$, and the computed sequence $\{\hat{v}_0(n)\}$ is a probability distribution concentrating on a finite number of nonnegative integers.

An appropriate measure of the truncation error is the quantity

$$(40) \qquad \Delta(\varepsilon) = \sum_{n=0}^{\infty} [\hat{v}_0(n) - v_0(n)] = \sum_{n=0}^{\infty} [1-v_0(n)] - \sum_{n=0}^{\infty} [1-\hat{v}_0(n)].$$

We see that $\Delta(\varepsilon)$ is the difference between the exact mean of the desired distribution $\{v_0(n)\}$ and the mean of the computed distribution $\{\hat{v}_0(n)\}$.

We shall now obtain an estimate of the quantity $\Delta(\varepsilon)$, and to this end we first make a number of preliminary observations. Let the first trimming occur after the computation of $\underline{v}(n)$ and let it reduce N to N'. The modified vector defined in (41) may then be written as $\underline{v}(n)Z_{N'}$, where the matrix $Z_{N'}$ is defined as an N x N matrix of m x m blocks, with

$$(41)\qquad \begin{aligned} Z_{N'}(i,i) &= I_m, && \text{for } 1 \le i \le N' \\ Z_{N'}(i,N') &= I_m, && \text{for } N' \le i \le N \\ Z_{N'}(i,j) &= 0, && \text{for all other pairs.} \end{aligned}$$

The mean M of the density $\underline{r}$ is given by

$$(42)\qquad \begin{aligned} M &= \underline{\gamma}[I - T^*(N)]^{-1}\underline{e} \\ &= \underline{\gamma} \cdot \sum_{k=0}^{n-1} T^{*k}(N)\ \underline{e} + \underline{\gamma}\ T^{*n}(N)[I-T^*(N)]^{-1}\underline{e}. \end{aligned}$$

Following the first trimming, the mean of the computed distribution is reduced to

$$(43)\qquad M_1 = \underline{\gamma} \sum_{k=0}^{n-1} T^{*k}(N)\underline{e} + \underline{\gamma}\ T^{*n}(N)\ Z_{N'}[I-T^*(N)]^{-1}\underline{e}.$$

Using the explicit form of the inverse, obtained in Theorem 4, we obtain that

$$(44)\qquad [I-T^*(N)]^{-1}\underline{e} = \begin{pmatrix} (I-T)^{-1}\underline{e} \\ (I-T)^{-1}\underline{e} + \mu\ \underline{e} \\ (I-T)^{-1}\underline{e} + 2\ \mu\ \underline{e} \\ \vdots \\ (I-T)^{-1}\underline{e} + (N-1)\ \mu\ \underline{e} \end{pmatrix},$$

so that

$$(45) \qquad Z_{N'}[I-T^*(N)]^{-1}\underline{e} = \begin{pmatrix} (I-T)^{-1}\underline{e} \\ (I-T)^{-1}\underline{e} + \mu\,\underline{e} \\ \vdots \\ (I-T)^{-1}\underline{e} + (N'-1)\,\mu\,\underline{e} \\ \vdots \\ (I-T)^{-1}\underline{e} + (N'-1)\,\mu\,\underline{e} \end{pmatrix}.$$

It follows that

$$(46) \qquad M - M_1 = \mu \sum_{i=1}^{N-N'} i\, \underline{v}^{(N'+i)}\underline{e} < \mu\, \varepsilon.$$

Since the computation after the first trimming is similar in nature to the original one, we see that the mean of the computed distribution differs from the mean M by at most $N\,\mu\,\varepsilon$, so that

$$(47) \qquad \Delta(\varepsilon) < N\,\mu\,\varepsilon.$$

Remark

It is of course possible to implement the original recurrence relations (33) up to the smallest index n^* for which

$$(48) \qquad \sum_{n=0}^{n^*} [1-v_0(n)] > M-\varepsilon,$$

thereby guaranteeing that the mean of the computed distribution differs from the exact one by at most ε. The advantage of the adaptive trimming procedure lies in the progressive reduction of the number of operations involved in the recurrence relations (33), which is particularly significant for stable Galton-Watson processes.

2.2. The Successive Generation Sizes

The probability densities $\{p_k(n)\}$ of the successive generation sizes of a Galton-Watson process, in which the density $\{p_k\}$ is of phase type, may be computed by repeated applications of the algorithm developed above for the mixtures defined by formula (30). We note that the recurrence relations (29) are valid for a single progenitor. If there are ν progenitors with probability a_ν, $0 \le \nu \le N$, then the first equation in (29) should be replaced by

$$(49) \qquad p_0(1) = \sum_{\nu=0}^{N} a_\nu \, p_0^\nu,$$

$$p_k(1) = \sum_{\nu=1}^{N} a_\nu \, p_k^{(\nu)} , \qquad \text{for } k \ge 1.$$

The computation of $\{p_k(1)\}$ is clearly of the type defined by formula (30). It results in a computed density $\{\hat{p}_k(1),\ 0 \le k \le N_1\}$, which plays the role of the density $\{a_\nu\}$ in the computation of $\{p_k(2)\}$ and so on. By using the second adaptive trimming procedure, discussed above, we may use the computed means $M_1(n)$, $n \ge 0$, of the successive generation sizes to keep track of the accumulated truncation and trimming errors. The means $M(n)$ of the exact distributions are of course given by

$$(50) \qquad M(n) = \sum_{\nu=1}^{N} \nu a_\nu \, \mu^n, \qquad \text{for } n \ge 0.$$

For Galton-Watson processes for which μ is significantly less than one and the maximum initial population size N not too large, this method permits us to study the successive generation sizes until the extinction probability $p_0(n)$ becomes close to one. Computation times are generally small, on the order of a few seconds per generation. If μ is close to one, and a fortiori when μ is greater than one, the support of the successive computed densities increases with n and the computation time per generation increases quite rapidly.

We also note that this computational method needs to be only trivially modified to handle cases where the probability density of the number of offspring depends on the index of each generation. Immigration or removals from the population can also be studied by routine modifications of the algorithm.

The mean μ alone does not provide much information on the size of possible large excursions of the Galton-Watson process before extinction. It is possible in many cases however, to compute the distribution of the maximum generation size before extinction. The appropriate algorithm is discussed in Section 4.

The matrix T is usually very sparse and a major reduction in the computation time can be achieved by writing special purpose routines to compute the products $\underline{v}^{(i)}(n)\, T$ in the last formula (33).

3. A Galton-Watson Process embedded in the $M^X/G/1$ Queue.

Consider an $M^X/G/1$ queue with group arrivals and let the probability generating function of the density $\{a_k\}$ of the group sizes be $\varphi(z)$, with $\varphi(0) = 0$. The arrival rate (of groups) is λ and the service time distribution is denoted by $H(\cdot)$ with Laplace-Stieltjes transform $h(s)$. If the mean service time is μ and the mean group size is η, then it is well-known that the queue is stable if and only if $\lambda\eta\mu \leq 1$. Stationary distributions of the relevant queue features exist if and only if $\lambda\eta\mu < 1$.

Let t=0, be the beginning of a service and let the queue length $\xi(0)$ at t=0, be equal to i_0. Let T_1 be the time when all customers present at $T_0=0$, have been served under the FIFO discipline and let $\xi(T_1)$ denote the number of arrivals during the interval (T_0,T_1). Similarly T_2 is the time when all $\xi(T_1)$ customers present at time T_1+ have been served and $\xi(T_2)$ denotes the number of arrivals in (T_1,T_2). This construction is repeated to yield a bivariate sequence of random variables $\{(T_n,\ \xi(T_n),\ n \geq 0\}$. We shall agree that if $\xi(T_n)=0$, then T_{n+1} is the time when the group of customers, who arrive during the idle period starting at T_n, have completed service. The marginal sequence $\{\xi(T_n),\ n \geq 0\}$ is known to be a Markov chain on the nonnegative integers.

[3,7]. Its transition probability matrix U is given by

$$(51) \qquad U = \begin{bmatrix} a_0' & a_1' & a_2' & a_3' & \cdots \\ p_0 & p_1 & p_2 & p_3 & \cdots \\ p_0^{(2)} & p_1^{(2)} & p_2^{(2)} & p_3^{(2)} & \cdots \\ p_0^{(3)} & p_1^{(3)} & p_2^{(3)} & p_3^{(3)} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix},$$

where $\{p_k\}$ is the probability density with generating function $P(z) = h[\lambda-\lambda\,\varphi(z)]$, $\{p_k^{(\nu)}\}$ is its ν-fold convolution, and the density $\{a_k'\}$ is defined by

$$(52) \qquad a_k' = \sum_{j=1}^{\infty} a_j\, p_k^{(j)}, \qquad \text{for } k \geq 0.$$

The probability generating function A(z) of $\{a_k'\}$ is clearly given by

$$(53) \qquad A(z) = \varphi\{h[\lambda-\lambda\,\varphi(z)]\}, \qquad \text{for } |z| \leq 1.$$

We note that if $H(\cdot)$ is a (continuous) distribution of phase type and $\{a_k\}$ a (discrete) density of phase type, then by Theorem 3, $\{p_k\}$ is of phase type, and by Theorem 1, $\{a_k'\}$ is of phase type. Representations for $\{p_k\}$ and $\{a_k'\}$ may easily be constructed from those of $H(\cdot)$ and $\{a_k\}$.

