AGE DISTRIBUTIONS FOR MULTIPHASE BRANCHING PROCESSES AND THEIR APPLICATIONS

Thesis for the Degree of Ph. D.
MICHIGAN STATE UNIVERSITY
WEN-HOU KUO
1973

This is to certify that the

thesis entitled

Age Distributions for Multiphase Branching
Processes and Their Applications

presented by

Wen-Hou Kuo

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Statistics

P. J. Brockwell
Major professor

Date 2/19/73

O-7639

ABSTRACT

AGE DISTRIBUTIONS FOR MULTIPHASE BRANCHING PROCESSES AND THEIR APPLICATIONS

By

Wen-Hou Kuo

A supercritical age-dependent branching process is considered in which the lifespan of each individual is composed of four phases whose durations have joint probability density $f(t_1,t_2,t_3,t_4)$. Starting with a single individual of age zero at time zero we consider the asymptotic behavior as $t \to \infty$ of the random variable $Z^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t)$ defined as the number of individuals in phase 4 at time t for which the elapsed phase durations $Y_{01},\ldots,Y_{04}$, $\ldots$, $Y_{i1},\ldots,Y_{i4},\ldots,Y_{n4}$ of the individual itself and its first n ancestors satisfy the inequalities $Y_{ij} \le a_{ij}$, $i = 0,\ldots,n$, $j = 1,\ldots,4$. We also state an analogous result which defines (conditional on eventual non-extinction of the population) the asymptotic joint distribution of the phase and elapsed phase durations of an individual drawn at random from the population and the phase durations of its ancestors.

The correlations between the lifespans of parent and daughters are considered in a more general model in which we treat the population as a multiphase, multitype branching process. Moreover, the possibility of random cell removal is also taken into account.

The application of the results to the analysis of cell labelling experiments is described. Finally we include some numerical results.

AGE DISTRIBUTIONS FOR MULTIPHASE BRANCHING PROCESSES AND THEIR APPLICATIONS

By

Wen-Hou Kuo

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Statistics and Probability

1973

G80297

TO MY PARENTS

ACKNOWLEDGEMENTS

I wish to express my thanks to Professor P.J. Brockwell for suggesting the problem and for his patient guidance before and during the preparation of this thesis. His comments and suggestions led the way to theorems and simplified proofs.

I am also grateful to the Department of Statistics and Probability, Michigan State University for financial support during my graduate studies.

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The life-cycle of many biological cells can be divided into four phases numbered 1,2,3,4. During the second phase the cell is synthesizing DNA and during the fourth phase the cell is in the process of dividing (cells in phase 4 are said to be mitotic and can be distinguished from cells in phases 1,2 and 3). The phases 1,2,3,4 are often referred to as G_1, S, G_3 and M respectively. In this paper we shall consider a population of cells each having a 4-phase life-cycle as described above. It will be assumed that the joint distribution of the four phase durations T_1, T_2, T_3 and T_4 for any new-born cell is specified by a probability density

$$f(t_1,t_2,t_3,t_4) = f(\underset{\sim}{t}) \ .$$

The distribution of total lifetime for any cell thus has the probability density

$$g(u) = \underset{t_1+t_2+t_3 \le u}{\int\int\int} f(t_1,t_2,t_3,u-t_1-t_2-t_3)dt_1dt_2dt_3 \ .$$

We shall suppose moreover that the process of cell multiplication is a supercritical age-dependent branching process (see Harris [7], Chapter 6) in which the lifetime distribution of any individual is specified by the probability density g and the distribution of the number of offspring per parent has probability generating function,

$$h(z) = \sum_{k=0}^{\infty} p_k z^k ,$$

where the factorial moments $h'(1)$ and $h''(1)$ are assumed to satisfy $1 < h'(1) < \infty$ and $h''(1) < \infty$ respectively. We shall assume also that the density g satisfies the condition $\int [g(y)]^p dy < \infty$ for some p such that $p > 1$. (The latter condition is not used until the proof of Theorem 4.2 of Chapter II.

The analysis of the model we have just described can easily be modified to allow for the removal of cells from the population if the removal process is as follows: for any cell, independently of its age or phase and independently of other cells, the probability of removal in any small time interval $(t, t+\delta t)$ is $\lambda\delta t + o(\delta t)$. The generalization to this case of the results derived in Chapter II are indicated in Section 6.

A powerful tool in the experimental study of the cell cycle is the class of so-called "labelling" experiments, in which a radioactive substance, e.g. tritiated thymidine, is injected into the population. If the injection is a pulse administered at epoch t_0 then the effect is to label all cells which at epoch t_0 are in phase 2 and no others. Subsequently when the labelled cells divide they pass the label on to their offspring. After injection of the pulse at t_0 we observe the cells in phase 4 and count the fraction at epoch $t_0 + t$ which are labelled. The resulting function $P_{t_0}(t)$, $t \geq 0$, is known as the FLM- (fraction-labelled-mitoses-) function. If instead of using a pulse at epoch t_0 we inject continuously from epoch t_0 onwards and again count the fraction of labelled cells in phase 4 at time $t_0 + t$ which are labelled, we obtain a function $C_{t_0}(t)$, $t \geq 0$, known as the continuous-labelling-function. For large t_0 the function P_{t_0} has been studied in the case of independent phase durations by Barrett [1], Trucco and Brockwell [17], and Takahashi [16] and for a more general model by Macdonald [9], who also considered the function C_{t_0}. These analyses all assume that the population is large, that a limiting age distribution has been attained by epoch t_0 and that the random variables involved may be replaced by their expected values. Note that in practice we do not observe the fraction of <u>all</u> cells in mitosis at time $t_0 + t$ which are labelled but only a sample of N mitotic cells. The number of labelled cells observed in such a sample will be (under our model) binomially distributed with parameters N and $P_{t_0}(t)$ or $C_{t_0}(t)$.