Assuming henceforth that $\lambda\eta\mu < 1$, we proceed to discuss the stationary density $\{\pi_k\}$ of the recurrent Markov chain U. The quantities π_k, $k \geq 0$, satisfy the system of equations

$$(54) \qquad \pi_0\, a_\nu' + \sum_{j=1}^{\infty} \pi_j\, p_\nu^{(j)} = \pi_\nu, \qquad \text{for } \nu \geq 0$$

$$\sum_{j=0}^{\infty} \pi_j = 1.$$

Denoting the generating function of $\{\pi_j\}$ by $\pi(z)$, we obtain

$$(55) \qquad \pi(z) = \pi[P(z)] - \pi_0[1-A(z)], \qquad \text{for } |z| \leq 1.$$

Theorem 5

The probability generating function $\pi(z)$ is given by

$$(56)\quad \pi(z) = 1-\{1+\sum_{j=0}^{\infty} [1-A[P_j(0)]]\}^{-1} \sum_{\nu=0}^{\infty} [1-A[P_\nu(z)]], \qquad \text{for } |z| \leq 1,$$

where $P_\nu(z)$ is the ν-th functional iterate of $P(z)$. $P_0(z) = z$.

Proof

Replacing z by $P_n(z)$ in (55), we obtain

$$(57)\quad \pi[P_n(z)] = \pi[P_{n+1}(z)] - \pi_0 [1-A[P_n(z)]], \qquad \text{for } n \geq 0,$$

and hence

$$(58)\quad \pi(z) = \pi[P_{n+1}(z)] - \pi_0 \sum_{j=0}^{n} [1-A[P_j(z)]].$$

Since $P'(1) = \lambda\eta\mu < 1$, we know that $P_n(z) \to 1$, for all $0 \leq z \leq 1$, as n tends to infinity. The series of analytic functions

$$\sum_{j=0}^{\infty} [1-A[P_j(z)]]$$

converges uniformly for all $0 \leq z \leq 1$. This follows from the Lebesgue dominated convergence theorem, since

$$(59)\quad 1-A[P_j(z)] \leq 1-A[P_j(0)] \leq \theta_j,$$

where

$$\theta_j = [\frac{d}{dz} A[P_j(z)]]_{z=1} = \eta(\lambda\eta\mu)^{j+1}.$$

The second inequality in (59) is obtained by noting that the graph of the convex increasing function $A[P_j(z)]$ lies for every in $0 \leq z < 1$, above its tangent at z=1.

Passing to the limit in (58), we obtain

$$(60)\quad \pi(z) = 1-\pi_0 \sum_{j=0}^{\infty} [1-A[P_j(z)]],$$

for $0 \leq z \leq 1$. By analytic continuation, the same formula is valid for

$|z| \leq 1$. Setting $z=0$, in formula (60) we obtain

$$\pi_0 = \{1+\sum_{j=0}^{\infty}[1-A[P_j(0)]]\}^{-1}. \tag{61}$$

Remark

We note that

$$\sum_{j=0}^{\infty}[1-A[P_j(0)]] = d, \tag{62}$$

is the mean number of generations till extinction in a Galton-Watson process with offspring density $\{p_k\}$ and initial population size density $\{a_k'\}$. Applying the second inequality in (59) we obtain

$$d \leq \frac{\lambda\eta^2\mu}{1-\lambda\eta\mu}, \tag{63}$$

and hence

$$\pi_0 \geq \frac{1-\lambda\eta\nu}{1-\lambda\eta\mu(1-\eta)}. \tag{64}$$

By differentiating k times in (60), we obtain the explicit formula

$$\pi_k = \pi_0 \sum_{j=0}^{\infty} P(j,k), \qquad \text{for } k \geq 1, \tag{65}$$

where $P(j,k)$ is the probability that there are k individuals in the j-th generation of a Galton-Watson process with offspring density $\{p_k\}$ and initial population density $\{a_k'\}$. The initial population is counted as generation 0.

3.1. Computational Aspects

The density $\{\pi_k\}$ may be accurately computed for queues for which the underlying distributions are of phase type, by means of the recursive algorithm developed in Section 2; this at least if $\lambda\eta\mu$ is not too close to one. The quantities $P_j(0)$ can be efficiently computed by successive substitutions in the probability generating function, but each step involves two matrix inversions. The value of π_0 is computed in terms of the $P_j(0)$.

The term-wise sums of the densities $\{P(j,k),\ k \geq 1\}$ over the index j are formed and serve in the computation of π_k, for $k \geq 1$. It is advisable to compute π_0 separately and to a high accuracy. If the value of π_0 is essentially correct, we may use the normalizing condition $\Sigma\ \pi_k=1$, to determine the number of generations needed in (65) to obtain the probabilities π_k, $k \geq 1$, to a sufficient degree of accuracy. The computation of the density $\{\pi_k\}$ is of interest in the numerical investigation of the priority rules discussed by Nair and Neuts [5,6].

4. The Maximum Generation Size before Extinction

The random variable $Y = \max\{X_n,\ n \geq 0\}$ of the successive generation sizes before extinction in a Galton-Watson process has been discussed by J. Bishir [1] and E. Seneta [9].

For each $k \geq 1$, the system of linear equations

$$(66) \qquad y_i^{(k)} = \sum_{\nu=1}^{k} p_\nu^{(i)}\ y_\nu^{(k)} + p_0^i, \qquad \text{for } 1 \leq i \leq k,$$

has a unique solution $[y_1^{(k)},\ldots,\ y_k^{(k)}]$ and

$$(67) \qquad P\{Y \leq k\} = \sum_{i=1}^{k} P\{X_0=i\}\ y_i^{(k)}, \qquad \text{for } k \geq 1.$$

For a subcritical or critical Galton-Watson process the distribution of Y is honest, but for a supercritical process we have

$$(68) \qquad P\{Y < \infty\} = \sum_{i=1}^{\infty} P\{X_0=i\}\ \rho^i,$$

where ρ is the probability of extinction for the line of a single progenitor.

Bishir's paper does not enter into the construction of an efficient algorithm for the computation of the distribution of Y. The examples of highly subcritical or highly supercritical cases, presented in [1], are somewhat misleading in assessing the computational effort involved. We examined the following two methods for a large number of examples:

4.1. The Gauss-Seidel Method

For each k, the system of linear equations (66) satisfies sufficient conditions for the convergence of the Gauss-Seidel iterative method. It is easy to show that the quantities $y_i^{(k)}$ satisfy $y_1^{(k)} \geq y_2^{(k)} \geq \ldots \geq y_k^{(k)}$, since $y_i^{(k)}$ is the probability that a Galton-Watson process with i progenitors becomes extinct without exceeding the population size k. After solving the system of equations for k, it is convenient to use the (k+1)-tuple $y_1^{(k)},\ldots, y_k^{(k)}, y_k^{(k)}$, as a starting solution for the computation of the quantities $y_1^{(k+1)},\ldots, y_{k+1}^{(k+1)}$.

For Galton-Watson processes which are close to critical, and in general when systems in excess of k=75, need to be solved, the computation time for the Gauss-Seidel method becomes substantial and exceeds one minute of central processing time on a CDC 65000 computer.

4.2. The Gauss Elimination Method

Writing the system (66) as

$$(68) \qquad \sum_{\nu=1}^{k} (\delta_{i\nu} - p_\nu^{(i)})\, y_\nu^{(k)} = p_0^i, \qquad 1 \leq i \leq k,$$

assume that the system has been reduced to upper triangular form by elementary row operations represented by the lower triangular matrix K_k. The resulting system is written in the form

$$(69) \qquad H_k\, \underline{y}^{(k)} = \underline{c}^{(k)},$$

where H_k is upper triangular with $H_{k,11} = 1$. The system (69) is readily solved and $P\{Y \leq k\}$ is computed.

The appealing feature of this method is the easy computation of the matrix H_{k+1} and the vector $\underline{c}_{k+1}$. This is described in the following algorithmic steps to go from k to k+1:

Step 1: Compute p_{k+1}.

Step 2: Compute $p_{k+1}^{(j)}$, for $j = 1, \ldots, k$.

Step 3: Compute the terms p_0^{k+1} and $p_i^{(k+1)}$, for $1 \leq i \leq k+1$.

Step 4: Left-multiply the vector computed in Step 2 by the matrix K_k, to obtain the first k entries in the (k+1)-st column of H_{k+1}.

Step 5: Perform Gauss elimination on the row computed in Step 3, to obtain the (k+1)-st row of H_{k+1} and the (k+1)-st entry of $\underline{c}_{k+1}$.

Step 6: Compute $y_i^{(k+1)}$, for $1 \leq i \leq k+1$.

Step 7: Compute $P\{Y \leq k+1\}$. If $P\{Y \leq k+1\}$ is sufficiently close to the probability of eventual extinction, stop. If not, set k equal to k+1 and go to Step 1.

This method is much faster than the Gauss-Seidel method, but may be sensitive to the accuracy problems usually associated with Gauss elimination. The strong diagonal dominance of the coefficient matrix in the system (68) suggests that these problems will be minor. Both methods were compared in single precision on the CDC 6500, which is a computer with large word length. Even in examples where k ran up to one hundred, all computed probabilities agreed to at least four decimal places, but on computers with a shorter word length it is probably advisable to perform the latter method in double precision.

For the computation of $P\{Y \leq k\}$, $k \geq 1$, there is no particular advantage in assuming that $\{p_j\}$ is of phase type, except for the easy computation of the terms of the density. A minor drawback of the Gauss elimination lies in the substantial storage requirements. If we allow values of k up to one hundred, two storage arrays of size 10000 are required, one to store the quantities $\{p_i^{(j)}\}$ and the second one to store the entries of the matrices K_k and H_k.

5. The Probability of Eventual Absorption

It is well-known that the probability of extinction of the lineage of a single progenitor is given by the smallest positive root ρ of the equation

$$(70) \qquad z = \alpha_{m+1} + z\,\underline{\alpha}\,(I-z\,T)^{-1}\,\underline{T}^{o},$$

and that $\rho=1$, if and only if $\mu = \underline{\alpha}(I-T)^{-1}\underline{e} \leq 1$. For $\mu > 1$, we may compute ρ by successive substitutions or more efficiently by Newton's method. Since the derivative of the right hand side is given by

$$(71) \qquad P'(z) = \underline{\alpha}(I-z\,T)^{-2}\,\underline{T}^{o},$$

the successive Newton approximations are given by

$$(72) \quad z_{\nu+1} = [1-\underline{\alpha}(I-z_{\nu}T)^{-2}\underline{T}^{o}]^{-1}[\alpha_{m+1}+z_{\nu}\underline{\alpha}(I-z_{\nu}T)^{-1}\underline{T}^{o}-z_{\nu}\underline{\alpha}(I-z_{\nu}T)^{-2}\underline{T}^{o}]$$

$$= [1-\underline{\alpha}(I-z_{\nu}T)^{-2}\underline{T}^{o}]^{-1}[\alpha_{m+1}-z_{\nu}^{2}\underline{\alpha}(I-z_{\nu}T)^{-2}T\,\underline{T}^{o}]$$

Since the function $P(z)$ is convex increasing the sequence $\{z_{\nu}\}$, $0 \leq z_0 < 1$, always converges to ρ. Caution is needed when μ is very close to one, since in this case the first factor in (72) becomes very large. Note that $1-\underline{\alpha}(I-T)^{-2}\underline{T}^{o} = 1-\mu$. In all other cases Newton's method converges rapidly. Each iteration may be most efficiently computed as follows:

<u>Step 1</u>: Compute $(I-z_{\nu}T)^{-1}$.

<u>Step 2</u>: Compute $\underline{\alpha}(I-z_{\nu}T)^{-1} = \underline{y}$.

<u>Step 3</u>: Evaluate $\underline{u} = \underline{y}(I-z_{\nu}T)^{-1}$.

<u>Step 4</u>: Evaluate $1-\underline{u}\,\underline{T}^{o}$, and the second factor in (72) and compute $z_{\nu+1}$.

Remark

The preprint of this paper is available as a Technical Report, which contains FORTRAN programs to compute the distributions of the successive generation sizes and of the maximum Y, in addition to sample output and information on computation times. This document may be obtained from the Department of Statistics, Purdue University, West Lafayette, IN, 47907.

The author expresses appreciation to Professor Herman Rubin and to Mr. Charles Carson for helpful discussions.

Bibliography

[1] J. Bishir
Maximum Population Size in a Branching Process
Biometrics, 18, 394-403, 1962

[2] T. E. Harris
The Theory of Branching Processes
Springer Verlag, Berlin, 1963

[3] D. G. Kendall
Some Problems in the Theory of Queues
J.R.S.S., B, 13, 151-185, 1951

[4] M. Marcus and H. Minc
A Survey of Matrix Theory and Matrix Inequalities
Allyn & Bacon, Boston, 1964

[5] S. S. Nair and M. F. Neuts
A Priority Rule based on the Ranking of the Service Times for the M/G/1 Queue
Oper. Research, 17, 466-477, 1969

[6] S. S. Nair and M. F. Neuts
An Exact Comparison of the Waitingtimes under three Priority Rules
Oper. Research, 19, 414-423, 1971

[7] M. F. Neuts

The Queue with Poisson Input and General Service Times, treated as a Branching Process

Duke Math. Journ, 36, 215-232, 1969

[8] M. F. Neuts

Probability Distributions of Phase Type

Purdue Mimeo Series No. 374, Dept. of Stat., Purdue Univ., West Lafayette, IN, 47907, 1974

[9] E. Seneta

On the Maxima of Absorbing Markov Chains

Australian J. of Stat., 9, 93-102, 1967

Appendix I

This appendix contains the FORTRAN program for the computation of the distributions of the successive generation sizes in a Galton-Watson process in which the number of offspring of an individual has a probability density of phase type. A sample of the output is also given.