In Chapter II we show how the asymptotic forms P and C of the functions P_{t_0} and C_{t_0} as $t_0 \to \infty$ can be derived for the model defined in the first paragraph. The functions P and C are obtained from the asymptotic joint distributions of the phase durations of a cell selected at random in phase 4 and its first n ancestors. (The cell itself is said to be its own zero^{th} ancestor, its parent is its 1^{st} ancestor, etc.) For a population which starts at epoch zero with one member of age zero we establish the (almost sure) existence of these asymptotic joint distributions given that the population does not die out and determine them explicitly. They generalize the so-called "carrier-distributions" of E.O. Powell [14] (see also Brockwell and Trucco [4], p. 172 for a heuristic derivation of the corresponding results for the model with independent phase durations).

The purpose of deriving such theoretical expressions for P and C is to use them to estimate parameters of the model from experimental data. We do not consider this question here but refer the reader to Barrett [2], Brockwell, Trucco and Fry [5], Macdonald [9], Steel and Hanes [15] and Mendelsohn and Takahashi [10] who deal with the estimation problem for a variety of particular models.

There is experimental evidence to suggest that in many populations correlations exist between the lifespans of parent and daughter cells. In order to take this into account we consider in Chapter III a more general model in which we treat the population as a multiphase, multi-type branching process. We assume that there are m types of individual, the joint phase distribution for the i^{th} type being specified by a joint density $f_i(t_1,t_2,t_3,t_4)$. The expected number of type j offspring from a type i parent is m_{ij} and the second moment of the total number of offspring per parent is assumed to be finite.

CHAPTER II

AGE DISTRIBUTIONS AND LABELLING FUNCTIONS FOR A SINGLE-TYPE MULTIPHASE BRANCHING PROCESS

§1. Notation

The model to be considered here (a single-type multiphase supercritical branching process with correlated phase durations) was defined in Chapter I. The following notation will be required in addition to that already introduced.

Consider the population at time t descended from a single ancestor of age zero at epoch zero. Define $Z_k^{(4)}(\underset{\sim}{a}_0,t)$ to be the number of cells in phase 4 at time t which are k^{th} generation descendants of the initial cell and for which the times $Y_{01},Y_{02},Y_{03},Y_{04}$ spent (prior to epoch t) in phases 1,2,3,4 satisfy

$$(1.1) \qquad Y_{0i} \le a_{0i} \text{ , } i = 1,\dots,4 \text{ .}$$

(We shall consistently use the vector notation $\underset{\sim}{Y}_0 \le \underset{\sim}{a}_0$ as an abbreviation for the set of inequalities $Y_{0i} \le a_{0i}$ between corresponding components of $\underset{\sim}{Y}_0$ and $\underset{\sim}{a}_0$. Note also that the 0^{th} generation descendant of a cell means the cell itself, a first generation descendant is a daughter cell, etc.) More generally we define $Z_k^{(4)}(\underset{\sim}{a}_0,\underset{\sim}{a}_1,\dots,\underset{\sim}{a}_n,t)$, $k \ge n$, to be the number of cells in phase 4 at time t which are k^{th} generation descendants of the initial cell and whose i^{th} ancestor, $i = 0,1,\dots,n$ has spent times $Y_{i1},\dots,Y_{i4}$ in phases $1,\dots,4$ satisfying $\underset{\sim}{Y}_i \le \underset{\sim}{a}_i$. The quantities

$Z_k^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t)$ are random variables defined on a probability space of family trees (whose construction is essentially the same as that in Harris [6], Chapter 6). We shall also make use of the random variables $Z(t)$, the total population size at time t, $Z^{(4)}(t)$, the number of cells in phase 4 at time t, and

$$Z^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t) = \sum_{k=n}^{\infty} Z_k^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t) \ .$$

Our primary interest is in the ratio $Z^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t)/Z^{(4)}(t)$, particularly in its asymptotic behaviour as $t \to \infty$ (given that $Z(t) \not\to 0$) which leads to the asymptotic joint distributions of the phase durations of a cell selected at random from those in mitosis and its first n ancestors. (Clearly the analysis which follows could easily be modified to treat the case of a cell selected at random from those in any particular phase but bearing in mind the application to labelling experiments we shall consider only cells selected from phase 4. A corresponding result for sampling from an arbitrary phase is given in Section 5.)

We shall use M and S with appropriate subscripts and superscripts to denote first and second moments of the corresponding random variables, e.g.

$$M_k^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t) = E\, Z_k^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t),$$

$$S_{k\ell}^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t,\tau) = E[Z_k^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t)Z_\ell^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t+\tau)],$$

$$M(t) = E\, Z(t),\ S(t,\tau) = E[Z(t)Z(t+\tau)],$$

$$M^{(4)}(t) = E\, Z^{(4)}(t),\ S^{(4)}(t,\tau) = E[Z^{(4)}(t)Z^{(4)}(t+\tau)], \text{ etc.}$$

These moments are known to be finite for all finite t, τ because of the finiteness of $M(t)$ and $S(t,\tau)$ (see Harris [7]).