This program requires 45000 words of central memory on the CDC 6500. The computation times are very substantial for supercritical processes. For subcritical processes, the computation time is modest, unless the initial population size is large with high probability. In this case, the distribution of the number of offspring in each generation of a single progenitor can be computed and the fast Fourier transform can be used to evaluate the distribution of each generation for the given initial conditions.

```
      PROGRAM CONVO(INPUT,OUTPUT,TAPE5=INPUT)
*******
*
* THIS PROGRAM WAS WRITTEN BY PROF. MARCEL F. NEUTS TO
* IMPLEMENT THE NUMERICAL COMPUTATIONS OF THE DISTRIBUTIONS
* OF THE SUCCESSIVE GENERATION SIZES IN A GALTON-WATSON PRO-
* CESS, DISCUSSED IN THE PAPER
*
*** SOME COMPUTATIONAL PROBLEMS RELATED TO THE GALTON - WATSON
*** PROCESS
*
* PURDUE MIMEO SERIES - DEPARTMENT OF STATISTICS - 1975
* WEST LAFAYETTE IN.  USA
*
*******
       COMMON/B1/N,A0,A(300),M,B0,B(20),B1(300),KG,T0(20),
      1T(20,20),TRIM,TEST,V0,VV(300),V(300,20),Y(300),X(20),
      1Z(20),AMEAN,TMEAN,RMEAN,NCOMP
       READ(5,1000) TRIM,TEST
*
*     TRIM AND TEST ARE PARAMETERS USED IN TRUNCATING THE COMPUTED
*     DISTRIBUTIONS. THE VALUE 0.00001 IS RECOMMENDED FOR BOTH
*
*     V0 IS THE INITIAL TERM OF THE COMPUTED DENSITY
*     VV(.) CONTAINS THE OTHER TERMS OF THE COMPUTED DISTRIBUTION
*     OR DENSITY
*     THE ARRAY V(.,.) IS USED FOR THE STORAGE OF THE VECTORS
*     V(I), WHICH ARISE IN THE RECURSIVE COMPUTATION OF THE DENSITY
*     OF INTEREST
*     THE ARRAYS X(.),Y(.) AND Z(.) ARE INTERMEDIATE STORAGE
*     ARRAYS, WHICH SERVE SEVERAL PURPOSES IN THE COURSE OF THE
*     PROGRAM. THEIR PARTICULAR USES ARE CLEAR BY EXAMINING THE
*     CONTEXTS IN WHICH THEY ARE USED.
*
   30 READ(5,1001) N,M,KG
*
* N  IS THE MAXIMUM NUMBER OF INDIVIDUALS IN THE ORIGINAL
*     GENERATION. IF THERE IS ONLY ONE PROGENITOR, SET N=1.
* M  IS THE ORDER OF THE MATRIX T ARISING IN THE REPRESENTA-
*     TION OF THE DISTRIBUTION OF THE NUMBER OF OFFSPRING OF AN
*     INDIVIDUAL. THIS DISTRIBUTION IS ASSUMED TO BE OF PHASE TYPE
* KG IS THE MAXIMUM NUMBER OF GENERATIONS FOR WHICH IT IS
*     DESIRABLE TO COMPUTE THE DENSITY. FOR SUPER-CRITICAL
*     PROCESSES THE COMPUTATION TIME IS SUBSTANTIAL IF KG IS
*     LARGE.
*
       IF(EOF,5) 31,32
   32 CALL SECOND(T1)
       READ(5,1002) A0,(A(I),I=1,N)
*
* A0, A(I), I=1,...,N, IS THE DENSITY OF THE INITIAL POPULATION
*
       READ(5,1002) B0,(B(I),I=1,M)
*
*     B0, B(I), I=1,...,M, IS THE INITIAL DENSITY NEEDED IN THE
*     REPRESENTATION OF THE DENSITY OF PHASE TYPE. B0 IS ALPHA(SUB
*     M+1) IN THE PAPER AND THE B(I) ARE THE COMPONENTS OF THE VECTOR
*     ALPHA USED IN THE PAPER.
*
       IF(KG.EQ.0) KG=1
       DO 1 I=1,M
```

```
      READ(5,1002) (T(I,J),J=1,M)
*
* T(I,J)  IS THE (I,J)-COMPONENT OF THE MATRIX T ARISING IN THE
*    REPRESENTATION OF THE DISTRIBUTICN OF PHASE TYPE.
*
      U=1.
*
* COMPUTATION OF THE VECTOR T(SUPER 0)
*
      DO 2 J=1,M
    2 U=U-T(I,J)
    1 T0(I)=U
*
* PRINT-OUT OF THE DATA
*
      PRINT 1003
      NO=0
      PRINT 1004,NO,A0,(I,A(I),I=1,N)
      PRINT 1005
      PRINT 1004,NO,B0,(I,B(I),I=1,M)
      PRINT 1006
*
* INCIDENTAL COMPUTATION OF THE MEAN OF THE DISTRIBUTION OF
* PHASE TYPE. NOTE EFFICIENT USE OF ARRAYS.
*
      DO 3 I=1,M
      PRINT 1004,NO,T0(I),(J,T(I,J),J=1,M)
      X(I)=-1.
    3 T(I,I)=T(I,I)-1.
*
* LINEQ1 IS A LIBRARY ROUTINE TO SOLVE THE SYSTEM OF LINEAR
* EQUATIONS TZ=X
*
      CALL LINEQ1(T,X,Z,20,M,1,LL)
      IF(LL.EQ.0) GO TO 4
      GO TO 33
    4 U=0.
      DO 5 I=1,M
      T(I,I)=T(I,I)+1.
    5 U=U+B(I)*Z(I)
*
* TMEAN IS THE MEAN OF THE DISTRIBUTION OF PHASE TYPE
*
      TMEAN=U
*
* COMPUTATION OF THE MEAN OF THE FIRST GENERATION SIZE
* SOME ELEMENTARY DIAGNOSTIC TESTS ON THE DATA
*
      U=1.-A0
      AMEAN=0.
*
* AMEAN IS THE MEAN OF THE DENSITY A0,A(1),....,A(N) - IN THE GALTON
* WATSON PROCESS, IT IS THE EXPECTED NUMBER OF PROGENITORS.
*
* BMEAN IS THE MEAN OF THE FIRST COMPUTED DENSITY AND LATER THAT
* OF THE SUCCESSIVE GENERATION SIZES.
*
      DO 6 I=1,N
      U=U-A(I)
    6 AMEAN=AMEAN+I*A(I)
      IF(ABS(U).GT.1.E-5) GO TO 33
```

```
      RMEAN=AMEAN*TMEAN
      PRINT 1007,AMEAN,TMEAN,RMEAN
      B1(1)=B0
      PRINT 1015,TRIM,TEST,KG
      DO 7 I=2,300
      B1(I)=B1(I-1)*B0
      IF(B1(I).LT.1.E-15) GO TO 8
    7 CONTINUE
    8 CALL SECOND(T3)
*
* SUBROUTINE INIT COMPUTES THE INITIAL CONDITIONS USED IN
* THE RECURSIVE SCHEME TO EVALUATE THE MIXTURE OF SUCESSIVE
* CONVOLUTIONS DISCUSSED IN SECTION 2 OF THE PAPER.
*
      CALL INIT
      DO 9 KG1=1,KG
      PRINT 1008,KG1
*
* SUBROUTINE MIXT EVALUATES THE DISTRIBUTION OF THE MIXTURE
* BY IMPLEMENTING THE RECURSIVE COMPUTATION DISCUSSED IN SECTION
* 2 OF THE PAPER.
*
* NCOMP IS THE HIGHEST INTEGER IN THE SUPPORT OF THE COMPUTED
* DISTRIBUTION OF THE MIXTURE.
*
      CALL MIXT
      UX=1.-VV(NCOMP)
*
* PRINT-OUT OF THE DISTRIBUTION OF THE MIXTURE, I.E. THE
* DISTRIBUTION OF THE KG1-ST GENERATION.
*
      PRINT 1004,N0,V0,(I,VV(I),I=1,NCOMP)
*
* COMPUTATION OF THE MEAN OF THE COMPUTED DISTRIBUTION AND
* OF ITS DENSITY
*
      RMN=1.-V0
      NX=NCOMP-1
      DO 12 J1=1,NX
      J=NCOMP-J1+1
      RMN=1.-VV(J)+RMN
   12 VV(J)=VV(J)-VV(J-1)
      RMN=1.-VV(1)+RMN
      VV(1)=VV(1)-V0
      PRINT 1011
*
* PRINT-OUT OF THE DENSITY
*
      PRINT 1004,N0,V0,(I,VV(I),I=1,NCOMP)
      RMEAN=AMEAN*TMEAN**KG1
*
* PRINT-OUT OF THE EXACT MEAN, THE COMPUTED MEAN AND THE
* PERCENTAGE ERROR BETWEEN BOTH.
*
      PRINT 1012,RMEAN,RMN
      PERR=100.*(1.-RMN/RMEAN)
      PRINT 1014,PERR
      IF(KG1.EQ.KG) GO TO 13
      N=NCOMP
*
* WITH THE PRESENT DIMENSIONS ONLY DENSITIES WHOSE MAXIMUM
```

```
*  SUPPORT POINT DOES NOT EXCEED THREE HUNDRED CAN BE HANDLED.
*
      IF(N.EQ.300) GO TO 13
*
*  INTERCHANGE OF THE COMPUTED DENSITY WITH THE DENSITY A(.)
*  FOR COMPUTATION OF THE NEXT GENERATION.
*
      A0=V0
      DO 11 I=1,NCOMP
   11 A(I)=VV(I)
      N=N+1
      A(N)=UX
      CALL SECOND(T4)
      T5=T4-T3
      PRINT 1009,T5
      T3=T4
      CALL INIT
    9 CONTINUE
      GO TO 10
*
*  STATEMENTS NEEDED TO TIME THE EXECUTION OF THE ALGORITHM,
*  TO PERFORM EXITS IN CASES OF CERTAIN ERRORS AND TO RETURN TO
*  THE INITIAL READ FOR THE NEXT SET OF DATA IF THERE
*  IS ONE PRESENT.
*
   13 PRINT 1013
   10 CALL SECOND(T2)
      U=T2-T1
      U1=T2-T3
      PRINT 1009,U1
      PRINT 1016,U
      GO TO 30
   31 CALL EXIT
   33 PRINT 1010
*
*  THE FORMAT STATEMENTS
*
 1000 FORMAT(2F7.0)
 1001 FORMAT(I4,I3,I4)
 1002 FORMAT(5F7.5)
 1003 FORMAT(*1*//3X*REPEATED MIXTURES OF THE SUCCESSIVE *
     1*CONVOLUTIONS OF A PROBABILITY */3X*DENSITY OF PHASE *
     1*TYPE*//3X*THE INITIAL MIXING DENSITY A0,A(1),...,*
     1*A(N)*//)
 1004 FORMAT(3X,8(I4,F8.5))
 1005 FORMAT(///3X*THE INITIAL PROBABILITY DENSITY B0,B(1),*
     1*...,B(M)*//)
 1006 FORMAT(///3X*THE PHASE MATRIX*//)
 1007 FORMAT(///3X*THE MEAN OF THE MIXING DENSITY =*F9.5//
     13X*THE MEAN OF THE PHASE DENSITY =*F9.5//3X*THE MEAN *
     1*OF THE MIXTURE =*F9.5)
 1008 FORMAT(*1*//3X*GENERATION NO.*I3///)
 1009 FORMAT(//3X*COMPUTATION TIME FOR THIS GENERATION =*F8.4)
 1010 FORMAT(///3X*INPUT ERROR*)
 1011 FORMAT(4(/))
 1012 FORMAT(//3X*THE EXACT MEAN =*F9.5//3X*THE APPROXIMATE *
     1*MEAN =*F9.5)
 1013 FORMAT(*2*3X*THE CAPACITY OF THE STORAGE ARRAYS IS *
     1*EXHAUSTED*)
 1014 FORMAT(//3X*THE PERCENTAGE ERROR IN THE COMPUTED MEAN =*
     1F9.5)
```

```
1015 FORMAT(//3X*TRIM =*E10.3,5X*TEST =*E10.3//3X*NUMBER OF *
    1*GENERATIONS =*I3)
1016 FORMAT(//3X*TOTAL PROCESSING TIME FOR KG GENERATIONS =*
    1F8.4)
     END
```

```
      SUBROUTINE INIT
*
* THE SUBROUTINE INIT INITIALIZES THE ARRAYS FOR THE RECURSIVE
* COMPUTATION OF THE MIXTURES DISCUSSED IN SECTION 2 OF THE PAPER
*
      COMMON/B1/N,A0,A(300),M,B0,B(20),B1(300),KG,T0(20),
     1T(20,20),TRIM,TEST,V0,VV(300),V(300,20),Y(300),X(20),
     1Z(20),AMEAN,TMEAN,RMEAN,NCOMP
      V0=A0
      DO 1 I=1,N
    1 V0=V0+A(I)*B1(I)
      Y(N)=A(N)
      IF(N.EQ.1) GO TO 6
      N1=N-1
      DO 2 I=1,N1
      U=A(I)
      I1=I+1
      DO 3 NU=I1,N
    3 U=U+B1(NU-I)*A(NU)
    2 Y(I)=U
    6 DO 4 I=1,N
      U=Y(I)
      DO 4 J=1,M
    4 V(I,J)=U*B(J)
      RETURN
      END
```

```
      SUBROUTINE MIXT
*
* THE SUBROUTINE MIXT EVALUATES THE DISTRIBUTION OF THE MIXTURE
* BY THE RECURSIVE METHOD DISCUSSED IN SECTION 2 OF THE PAPER
*
      COMMON/B1/N,A0,A(300),M,B0,B(20),B1(300),KG,T0(20),
     1T(20,20),TRIM,TEST,V0,VV(300),V(300,20),Y(300),X(20),
     1Z(20),AMEAN,TMEAN,BMEAN,NCOMP
*
* THE QUANTITY XX KEEPS TRACK OF THE TOTAL AMOUNT OF PROBABILITY
* ALREADY COMPUTED OF THE DENSITY OF THE MIXTURE
*
      XX=V0
      DO 1 NV=1,300
      DO 2 NU=1,N
      U=0.
      DO 3 J=1,M
    3 U=U+V(NU,J)*T0(J)
    2 Y(NU)=U
      U=XX+Y(1)
      IF(N.EQ.1) GO TO 4
      DO 5 J=2,N
    5 U=U+Y(J)*B1(J-1)
    4 XX=VV(NV)=U
      XXX=1.-XX
      IF(XXX.LE.TEST) GO TO 6
      IF(N.EQ.1) GO TO 7
      IF(N.EQ.2) GO TO 8
      N1=N-2
      DO 9 I=1,N1
      U=Y(I+1)
      I1=I+2
      DO 10 J1=I1,N
   10 U=U+Y(J1)*B1(J1-I-1)
    9 Y(I)=U
    8 Y(N-1)=Y(N)
    7 DO 11 J=1,M
      U=0.
*
* THE SPARSITY OF THE MATRIX T CAN BE EXPLOITED HERE. THE NEXT
* FEW LINES SHOULD THEN BE REPROGRAMMED TO USE THE SPARSITY OF
* THE MATRIX T.
*
      DO 12 NU=1,M
   12 U=U+V(N,NU)*T(NU,J)
   11 X(J)=U
      DO 13 J=1,M
   13 V(N,J)=X(J)
      IF(N.EQ.1) GO TO 20
      N1=N-1
      DO 14 I=1,N1
      U1=Y(I)
*
* THE SPARSITY OF THE MATRIX T CAN BE EXPLOITED HERE. THE NEXT
* FEW LINES SHOULD THEN BE REPROGRAMMED TO USE THE SPARSITY OF
* THE MATRIX T.
*
      DO 15 J=1,M
      U=U1*B(J)
      DO 16 NU=1,M
   16 U=U+V(I,NU)*T(NU,J)
```

```
   15 X(J)=U
      DO 17 J=1,M
   17 V(I,J)=X(J)
   14 CONTINUE
   20 CONTINUE
*
* THE SUBROUTINE TRIMS IS CALLED ONLY AT EVERY M-TH STEP OF THE
* ITERATIVE COMPUTATION. IT IS NOT CLEAR THAT THIS RESULTS IN
* A SIGNIFICANT SAVING IN COMPUTATION TIME IN GENERAL. BY
* MODIFYING THE NEXT TWO LINES IT IS POSSIBLE TO CALL TRIMS
* EVERY TIME.
*
      NTST=MOD(NV,M)
      IF(NTST.EQ.0) CALL TRIMS
    1 CONTINUE
    6 NCOMP=NV
      RETURN
      END
```

```
      SUBROUTINE TRIMS
*
* THE SUBROUTINE TRIMS IMPLEMENTS THE SECOND ADAPTIVE TRIMMING
* PROCEDURE DISCUSSED IN SECTION 2 OF THE PAPER.
*
      COMMON/B1/N,A0,A(300),M,B0,B(20),B1(300),KG,T0(20),
     1T(20,20),TRIM,TEST,V0,VV(300),V(300,20),Y(300),X(20),
     1Z(20),AMEAN,TMEAN,RMEAN,NCOMP
      IF(N.EQ.1)RETURN
      U=U1=U2=0.0
      DO 1 N1=1,N
      NN=N-N1+1
      U1=0.0
      DO 2 J=1,M
    2 U1=U1+V(NN,J)
      U2=U2+U1
      U=U+U2
      IF(U.GT.TRIM.OR.NN.EQ.1) GO TO 4
      DO 3 J=1,M
    3 V(NN-1,J)=V(NN-1,J)+V(NN,J)
    1 CONTINUE
    4 N=NN
      RETURN
      END
```

REPEATED MIXTURES OF THE SUCCESSIVE CONVOLUTICNS OF A PROBABILITY DENSITY OF PHASE TYPE

THE INITIAL MIXING DENSITY A0,A(1),....,A(N)

```
0 0.00000   1  .25000   2  .75000
```

THE INITIAL PROBABILITY DENSITY B0,B(1),....,B(M)

```
0  .45000   1  .25000   2  .30000
```

THE PHASE MATRIX

```
0  .60000   1  .10000   2  .30000
0  .70000   1  .15000   2  .15000
```

THE MEAN OF THE MIXING DENSITY = 1.75000

THE MEAN OF THE PHASE DENSITY =   .83681

THE MEAN OF THE MIXTURE = 1.46441

TRIM = 1.000E-05   TEST = 1.000E-05

NUMBER OF GENERATIONS = 5

GENERATION NO. 4

0	.69913	1	.80320	2	.87219	3	.91719	4	.94649	5	.96550	6	.97780	7	.98574
8	.99086	9	.99415	10	.99626	11	.99761	12	.99848	13	.99903	14	.99939	15	.99961
16	.99976	17	.99985	18	.99991	19	.99994	20	.99997	21	.99998	22	.99999		

0	.69913	1	.10407	2	.06899	3	.04501	4	.02930	5	.01901	6	.01230	7	.00794
8	.00512	9	.00329	10	.00211	11	.00135	12	.00087	13	.00055	14	.00035	15	.00023
16	.00014	17	.00009	18	.00006	19	.00004	20	.00002	21	.00001	22	.00001		

THE EXACT MEAN = .85810

THE APPROXIMATE MEAN = .85799

THE PERCENTAGE ERROR IN THE COMPUTED MEAN = .01286

COMPUTATION TIME FOR THIS GENERATION = .1650

Appendix II

This appendix contains the FORTRAN program for the computation of the distribution of the maximum generation size before extinction in a Galton-Watson process in which the number of offspring of an individual has a probability density of phase type. The density of the number of offspring of a single individual, and that of the number of generations till extinction are also computed. A sample output is given.