§2. Asymptotic Behaviour of $M^{(4)}(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_n,t)$

The following decomposition,

$$(2.1)\qquad M^{(4)}(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_n,t) = M_n^{(4)}(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_n,t) + \sum_{k=n+1}^{\infty} M_k^{(4)}(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_n,t),$$

leads at once to a renewal equation for $M^{(4)}$,

$$(2.2)\qquad M^{(4)}(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_n,t) = M_n^{(4)}(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_n,t) + A\int_0^t M^{(4)}(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_n,t-u)g(u)\,du,$$

where $A = h'(1) > 1$, g is the probability density of the total lifetime of a newborn cell, and $M_n^{(4)}$ is easily shown (by induction) to be

$$(2.3)\qquad M_n^{(4)}(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_n,t) =$$

$$\underset{\substack{\underset{\sim}{y}_i\le\underset{\sim}{a}_i,\ i=1,\dots,n,\\ y_{0j}\le a_{0j},\ j=1,\dots,3}}{\int\cdots\int} A^n f(\underset{\sim}{y}_0)\cdots f(\underset{\sim}{y}_n)\, I_{[0,y_{04}\wedge a_{04})}(t-y_{03}-y_{02}-\dots-y_{n1})\; d\underset{\sim}{y}_0\cdots d\underset{\sim}{y}_n .$$

Applying the standard asymptotic theory of the renewal equation to equation (2.2) (see Feller [6], p. 468 or Harris [7], p. 161) we find that as $t\to\infty$

$$(2.4)\qquad M^{(4)}(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_n,t) \sim K^{(4)}e^{ct}\prod_{i=0}^{n}\Psi_i(\underset{\sim}{a}_i)$$

where

$$(2.5)\qquad \Psi_i(\underset{\sim}{a}_i) = \int_{\underset{\sim}{y}\le\underset{\sim}{a}_i}\psi_i(\underset{\sim}{y})\,d\underset{\sim}{y}\Big/\int\psi_i(\underset{\sim}{y})\,d\underset{\sim}{y}\ ,$$

$$(2.6)\qquad \psi_0(\underset{\sim}{y}) = e^{-c(y_1+\dots+y_4)}\int_{y_4}^{\infty} f(y_1,y_2,y_3,u)\,du,$$

$$(2.7)\qquad \psi_i(\underset{\sim}{y}) = e^{-c(y_1+\dots+y_4)}f(\underset{\sim}{y}),\ i = 1,2,\dots,$$

$$(2.8)\qquad K^{(4)} = \frac{\int e^{-c(y_1+\dots+y_4)}\left[\int_{y_4}^{\infty} f(y_1,y_2,y_3,u)\,du\right] d\underset{\sim}{y}}{A\int y\, e^{-cy} g(y)\,dy}$$

and c is the unique positive root of the equation

$$(2.9)\qquad A\int e^{-cu} g(u)\,du = 1\ .$$

It is easy to show, using precisely the same argument (applied to $M^{(4)}(\underset{\sim}{a}_0,t)$ with all $a_{0j} = \infty$), that the expected number of cells in phase 4 at time t is given asymptotically, as $t \to \infty$, by

$$(2.10)\qquad M^{(4)}(t) \sim K^{(4)} e^{ct}$$

with $K^{(4)}$ and c defined as above. Equations (2.9) and (2.10) immediately suggest the following. If a cell is selected at random from those in phase 4 at time t then the joint distribution of the times Y_{ij}, $j = 1,\dots,4$, $i = 0,\dots,n$ spent by the i^{th} ancestor of the cell in phase j prior to time t has the following asymptotic form as $t \to \infty$:

$$(2.11)\qquad F_n(\underset{\sim}{y}_0,\dots,\underset{\sim}{y}_n) = P[\underset{\sim}{Y}_0 \le \underset{\sim}{y}_0,\dots,\underset{\sim}{Y}_n \le \underset{\sim}{y}_n] = \prod_{i=0}^{n} \Psi_i(\underset{\sim}{y}_i)$$

where the distribution functions $\Psi_i(\underset{\sim}{y}_i)$ are defined by equations (2.5)-(2.7). (Note that the functions Ψ_i, $i = 1,2,\dots$, are identical.)

We shall refer to the function F_n defined by (2.11) as the asymptotic <u>n-fold age distribution</u> of cells in phase 4. It is reasonable to suppose that conditional upon non-extinction of the population; the <u>actual fraction</u> of cells in phase 4 satisfying $\underset{\sim}{Y}_0 \le \underset{\sim}{a}_0,\dots,\underset{\sim}{Y}_n \le \underset{\sim}{a}_n$ converges to F_n for each n. The purpose of the following sections is to show that this is indeed true in the following sense: First we recall (Harris [7], Chapter 6) that under the conditions specified in Chapter I the <u>total</u>

population size $Z(t)$ has the property that

(2.12) $Z(t)e^{-ct} \to KW$ both in mean square (m.s.) and almost surely (a.s.),

where

$$K = \frac{A - 1}{c\, A^2 \int y\, e^{-cy} g(y)\,dy}$$

and W is a random variable with expectation

$$EW = 1 ,$$

and variance

$$\sigma_W^2 > 0 ;$$

moreover the events $\{Z(t) \to 0\}$ and $\{W = 0\}$ are a.s. equal. We shall show that analogous results hold for the random variables $Z^{(4)}(t)$ and $Z^{(4)}(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_m,t)$; more precisely we shall show that

$$Z^{(4)}(t)e^{-ct} \overset{\text{a.s.}}{\underset{\text{m.s.}}{\to}} K^{(4)}W ,$$

$$Z^{(4)}(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_n,t)e^{-ct} \overset{\text{a.s.}}{\underset{\text{m.s.}}{\to}} K^{(4)}F_n(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_n)W$$

for all $\underset{\sim}{a}_i \geq \underset{\sim}{0}$, $i = 0,1,\dots,n$ and for all $n = 0,1,2,\dots$, and in fact

$$\{Z^{(4)}(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_n,t)e^{-ct} \to K^{(4)}F_n(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_n)W$$
$$\text{for all } \underset{\sim}{a}_i \geq 0,\ i = 0,1,\dots,n \text{ and for all } n = 0,1,2,\dots\}\text{a.s.}$$

The methods used are similar to those in Harris [7], Chapter 6, but the arguments are somewhat complicated by (a) the multiphase structure of the process and (b) the simultaneous consideration of the elapsed phase durations of cells and their ancestors.

§3. Mean Square Convergence of $Z^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t)e^{-ct}$.

It is convenient to decompose the second moment $S^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t,\tau) = E[Z^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t)Z^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t+\tau)]$ in the following manner:

$$(3.1) \quad S^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t,\tau) = Q(t,\tau) + \sum_{i=n+1}^{\infty}\sum_{j=n+1}^{\infty} S_{ij}^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t,\tau),$$

where $S_{ij}^{(4)}$ was defined in Section 2 and

$$(3.2) \quad 0 \le Q(t,\tau) \le E[Z_n(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t)Z(t+\tau) + Z(t)Z_n(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t+\tau)] .$$

(Here $Z_n(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t)$ is the <u>total</u> number of n^{th} generation descendants of the initial ancestor in the population at time t.)

From equation (3.1) we obtain the renewal equation

$$(3.3) \quad S^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_m,t,\tau) = Q(t,\tau)$$
$$+ h''(1)\int_0^t M^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_m,t-u)M^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_m,t+\tau-u)g(u)\,du$$
$$+ A\int_0^t S^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_m,t-u,\tau)g(u)\,du .$$

Writing $R(t,\tau)$ for the sum of the first two terms on the right side of (3.3) it is not difficult to show that $R(t,\tau)e^{-2ct-c\tau} \to h''(1)[K^{(4)}F_m(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_m)]^2\int_0^\infty e^{-2cy}g(y)\,dy$ uniformly for $\tau \ge 0$ as $t \to \infty$. (This follows from equation (2.4) and Schwarz's inequality applied to the right-hand side of (2.2), noting that $Z_n(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t) \le Z_n(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n, \sum_{i=0}^{n}\sum_{j=1}^{4} a_{ij})$ and that $E[Z(t)^2]e^{-2ct} \to M < \infty$ (Harris [7], p. 144).)

Writing $\bar{S}^{(4)}$ for $S^{(4)}e^{-2ct-c\tau}$ we obtain from (3.3)

$$\bar{S}^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_m,t,\tau) = R(t,\tau)e^{-2ct-c\tau}$$
$$+ \int_0^t \bar{S}^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_m,t-u,\tau)A\,e^{-2cu}g(u)\,du$$

whence, by Lemma 4 of Harris [7], p. 163, we obtain

$$(3.4)\quad \lim_{t\to\infty} \bar{S}^{(4)}(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_m,t,\tau) = \frac{h''(1)[K^{(4)}F_m(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_m)]^2\int_0^\infty e^{-2cy}g(y)\,dy}{1-\int_0^\infty A\,e^{-2cy}g(y)\,dy}$$

uniformly in $\tau \geq 0$.

Starting from a renewal equation similar to (3.3) for $E[Z^{(4)}(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_m,t)Z(t+\tau)] = C(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_m,t,\tau)$ we can proceed in the same way as above to obtain

$$(3.5)\quad \lim_{t\to\infty} C(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_m,t,\tau)e^{-2ct-c\tau} = \frac{h''(1)[K^{(4)}F_m(a_0,\dots,a_m)]K\int_0^\infty e^{-2cy}g(y)\,dy}{1-\int_0^\infty A\,e^{-2cy}g(y)\,dy}$$

uniformly for $\tau \geq 0$.

Equations (3.4) and (3.5) imply that as $t \to \infty$,

$$(3.6)\quad E[(KZ^{(4)}(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_m,t) - K^{(4)}F_m(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_m)Z(t))^2e^{-2ct}] \to 0\ ,$$

from which it follows that

$$(3.7)\quad Z^{(4)}(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_m,t)e^{-ct} \xrightarrow{m.s.} K^{(4)}F_m(\underset{\sim}{a}_0,\dots,\underset{\sim}{a}_m)W\ ,$$

where W is the random variable appearing in (2.12).

§4. The Almost Sure Convergence of $Z^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t)e^{-ct}$

We shall need the following decomposition of the random variables $Z^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t)$:

$$(4.1)\quad Z^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t) = Z^{E}(\underset{\sim}{a}_0',\underset{\sim}{a}_1,\ldots,\underset{\sim}{a}_n,t) - Z^{e}(\underset{\sim}{a}_0',\underset{\sim}{a}_1,\ldots,\underset{\sim}{a}_n,t-a_{04})$$

$$- X(\underset{\sim}{a}_0',\underset{\sim}{a}_1,\ldots,\underset{\sim}{a}_n,t,a_{04}) ,$$

where $\underset{\sim}{a}_0'$ denotes a three-component vector (a_{03},a_{02},a_{01}) and $Z^{E}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t)$, $Z^{e}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t)$ are respectively the numbers of cells entering phase 4 in $[0,t]$, $[0,t)$ with ancestral phase durations $Y_{03},Y_{02},\ldots,Y_{n3},Y_{n2},Y_{n1}$ satisfying $\underset{\sim}{Y}_0' \le \underset{\sim}{a}_0'$, $\underset{\sim}{Y}_1 \le \underset{\sim}{a}_1,\ldots,\underset{\sim}{Y}_n \le \underset{\sim}{a}_n$. Similarly $X(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t,a)$ is the number of cells entering and leaving phase 4 in $[t-a,t]$ with $\underset{\sim}{Y}_0' \le \underset{\sim}{a}_0',\ldots,\underset{\sim}{Y}_n \le \underset{\sim}{a}_n$.