This program requires 76000 words of central memory on the CDC 6500. This permits to compute up to one hundred terms of the densities of interest. The computation time is under one minute, and is not sensitive to the order m of the phase matrix T. This program was validated by comparison with a different algorithm, based on Gauss-Seidel solution of the successive systems of equations. The latter procedure requires a much larger processing time.

```
      PROGRAM CONVO(INPUT,OUTPUT,TAPE5=INPUT)
*******
*
* THIS PROGRAM WAS WRITTEN BY PROF. MARCEL F. NEUTS TO
* IMPLEMENT THE NUMERICAL COMPUTATIONS OF THE DISTRIBUTIONS
* OF THE MAXIMUM GENERATION SIZE BEFORE EXTINCTION AND OF
* THE NUMBER OF GENERATIONS BEFORE EXTINCTION IN A GALTON-
* WATSON PROCESS. FOR CONVENIENCE IT IS ASSUMED THE DENSITY
* OF THE NUMBER OF OFFSPRING OF EACH INDIVIDUAL IS OF PHASE TYPE.
* THESE DISTRIBUTIONS ARE DISCUSSED IN THE PAPER
*
*** SOME COMPUTATIONAL PROBLEMS RELATED TO THE GALTON-WATSON
*** PROCESS
*
* PURDUE MIMEO SERIES - DEPARTMENT OF STATISTICS - 1975
* WEST LAFAYETTE IN. USA
*
*******
      COMMON/BL/P1(100),P(100,100),A(10),B0,B(20),T0(20),
     1T(20,20),TT(20,20),E(20,20),F(20,20),X(100),Y(20),Z(20),
     1G(100),C(100),H(100,100),H1(100),M,N,TEST,TEST1,TEST2,MAXITER,NCRIT
*
*    THE ARRAYS P1(.) AND P(.,.) CONTAIN THE RELEVANT TERMS
*    OF THE SUCCESSIVE CONVOLUTIONS OF THE DENSITY (P-SUB K)
*    EXCEPT FOR INITIAL TERMS THE ENTRIES OF THESE ARRAYS
*    ARE COMPUTED BY THE SUBROUTINES TERM AND NEXT
*
*    THE ARRAYS H1(.) AND H(.,.) CONTAIN THE QUANTITIES OBTAINED
*    BY APPLYING THE GAUSS ELIMINATION TO THE SUCCESSIVE
*    SYSTEMS OF LINEAR EQUATIONS WHICH NEED TO BE SOLVED.
*    THE ENTRIES STRICTLY ABOVE THE DIAGONAL IN H(.,.) ARE
*    THE COEFFICIENTS OF THE MATRICES OF THE SYSTEMS AFTER
*    REDUCTION TO UPPER TRANGULAR FORM. THE DIAGONAL ENTRIES AND
*    THE LOWER DIAGONAL ENTRIES KEEP TRACK OF THE ELEMENTARY
*    ROW OPERATIONS ARISING FROM THE GAUSS ELIMINATION.
*
*    THE ARRAY G(.) CONTAINS THE DISTRIBUTION AND LATER THE DENSITY
*    OF THE MAXIMUM GENERATION SIZE BEFORE EXTINCTION, STARTING
*    WITH THE GIVEN INITIAL POPULATION SIZE.
*
*    THE ARRAY C(.) CONTAINS THE DENSITY OF THE NUMBER OF GENERATIONS
*    TILL ABSORPTION.
*
*    THE ARRAYS E(.,.), F(.,.), X(.), Y(.), Z(.) ARE AUXILIARY
*    ARRAYS, WHICH CONTAIN VARIOUS INTERMEDIARY RESULTS. THEIR
*    SIGNIFICANCE IS CLEAR FROM THE CONTEXT IN WHICH THEY OCCUR.
*
      MAXITER=40
*
*    MAXITER CONTAINS THE MAXIMUM NUMBER OF ITERATIONS TO BE
*    PERFORMED IN COMPUTING THE PROBABILITY OF EVENTUAL EXTINCTION
*    BY MEANS OF NEWTON-S METHOD.
*
      TEST=0.0005
      NCRIT=0
      TEST2=TEST/10.
*
* N  IS THE MAXIMUM NUMBER OF INDIVIDUALS IN THE ORIGINAL
*    GENERATION OF PROGENITORS. IF THERE IS ONLY ONE PROGENITOR
*    SET N=1
*
```

```
* M  IS THE ORDER OF THE MATRIX ARISING IN THE REPRESENTATION
*     OF THE DISTRIBUTION OF THE NUMBER OF OFFSPRING OF A
*     SINGLE INDIVIDUAL. THIS DISTRIBUTION IS ASSUMED TO BE OF PHASE
*     TYPE.
*
   30 READ(5,1001) N,M
      IF(EOF,5) 31,32
   32 CALL SECOND(T1)
*
* A0, A(I), I=1,...,N IS THE DENSITY OF THE INITIAL POPULATION
*
      READ(5,1002) (A(I),I=1,N)
*
* B0, B(I), I=1,...,M IS THE INITIAL DENSITY NEEDED IN THE
*    REPRESENTATION OF THE DENSITY OF PHASE TYPE. B0 IS ALPHA(SUB
*    M+1) IN THE PAPER AND THE B(I) ARE THE COMPONENTS OF THE
*    VECTOR ALPHA USED IN THE PAPER.
*
      READ(5,1002) B0,(B(I),I=1,M)
*
*    THE MATRIX T, WITH ENTRIES T(I,J) IS THE MATRIX ARISING
*    IN THE REPRESENTATION OF THE DENSITY OF PHASE TYPE, WHICH
*    IS THE DENSITY OF THE NUMBER OF OFFSPRING OF A SINGLE INDIVIDUAL.
*
*    THE FOLLOWING DO-LOOP ALSO COMPUTES THE VECTOR T(SUPER 0)
*
      DO 1 I=1,M
      READ(5,1002) (T(I,J),J=1,M)
      U=1.
      DO 2 J=1,M
    2 U=U-T(I,J)
    1 T0(I)=U
*
* PRINT-OUT OF THE DATA OF THE PROBLEM
*
      PRINT 1003
      N0=0
      PRINT 1004,(I,A(I),I=1,N)
      PRINT 1005
      PRINT 1004,N0,B0,(I,B(I),I=1,M)
      PRINT 1006
      DO 3 I=1,M
      PRINT 1004,N0,T0(I),(J,T(I,J),J=1,M)
      X(I)=-1.
    3 T(I,I)=T(I,I)-1.
*
* AT THIS STAGE THE VECTOR Z = (I-T)**(-1)*VECTOR(E) IS COMPUTED
*
      CALL LINEQ1(T,X,Z,20,M,1,LL)
      IF(LL.EQ.0) GO TO 4
      GO TO 33
    4 U=0.
      U1=1.-B0
      DO 5 I=1,M
      U1=U1-B(I)
      T(I,I)=T(I,I)+1.
*
* COMPUTATION OF THE MEAN TMEAN OF THE DENSITY OF PHASE TYPE
*
    5 U=U+B(I)*Z(I)
      IF(ABS(U1).GT.1.E-5) GO TO 33
```

```
      TMEAN=U
      PRINT 1011,U
      U=1.
      DO 6 I=1,N
    6 U=U-A(I)
      IF(ABS(U).GT.1.E-5) GO TO 33
      IF(TMEAN.LT.1.0001) GO TO 7
*
* THE SUBROUTINE NEWTON COMPUTES THE PROBABILITY OF EXTINCTION
* OF THE LINE OF A SINGLE PROGENITOR. NOTE THAT WE SET
* RHO EQUAL TO ONE IF TMEAN IS LESS THAN 1.0001 IN ORDER TO
* AVOID NUMERICAL PROBLEMS IN THE NEWTON ITERATIONS FOR
* PROBLEMS WHERE TMEAN IS VERY CLOSE TO ONE
*
      CALL NEWTON(RHO)
      IF(NCRIT.EQ.1) CALL EXIT
      YY=0.0
      DO 18 J=1,N
   18 YY=YY+A(J)*RHO**J
      GO TO 8
    7 RHO=YY=1.
*
* YY IS THE PROBABILITY OF EVENTUAL EXTINCTION FOR THE GIVEN
* INITIAL CONDITIONS
*
    8 PRINT 1009,YY
      TEST1=YY-TEST
      P1(1)=B0
      P1(2)=B0**2
      DO 9 J=1,M
      U=0.0
      DO 10 I=1,M
   10 U=U+B(I)*T(I,J)
    9 Y(J)=U
*
* FOR CONVENIENCE THE FIRST TWO TERMS OF THE DENSITY AND OF THE
* SECOND CONVOLUTION ARE COMPUTED HERE, EXPLICITLY. THE CONVENIENCE
* IS IN THE INDEXING FOR THE SUBROUTINE NEXT.
*
      U=0.0
      U1=0.0
      DO 11 I=1,M
      U1=U1+B(I)*T0(I)
   11 U=U+Y(I)*T0(I)
      P(1,1)=U1
      P(1,2)=U
      P(2,1)=2.*B0*U1
      P(2,2)=2.*B0*U+U1**2
*
* THE NEXT 15 LINES COMPUTE THE SOLUTION OF THE 2X2 SYSTEM,
* WHICH CORRESPONDS TO K=2, AND RETAINS ITS TRIANGULAR FORM
* AND THE COEFFICIENTS OF THE ELEMENTARY ROW OPERATIONS IN
* THE ARRAY H(.,.). THE MODIFIED FORM OF THE RIGHT HAND VECTOR
* APPEARS IN THE ARRAY H1(.).
*
      H(1,1)=U=1./(1.-P(1,1))
      H(1,2)=-P(1,2)*U
      H1(1)=P1(1)*U
      U1=1.-P(2,2)-U*P(1,2)*P(2,1)
      H(2,2)=U1=1./U1
      H(2,1)=U*U1*P(2,1)
```

```
      H1(2)=(P1(2)+P(2,1)*H1(1))*U1
      X(2)=H1(2)
      X(1)=H1(1)-H(1,2)*X(2)
      NM=MIN0(2,N)
      U=0.0
      DO 12 I=1,NM
   12 U=U+A(I)*X(I)
*
* G(1) AND G(2) ARE THE FIRST TWO TERMS OF THE DISTRIBUTION OF
* THE MAXIMUM GENERATION SIZE BEFORE EXTINCTION.
*
      G(2)=U
      G(1)=U0*A(1)/(1.-P(1,1))
      DO 13 KR=3,100
      KR1=KR-1
*
* THE SUBROUTINE FILL COMPUTES THE (KR+1)-ST ROW AND COLUMN OF
* THE TRANSITION PROBABILITY MATRIX. FOR FURTHER DETAILS SEE
* COMMENTS GIVEN IN THE SUBROUTINE LISTING ITSELF
*
      CALL FILL(KR1)
*
* THE SUBROUTINE TRIANG PERFORMS GAUSS ELIMINATION TO SOLVE THE
* SYSTEM OF LINEAR EQUATIONS CORRESPONDING TO K=KR. IT
* ALSO UPDATES ALL THE INFORMATION STORED IN THE H(.,.) AND H1(.)
* ARRAYS.
*
      CALL TRIANG(KR)
      IF(NCRIT.EQ.1) GO TO 15
      NM=MIN0(KR,N)
      U=0.0
      DO 14 I=1,NM
   14 U=U+A(I)*X(I)
*
* G(KR) IS THE TERM WITH INDEX KR OF THE DISTRIBUTION OF THE
* MAXIMUM GENERATION SIZE BEFORE EXTINCTION.
*
      G(KR)=U
      IF(U.GT.TEST1) GO TO 15
   13 CONTINUE
*
* PRINT-OUT OF THE DISTRIBTION OF THE MAXIMUM GENERATION SIZE
*
   15 PRINT 1007
      PRINT 1004,(I,G(I),I=1,KR)
*
* COMPUTATION AND PRINT-OUT OF THE DENSITY OF THE DISTRIBUTION
* OF THE MAXIMUM GENERATION SIZE
*
      DO 16 I=2,KR
      I1=KR-I+2
   16 G(I1)=G(I1)-G(I1-1)
      PRINT 1008
      PRINT 1004,(I,G(I),I=1,KR)
*
* COMPUTATION AND PRINT-OUT OF THE DISTRIBUTION OF THE TIME
* TILL EXTINCTION, BY MEANS OF THE SUBROUTINE ABSORB
*
      CALL ABSORB(RHO,NV)
      PRINT 1014
      PRINT 1004,(I,C(I),I=1,NV)
```

```
      PRINT 1015
      PRINT 1004,N0,P1(1),(I,P(1,I),I=1,KR)
* PRINT-OUT OF THE DENSITY OF THE NUMBER OF OFFSPRING OF A