Derivation of the asymptotic behaviour of the first and second moments $M^{E}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t) = E[Z^{E}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t)]$ and $S^{E}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t,\tau) = E[Z^{E}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t)Z^{E}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t+\tau)]$ follows the pattern of Sections 2 and 3 so we shall simply write down the relevant renewal equations and state the final results without proofs.

$$(4.2)\quad M^{E}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t) = M_n^{E}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t) + \int_0^t M^{E}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t-u)g(u)\,du ,$$

where

$$(4.3)\quad M_n^{E}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t) = \underset{\substack{\underset{\sim}{y}_i\le\underset{\sim}{a}_i,\ i=1,\ldots,n\\ \underset{\sim}{y}_0'\le\underset{\sim}{a}_0'}}{\int\cdots\int} A^n f(\underset{\sim}{y}_0)\cdots f(\underset{\sim}{y}_n)I_{[0,\infty)}(t-y_{03}-y_{02}-$$

$$\cdots-y_{n1})\,d\underset{\sim}{y}_0\cdots d\underset{\sim}{y}_n ,$$

and

$$(4.4)\quad S^{E}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t,\tau) = Q^{E}(t,\tau)$$

$$+ h''(1)\int_0^t M^{E}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t-u)M^{E}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t+\tau-u)g(u)\,du$$

$$+ A\int_0^t S^{E}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t-u,\tau)g(u)\,du$$

where

$$(4.5)\qquad 0 \le Q^E(t,\tau) \le E[Z_n^E(\underset{\sim}{a}_0',\dots,\underset{\sim}{a}_n,t)Z^B(t+\tau) + Z^B(t)Z_n^E(\underset{\sim}{a}_0',\dots,\underset{\sim}{a}_n,t+\tau)]$$

and $Z^B(t)$ is the total number of births in $[0,t]$, and $Z_n^E(\underset{\sim}{a}_0',\dots,\underset{\sim}{a}_n,t)$ is the contribution to $Z^E(\underset{\sim}{a}_0',\dots,\underset{\sim}{a}_n,t)$ of n^{th} generation descendants of the initial cell.

<u>Lemma 4.1</u>. Under the conditions specified in Section 1, as $t \to \infty$,

$$(4.6)\qquad M^E(\underset{\sim}{a}_0',\dots,\underset{\sim}{a}_n,t)e^{-ct} = \rho^E(\underset{\sim}{a}_0',\dots,\underset{\sim}{a}_n)[1 + o(1)] ,$$

$$(4.7)\qquad S^E(\underset{\sim}{a}_0',\dots,\underset{\sim}{a}_n,t,\tau)e^{-2ct-c\tau} = \frac{h''(1)\rho^E(\underset{\sim}{a}_0',\dots,\underset{\sim}{a}_n)^2\int e^{-2cu}g(u)\,du}{1 - A\int e^{-2cu}g(u)\,du}[1 + o(1)],$$

and

$$(4.8)\qquad Z^E(\underset{\sim}{a}_0',\dots,\underset{\sim}{a}_n,t)e^{-ct} \xrightarrow{m.s.} \rho^E(\underset{\sim}{a}_0',\dots,\underset{\sim}{a}_n)W ,$$

where $$\rho^E(\underset{\sim}{a}_0',\dots,\underset{\sim}{a}_n) = \frac{\int_{\underset{\sim}{y}_0' \le \underset{\sim}{a}_0'} e^{-c(y_{03}+y_{02}+y_{01})} f(\underset{\sim}{y}_0)\,d\underset{\sim}{y}_0}{cA\int y\, e^{-cy}g(y)\,dy} \prod_{i=1}^{n} \Psi_i(\underset{\sim}{a}_i) .$$

<u>Remark 4.1</u>. If we apply Lemmas 3, 4 and 5 of Harris [7], p. 162, to equations (4.2) and (4.4) we obtain the stronger result that the terms $o(1)$ in (4.6) and (4.7) can be replaced by $o(e^{-\epsilon t})$ where ϵ is independent of t and τ and $\epsilon > 0$. (It is here that we use the condition $\int(g(t))^p dt < \infty$ for some $p > 1$.) This fact is crucial in the proof of almost sure convergence which follows. The argument follows that of Harris [7], Theorem 6.21.1.

<u>Theorem 4.1</u>. As $t \to \infty$, $Z^E(\underset{\sim}{a}_0',\dots,\underset{\sim}{a}_n,t)e^{-ct} \xrightarrow{a.s.} \rho^E(\underset{\sim}{a}_0',\dots,\underset{\sim}{a}_n)W$ for all $\underset{\sim}{a}_0' \ge 0$, $\underset{\sim}{a}_i \ge 0$, $i = 1,\dots,n$ and for all $n = 0,1,2,\dots$.

<u>Proof</u>. For fixed $\underset{\sim}{a}_0',\dots,\underset{\sim}{a}_n$ let $Y_t = Z^E(\underset{\sim}{a}_0',\dots,\underset{\sim}{a}_n,t)e^{-ct}$ and $Y = \rho^E(a_0',\dots,a_n)W$. Then from Lemma 4.1 and Remark 4.1 it follows that

$\int_0^\infty E(Y_t - Y)^2 dt < \infty$ and hence

$$(4.9) \qquad \int_0^\infty (Y_t - Y)^2 dt < \infty \text{ a.s.}$$

Since $Z^E(\underset{\sim}{a}_0', \ldots, \underset{\sim}{a}_n, t)$ is non-decreasing in t it follows that,

$$(4.10) \qquad Y_{t+\tau} \geq e^{-\alpha\tau} Y_t \; .$$

Denote by $(\Omega, \mathcal{F}, P)$ the probability space of family trees on which the random variables are defined and let ω denote a point of Ω. Then if $\overline{\lim_{t\to\infty}} Y_t(\omega) > Y(\omega)$ it follows from (4.10) that

$$(4.11) \qquad \int (Y_t(\omega) - Y(\omega))^2 dt = \infty \; .$$

Similarly (4.11) holds if $\underline{\lim}_{t\to\infty} Y_t(\omega) < Y(\omega)$. Hence by (4.9) the set $\{\omega \in \Omega : \overline{\lim}\, Y_t(\omega) = \underline{\lim}\, Y_t(\omega) = Y(\omega)\}$ has probability 1.

Remark 4.2. $Z^e(\underset{\sim}{a}_0', \ldots, \underset{\sim}{a}_n, t)e^{-ct} \overset{a.s.}{\rightarrow} \rho^E(\underset{\sim}{a}_0', \ldots, \underset{\sim}{a}_n)W$ for all $\underset{\sim}{a}_0' \geq \underset{\sim}{0}$, $\underset{\sim}{a}_i \geq 0$, $i = 1, \ldots, n$ and for all $n = 0, 1, 2, \ldots$, the proof of this assertion being identical to that of Theorem 4.1.

We now turn to the somewhat more complicated problem of establishing the a.s. convergence of $X(\underset{\sim}{a}_0', \ldots, \underset{\sim}{a}_n, t, x)e^{-ct}$ (see equation (4.1)). Let φ be the function defined on $[0, \infty)$ as follows:

$$(4.12) \qquad \varphi(u) = \begin{cases} 1 & , u \leq \beta \\ e^{-\alpha(u-\beta)} & , u > \beta \end{cases}$$

where $\alpha, \beta > 0$. Consider now the random variable $X^\varphi(\underset{\sim}{a}_0', \ldots, \underset{\sim}{a}_n, t)$ defined (a.s.) as the Lebesgue-Stieltjes integral,

$$(4.13) \qquad X^\varphi(\underset{\sim}{a}_0', \ldots, \underset{\sim}{a}_n, t) = \int \varphi(u) X(\underset{\sim}{a}_0', \ldots, \underset{\sim}{a}_n, t, du)$$

$$= \alpha \int e^{-\alpha u} X(\underset{\sim}{a}_0', \ldots, \underset{\sim}{a}_n, t, u+\beta) du ,$$

(the second equality can be established by integrating by parts). Notice also that if $0 \le \beta < x$,

$$(4.14) \quad X^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t) \le X(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t,x) + e^{-\alpha(x-\beta)}X(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t,\infty) .$$

Defining $M^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t) = E[X^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t)]$ and $S^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t,\tau) = E[X^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t)X^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t+\tau)]$ we can write down the renewal equations

$$(4.15) \quad M^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t) = M_n^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t) + \int_0^t M^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t-u)g(u)\,du ,$$

where

$$(4.16) \quad M_n^{\varphi}(a_0',\ldots,a_n,t) = \underset{\substack{\underset{\sim}{y}_0'\le \underset{\sim}{a}_0',\ \underset{\sim}{y}_i\le \underset{\sim}{a}_i,\ i=1,\ldots,n,\\ y_{04}+y_{03}+\ldots+y_{n2}+y_{n1}\le t}}{\int \cdots\cdots \int} A^n f(\underset{\sim}{y}_n)\cdots f(\underset{\sim}{y}_0)\varphi(t-y_{03}-y_{02}- \\ \ldots -y_{n1})\,d\underset{\sim}{y}_0\ldots d\underset{\sim}{y}_n ,$$

and

$$(4.17) \quad S^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t,\tau) = Q^{\varphi}(t,\tau) \\ + h''(1)\int_0^t M^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t-u)M^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t+\tau-u)g(u)\,du \\ + A\int_0^t S^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t-u,\tau)g(u)\,du ,$$

where

$$0 \le Q^{\varphi}(t,\tau) \le E[Z_n^E(\underset{\sim}{a}_0'',\ldots,\underset{\sim}{a}_n,t)Z^B(t+\tau) + Z^B(t)Z_n^E(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t+\tau)] .$$

The upper bound here is the same as in equation (4.5). Although it could easily be improved it is all we need to establish the following analogue of Lemma 4.1.

<u>Lemma 4.2</u>. Under the conditions specified in Section 1, the quantities M^{φ}, S^{φ} and X^{φ} satisfy the relations (4.6), (4.7) and (4.8) respectively

provided $\rho^{E}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n)$ is replaced by

$$(4.18)\quad \rho^{\varphi}(a_0',\ldots,a_n) = \frac{\int\limits_{\underset{\sim}{y}_0' \le \underset{\sim}{a}_0'} e^{-c(y_{04}+\ldots+y_{01})} \left[\int_0^{\infty} e^{-cu}\varphi(u+y_{04})\,du\right] f(\underset{\sim}{y}_0)\,d\underset{\sim}{y}_0}{A\int y\, e^{-cy} g(y)\,dy} \prod_{i=1}^{n} \Psi_i(\underset{\sim}{a}_i).$$

As in Lemma 4.1 the terms $o(1)$ may also be replaced by $o(e^{-\epsilon t})$ where $\epsilon > 0$ and ϵ does not depend on t or τ.