* SINGLE INDIVIDUAL. SINCE THIS DENSITY IS ALREADY COMPUTED
* BY THE REPEATED EARLIER CALLS ON THE SUBROUTINE TERM, NO
* FURTHER COMPUTATION IS REQUIRED HERE.
*
      IF(NCRIT.EQ.1) CALL EXIT
      IF(KR.LT.100) GO TO 17
      PRINT 1013
      CALL EXIT
   17 CALL SECOND(T2)
      U=T2-T1
      PRINT 1012,U
      GO TO 30
   31 CALL EXIT
   33 PRINT 1010
      CALL EXIT
*
* THE FORMAT STATEMENTS
*
 1001 FORMAT(2I3)
 1002 FORMAT(5F7.5)
 1003 FORMAT(*1*//3X*THE DISTRIBUTION OF THE MAXIMUM OF A *
     1*GALTON-WATSON PROCESS BEFORE EXTINCTION*//3X*THE *
     1*NUMBER OF OFFSPRING OF EACH INDIVIDUAL HAS A *
     1*PROBABILITY DENSITY OF PHASE TYPE*//3X*THE INITIAL *
     1*POPULATION SIZE HAS THE DENSITY A(1),...A(N) GIVEN BY*//)
 1004 FORMAT(3X,8(I4,F8.5))
 1005 FORMAT(///3X*THE INITIAL PROBABILITY DENSITY B0,B(1),*
     1*...,B(M)*//)
 1006 FORMAT(///3X*THE PHASE MATRIX*//)
 1007 FORMAT(*1*//3X*THE DISTRIBUTION OF THE MAXIMUM *
     1*GENERATION AMONG THOSE G-W PROCESSES WHICH BECOME *
     1/3X*EXTINCT EVENTUALLY*//)
 1008 FORMAT(*2*/3X*THE DENSITY OF THE MAXIMUM GENERATION SIZE*//)
 1009 FORMAT(//3X*THE PROBABILITY OF EVENTUAL EXTINCTION =*F9.5)
 1010 FORMAT(///3X*INPUT ERROR*)
 1011 FORMAT(//3X*THE MEAN NUMBER OF OFFSPRING PER INDIVIDUAL =*
     1F7.4)
 1012 FORMAT(//3X*THE TOTAL COMPUTATION TIME FOR THIS CASE IS*
     1F9.4)
 1013 FORMAT(*2*3X*THE CAPACITY OF THE STORAGE ARRAYS IS *
     1*EXHAUSTED*)
 1014 FORMAT(*1*//3X*THE DISTRIBUTION OF THE NUMBER OF *
     1*GENERATIONS TILL EXTINCTION*/3X*WITH THE GIVEN *
     1*INITIAL CONDITIONS A(1),...,A(N)*//)
 1015 FORMAT(*2*//3X*THE DENSITY OF THE NUMBER OF OFFSPRING *
     1*OF EACH INDIVIDUAL*//)
      END
```

```
      SUBROUTINE TERM(XX)
*
* THE SUBROUTINE TERM EVALUATES THE NEXT TERM OF THE DENSITY P
* OF THE NUMBER OF OFFSPRING OF A SINGLE INDIVIDUAL. THE COMPUTATION
* IS BY MEANS OF THE STANDARD RECURRENCE RELATIONS FOR A DENSITY
* OF PHASE TYPE.
*
      COMMON/BL/P1(100),P(100,100),A(10),B0,B(20),T0(20),
     1T(20,20),TT(20,20),E(20,20),F(20,20),X(100),Y(20),Z(20),
     1G(100),C(100),H(100,100),H1(100),M,N,TEST,TEST1,TEST2,MAXITER,NCRIT
      DO 1 J=1,M
      U=0.0
      DO 2 I=1,M
    2 U=U+Y(I)*T(I,J)
    1 Z(J)=U
      U=0.0
      DO 3 I=1,M
      U=U+Z(I)*T0(I)
    3 Y(I)=Z(I)
      XX=U
      RETURN
      END
```

```
      SUBROUTINE NEWTON(RHO)
*
* THE SUBROUTINE NEWTON COMPUTES THE PROBABILITY OF EXTINCTION
* FOR A SUPER-CRITICAL GALTON-WATSON PROCESS WITH A SINGLE
* PROGENITOR. NOTE THE EFFICIENT PLANNING OF THE COMPUTATION
* WHICH IS BASED ON THE REPRESENTATION OF THE PROBABILITY
* GENERATING FUNCTION FOR DENSITIES OF PHASE TYPE.
      COMMON/BL/P1(100),P(100,100),A(10),B0,B(20),T0(20),
     1T(20,20),TT(20,20),E(20,20),F(20,20),X(100),Y(20),Z(20),
     1G(100),C(100),H(100,100),H1(100),M,N,TEST,TEST1,TEST2,MAXITER,NCRIT
      DO 4 I=1,M
    4 E(I,I)=1.
      Z1=1.
      DO 1 I=1,M
    1 Z1=Z1-B(I)*T0(I)
      Z1=B0/Z1
      DO 2 NN=1,MAXITER
      DO 3 I=1,M
      DO 3 J=1,M
    3 TT(I,J)=-Z1*T(I,J)
      DO 5 I=1,M
    5 TT(I,I)=1.+TT(I,I)
      CALL LINEQ1(TT,E,F,20,M,M,LL)
      IF(LL.EQ.0) GO TO 6
      GO TO 33
    6 DO 7 J=1,M
      U=0.0
      DO 8 I=1,M
    8 U=U+B(I)*F(I,J)
    7 X(J)=U
      DO 9 J=1,M
      U=0.0
      DO 10 I=1,M
   10 U=U+X(I)*F(I,J)
    9 Z(J)=U
      U1=1.
      DO 11 I=1,M
   11 U1=U1-Z(I)*T0(I)
      DO 12 J=1,M
      U=0.0
      DO 13 I=1,M
   13 U=U+Z(I)*T(I,J)
   12 X(J)=U
      U=0.0
      DO 14 I=1,M
   14 U=U+X(I)*T0(I)
      U=B0-U*Z1**2
      Z2=U/U1
      IF(ABS(Z2-Z1).LT.TEST2) GO TO 15
    2 Z1=Z2
      PRINT 1000,MAXITER
      NCRIT=1
      RETURN
   15 RHO=Z2
      RETURN
   33 PRINT 1001
      NCRIT=1
 1000 FORMAT(//3X,I3,*ITERATIONS ARE NOT SUFFICIENT TO MEET*
     1* THE CONVERGENCE CRITERION*)
 1001 FORMAT(//3X*SINGULAR MATRIX ENCOUNTERED*)
      END
```

```
      SUBROUTINE FILL(K)
*
* THE SUBROUTINE FILL COMPUTES THE (K+1)-ST COLUMN AND THE
* (K+1)-ST ROW OF THE BASIC COEFFICIENT MATRIX
*
      COMMON/BL/P1(100),P(100,100),A(10),B0,B(20),TC(20),
     1T(20,20),TT(20,20),E(20,20),F(20,20),X(100),Y(20),Z(20),
     1G(100),C(100),H(100,100),H1(100),M,N,TEST,TEST1,TEST2,MAXITER,NCRIT
      K1=K+1
      P1(K1)=P1(K)*B0
      CALL TERM(XX)
      P(1,K1)=XX
*
* COMPUTATION BY MEANS OF THE DEFINITION OF THE CONVOLUTION
* PRODUCT OF THE ENTRIES P(I,K+1) FOR I=1,2,...,K
*
      DO 1 I=2,K
      I1=I-1
      U=P1(I1)*P(1,K1)+B0*P(I1,K1)
      DO 2 NU=1,K
    2 U=U+P(1,NU)*P(I1,K1-NU)
    1 P(I,K1)=U
*
* COMPUTATION BY MEANS OF THE DEFINITION OF THE CONVOLUTION PRODUCT
* OF THE ENTRIES P(K+1,J), FOR J=1,2,...,K+1
*
      P(K1,1)=B0*P(K,1)+P(1,1)*P1(K)
      DO 3 J=2,K1
      J1=J-1
      U=P(1,J)*P1(K)+B0*P(K,J)
      DO 4 NU=1,J1
    4 U=U+P(1,NU)*P(K,J-NU)
    3 P(K1,J)=U
      RETURN
      END
```

```
      SUBROUTINE TRIANG(K)
*
* THE SUBROUTINE TRIANG PERFORMS THE ADDITIONAL ROW OPERATIONS
* NECESSARY TO SOLVE THE LINEAR SYSTEM, CORRESPONDING TO THE
* INDEX K, ARISING IN THE COMPUTATION OF THE DISTRIBUTION OF THE
* MAXIMUM GENERATION SIZE BEFORE EXTINCTION.
*
      COMMON/BL/P1(100),P(100,100),A(10),B0,B(20),T0(20),
     1T(20,20),TT(20,20),E(20,20),F(20,20),X(100),Y(20),Z(20),
     1G(100),C(100),H(100,100),H1(100),M,N,TEST,TEST1,TEST2,MAXITER,NCRIT
      K1=K-1
*
* THE NEW ROW TO BE REDUCED IS STORED IN SCRATCH ARRAYS
*
      DO 1 J=1,K1
      X(J)=-P(J,K)
      C(J)=-P(K,J)
      U=0.0
*
* THE DO-2 LOOP COMPUTES THE K-TH COLUMN OF THE H(.,.) ARRAY
*
      DO 2 I=1,J
    2 U=U+H(J,I)*X(I)
    1 H(J,K)=U
      DO 9 I=1,K1
    9 X(I)=0.0
      C(K)=1.-P(K,K)
      H1(K)=P1(K)
*
* THE DO-LOOPS 3-6 CARRY OUT THE ELEMENTARY ROW OPERATIONS
*
      DO 3 I=1,K1
      U=C(I)
      DO 4 J=1,I
    4 X(J)=X(J)-U*H(I,J)
      I1=I+1
      DO 5 J=I1,K
    5 C(J)=C(J)-U*H(I,J)
    3 H1(K)=H1(K)-U*H1(I)
      U=C(K)
      H(K,K)=1./U
      H1(K)=H1(K)/U
      DO 6 I=1,K1
    6 H(K,I)=X(I)/U
*
* COMPUTATION OF THE SOLUTION TO THE TRIANGULAR SYTEM OF
* EQUATIONS
*
      X(K)=H1(K)
      DO 7 I1=1,K1
      I=K-I1
      U=H1(I)
      DO 8 I2=1,I1
    8 U=U-H(I,I+I2)*X(I+I2)
    7 X(I)=U
      RETURN
      END
```

```
      SUBROUTINE ABSORB(RHO,NV)
*
* THE SUBROUTINE ABSORB COMPUTES THE DISTRIBUTION OF THE
* NUMBER OF GENERATIONS TILL EXTINCTION. FOR SUPER-CRITICAL
* GALTON-WATSON PROCESSES WE OBTAIN THE PROBABILITY THAT THE
* PROCESS BECOMES EXTINCT AFTER J GENERATIONS. THE DISTRIBUTION
* OF THE TIME TILL EXTINCTION IS DEFECTIVE IN THIS CASE.
*
      COMMON/BL/P1(100),P(100,100),A(10),B0,B(20),T0(20),
     1T(20,20),TT(20,20),E(20,20),F(20,20),X(100),Y(20),Z(20),
     1G(100),C(100),H(100,100),H1(100),M,N,TEST,TEST1,TEST2,MAXITER,NCRIT
      ZX=B0
      DO 1 NU=2,100
      DO 2 I=1,M
      DO 3 J=1,M
    3 TT(I,J)=-ZX*T(J,I)
    2 TT(I,I)=1.+TT(I,I)
      CALL LINEQ1(TT,B,Y,20,M,1,LL)
      IF(LL.EQ.0) GO TO 4
      NV=NU-1
      GO TO 6
    4 U=0.0
      DO 5 I=1,M
    5 U=U+Y(I)*T0(I)
      ZX=B0+ZX*U
      U=0.0
      DO 8 I=1,N
    8 U=U+A(I)*(ZX**I)
      C(NU)=U
      IF(ABS(RHO-ZX).LT.TEST2) GO TO 7
    1 CONTINUE
    7 NV=NU
    6 U=0.0
      DO 9 I=1,N
    9 U=U+A(I)*B0**I
      C(1)=U
      RETURN
      END
*EOR
```

```
THE DISTRIBUTION OF THE MAXIMUM OF A GALTON-WATSON PROCESS BEFORE EXTINCTION

THE NUMBER OF OFFSPRING OF EACH INDIVIDUAL HAS A PROBABILITY DENSITY OF PHASE TYPE

THE INITIAL POPULATION SIZE HAS THE DENSITY A(1),...A(N) GIVEN BY

   1  .25000   2  .75000


THE INITIAL PROBABILITY DENSITY B0,B(1),...,B(M)

   0  .45000   1  .25000   2  .30000


THE PHASE MATRIX

   0  .60000   1  .10000   2  .30000
   0  .70000   1  .15000   2  .15000

THE MEAN NUMBER OF OFFSPRING PER INDIVIDUAL =  .8368

THE PROBABILITY OF EVENTUAL EXTINCTION =  1.00000
```