Proof. Exactly the same as that of Lemma 4.1 and Remark 4.1.

Theorem 4.2. $X^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t)e^{-ct} \overset{a.s.}{\to} \rho^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n)W$ for all $\underset{\sim}{a}_0' \ge \underset{\sim}{0}$, $\underset{\sim}{a}_i \ge \underset{\sim}{0}$, $i = 1,\ldots,n$ and for all $n = 0,1,2,\ldots$.

Proof. The proof of Theorem 4.1 carries over without change once we have established an inequality to play the role of (4.10). If we write $V_t = X^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t)e^{-ct}$ and $V = \rho^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n)W$ then the required inequality is

$$(4.19)\qquad V_{t+\tau} \ge e^{-(c+\alpha)\tau}\, V_t$$

which follows from (4.13) and the fact that

$$X^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t+\tau) \ge e^{-c\tau}\, X^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t) \quad \text{for all } \tau \ge 0 \ .$$

With (4.19) replacing (4.10) the proof is now identical to that of Theorem 4.1.

Corollary. $X(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t,\infty)e^{-ct} \overset{a.s.}{\to} \rho^{I_{[0,\infty)}}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n)W$ where $\rho^{I_{[0,\infty)}}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n)$ is obtained from (4.18) on replacing φ by the indicator function $I_{[0,\infty)}$.

Proof. Take $\beta = +\infty$ in the definition (4.12) of φ, in which case $X^{\varphi}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t)$ becomes $X(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t,\infty)$ and the arguments leading to Theorem 4.2 all remain valid.

Theorem 4.3. $X(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n,t,x) \overset{a.s.}{\to} \rho^{I_{[0,x]}}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n)W$ where $\rho^{I_{[0,x]}}(\underset{\sim}{a}_0',\ldots,\underset{\sim}{a}_n)$ is obtained from (4.18) on replacing φ by the indicator function $I_{[0,x]}$.

Proof. Observe that for $0 \le \beta < x$, equation (4.13) implies that

$$(4.20) \quad e^{-ct} X^{\varphi}(\underset{\sim}{a}_0', \ldots, \underset{\sim}{a}_n, t) - e^{-\alpha(x-\beta)} e^{-ct} X(\underset{\sim}{a}_0', \ldots, \underset{\sim}{a}_n, t, \infty)$$

$$\le e^{-ct} X(\underset{\sim}{a}_0', \ldots, \underset{\sim}{a}_n, t, x) \le e^{-ct} X^{\bar{\varphi}}(\underset{\sim}{a}_0', \ldots, \underset{\sim}{a}_n, t)$$

where $\bar{\varphi}(u) = 1$ if $u \le x$, $e^{-\alpha(u-x)}$ if $u > x$.

Letting $t \uparrow \infty$ in equation (4.20) we find that a.s.

$$\rho^{\varphi}(\underset{\sim}{a}_0', \ldots, \underset{\sim}{a}_n) W - e^{-\alpha(x-\beta)} \rho^{I_{[0,\infty)}}(\underset{\sim}{a}_0', \ldots, \underset{\sim}{a}_n) W \le \underline{\lim}_{t \to \infty} e^{-ct} X(\underset{\sim}{a}_0', \ldots, \underset{\sim}{a}_n, t, x)$$

$$\le \overline{\lim}_{t \to \infty} e^{-ct} X(\underset{\sim}{a}_0', \ldots, \underset{\sim}{a}_n, t, x) \le \rho^{\bar{\varphi}}(\underset{\sim}{a}_0', \ldots, \underset{\sim}{a}_n) W .$$

Now letting $\alpha \uparrow \infty$ then $\beta \uparrow x$ we obtain the assertion of the Theorem.

Theorem 4.4. $P[\lim_{t \to \infty} Z^{(4)}(\underset{\sim}{a}_0, \ldots, \underset{\sim}{a}_n, t) e^{-ct} = K^{(4)} F_n(\underset{\sim}{a}_0, \ldots, \underset{\sim}{a}_n) W$ for all $\underset{\sim}{a}_i \ge \underset{\sim}{0}$, $i = 0, 1, \ldots, n$ and for all $n = 0, 1, 2, \ldots] = 1$, where $K^{(4)}$ and F_n were defined by equations (4.4) and (4.11).

Proof. The fact that for each fixed $\underset{\sim}{a}_0, \ldots, \underset{\sim}{a}_n$,

$$Z^{(4)}(\underset{\sim}{a}_0, \ldots, \underset{\sim}{a}_n, t) e^{-ct} \overset{a.s.}{\to} K^{(4)} F_n(\underset{\sim}{a}_0, \ldots, \underset{\sim}{a}_n) W$$

as $t \to \infty$ follows at once from Theorems 4.1 and 4.3 and Remark 4.2. The stronger assertion made by Theorem 4.4 is a consequence of the a.s. monotonicity properties of $Z^{(4)}(\underset{\sim}{a}_0, \ldots, \underset{\sim}{a}_n, t)$ for each t, the continuity of F_n and the denseness of the rationals in the real line.

Corollary. Given that $Z(t) \not\to 0$ as $t \to \infty$, the ratio

$$\frac{Z^{(4)}(\underset{\sim}{a}_0, \ldots, \underset{\sim}{a}_n, t)}{Z^{(4)}(t)}$$ converges to the joint distribution function $F_n(\underset{\sim}{a}_0, \ldots, \underset{\sim}{a}_n)$ for all $\underset{\sim}{a}_0 \ge \underset{\sim}{0}, \ldots, \underset{\sim}{a}_n \ge \underset{\sim}{0}$ and for all $n = 0, 1, 2, \ldots$, with probability 1.