THE DISTRIBUTION OF THE MAXIMUM GENERATION AMONG THOSE G-W PROCESSES WHICH BECOME EXTINCT EVENTUALLY

1	.17578	2	.67190	3	.78353	4	.85401	5	.89927	6	.92916	7	.94945	8	.96353
9	.97349	10	.98062	11	.98577	12	.98951	13	.99225	14	.99427	15	.99575	16	.99685
17	.99766	18	.99826	19	.99870	20	.99903	21	.99928	22	.99946	23	.99960		

THE DENSITY OF THE MAXIMUM GENERATION SIZE

1	.17578	2	.49612	3	.11163	4	.07048	5	.04526	6	.02989	7	.02029	8	.01409
9	.00996	10	.00713	11	.00515	12	.00375	13	.00274	14	.00201	15	.00148	16	.00110
17	.00081	18	.00060	19	.00045	20	.00033	21	.00025	22	.00018	23	.00014		

THE DISTRIBUTION OF THE NUMBER OF GENERATIONS TILL EXTINCTION
WITH THE GIVEN INITIAL CONDITIONS A(1),...,A(N)

1	.26437	2	.46970	3	.60580	4	.69913	5	.76573	6	.81482	7	.85195	8	.88058
9	.90301	10	.92079	11	.93502	12	.94651	13	.95584	14	.96346	15	.96971	16	.97484
17	.97908	18	.98259	19	.98550	20	.98791	21	.98991	22	.99158	23	.99297	24	.99413
25	.99509	26	.99590	27	.99657	28	.99713	29	.99760	30	.99800	31	.99832	32	.99860
33	.99883	34	.99902	35	.99918	36	.99931	37	.99943	38	.99952	39	.99960	40	.99966
41	.99972	42	.99976	43	.99980	44	.99984	45	.99986	46	.99988	47	.99990	48	.99992

THE DENSITY OF THE NUMBER OF OFFSPRING OF EACH INDIVIDUAL

0	.45000	1	.36000	2	.12600	3	.04230	4	.01436	5	.00486	6	.00165	7	.00056
8	.00019	9	.00006	10	.00002	11	.00001	12	.00000	13	.00000	14	.00000	15	.00000
16	.00000	17	.00000	18	.00000	19	.00000	20	.00000	21	.00000	22	.00000	23	.00000

THE TOTAL COMPUTATION TIME FOR THIS CASE IS .7390

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Computational Problems related to the Galton-Watson Process		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER Mimeo Series #402
7. AUTHOR(s) Marcel F. Neuts		8. CONTRACT OR GRANT NUMBER(s) AFOSR-72-2350 B
9. PERFORMING ORGANIZATION NAME AND ADDRESS Purdue Research Foundation West Lafayette, IN 47907		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Department of Statistics, Purdue University West Lafayette, IN 47907		12. REPORT DATE February, 1975
		13. NUMBER OF PAGES 53
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Computational probability
Distributions of phase type
Galton-Watson process
M/G/1 queue

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An algorithm to compute the probability distributions of the successive generation sizes in a Galton-Watson process is presented. The distribution of the number of offspring of each individual is assumed to be of phase type. A probability distribution is of phase type if it can be identified as the distribution of the time till absorption in an absorbing finite Markov

chain with appropriate initial conditions. A detailed analysis of the error due to truncation is given as well as an application in a problem related to the M/G/1 queue.

A second algorithm deals with the distribution of the maximum generation size before extinction. Several theorems on probability distributions of phase type are proved.