Proof. We note first that $Z^{(4)}(t)e^{-ct} \xrightarrow{a.s.} K^{(4)}W$. (The proof of this fact is similar to but much easier than the proofs leading to Theorem 4.4 above.) Moreover it is shown by Harris [7], p. 147 that the events $\{Z(t) \to 0\}$ and $\{W = 0\}$ are a.s. equal. The assertion of the corollary is an immediate consequence of these facts.

§5. Random Selection of a Cell From the Population as a Whole.

The corollary of Theorem 4.4 gives the asymptotic joint distribution of $\underset{\sim}{Y}_0,\ldots,\underset{\sim}{Y}_n$ for a cell selected randomly from those in phase 4 at time t. Analogous results could clearly be derived for sampling from any particular phase. We now indicate the corresponding result for sampling from the entire cell population.

Let $Z^{(i)}(\underset{\sim}{a}_0^{(i)},\underset{\sim}{a}_1,\ldots,\underset{\sim}{a}_n,t)$ be the number of cells in phase i at time t for which the elapsed phase durations $Y_{0i},\ldots,Y_{01},\ldots,Y_{n4},\ldots,Y_{n1}$ of the cell itself and its first n ancestors satisfy $Y_{0i} \le a_{0i}^{(i)},\ldots,$ $Y_{01} \le a_{01}^{(i)},\ldots,Y_{n4} \le a_{n4},\ldots,Y_{n1} \le a_{n1}$.

Then by the same arguments as we have used for the case $i = 4$ it can be shown that, conditional on $Z(t) \not\to 0$,

$$(5.1) \qquad \frac{Z^{(i)}(\underset{\sim}{a}_0^{(i)},\underset{\sim}{a}_1,\ldots,\underset{\sim}{a}_n,t)}{Z(t)} \to \frac{K^{(i)}}{K}\, F_n^{(i)}(\underset{\sim}{a}_0^{(i)},\ldots,\underset{\sim}{a}_n) \quad \text{for all}$$

$$\underset{\sim}{a}_0^{(i)} \ge \underset{\sim}{0},\ldots,\underset{\sim}{a}_n \ge \underset{\sim}{0} \quad \text{for all} \quad n = 0,1,2,\ldots \quad \text{almost surely,}$$

where

$$(5.2) \qquad K^{(i)} = \frac{\int e^{-c(y_1+\ldots+y_i)}\left[\int_{y_i}^{\infty}[f(\underset{\sim}{y})]_{y_i=u}\,du\right]d\underset{\sim}{y}}{A\int y\, e^{-cy} g(y)\,dy}$$

and

$$(5.3) \qquad F_n^{(i)}(\underset{\sim}{a}_0^{(i)},\ldots,\underset{\sim}{a}_n) = \frac{\displaystyle\int_{(y_1,\ldots,y_i)\le \underset{\sim}{a}_0^{(i)}} e^{-c(y_1+\ldots+y_i)}\left[\int_{y_i}^{\infty}[f(\underset{\sim}{y})]_{y_i=u}\,du\right]d\underset{\sim}{y}}{\displaystyle\int e^{-c(y_1+\ldots+y_i)}\left[\int_{y_i}^{\infty}[f(\underset{\sim}{y})]_{y_i=u}\,du\right]d\underset{\sim}{y}} \prod_{k=1}^{n} \Psi_k(\underset{\sim}{a}_k)\,.$$

(The function Ψ_k was defined in Section 2.)

Equation (5.1) defines in considerable detail the asymptotic age-phase structure of the population as $t \to \infty$, conditional upon non-extinction. We note that $F_n^{(i)}(\underset{\sim}{a}_0^{(i)},\ldots,\underset{\sim}{a}_n)$ defined in (5.3) gives the asymptotic joint distribution of $Y_{0i},\ldots,Y_{01},\underset{\sim}{Y}_1,\ldots,\underset{\sim}{Y}_n$ for a cell selected randomly from those in phase i at time t.

§6. Generalization to Include Random Cell Removal.

We now consider the following modification of the process considered in Sections 1-5. Suppose that in addition to elimination of cells by death at the end of the cycle there is "removal" of cells in the following manner: for any cell, independently of its age or phase and independently of other cells, the probability of removal in any small time interval $(t, t+\delta t)$ is $\lambda\delta t + o(\delta t)$; the probability that more than one cell is removed in $(t, t+\delta t)$ is $o(\delta t)$. This model describes the growth of the population of proliferating cells in a population where cells may differentiate and become non-proliferative during the cycle. As before, c stands for the unique positive root of equation (2.9). We consider the following three cases.

Case 1. $\lambda < c$. It is not difficult to check that if we again start with a single cell with age zero at epoch zero but now define $\hat{Z}(t)$ to be

$$\hat{Z}(t) = e^{\lambda t} \cdot Z(t)$$

where $Z(t)$ is the number of cells in the population at time t, and likewise $\hat{Z}^4(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t)$, $\hat{Z}^{(4)}(t)$, etc. each to be $e^{\lambda t}$ times their definitions given in Section 1, then the whole of Sections 2-5 carries over verbatim to the new process. We end this case by remarking that the condition $\lambda < c$ is used at the place where we establish the m.s. convergence.

Case 2. $\lambda = c$. This is the case corresponding to the so called critical branching process. The procedures used in Section 3 do not hold in this case. In order to discuss the asymptotic behavior of the random variable $Z(t)$ we define I_k to be the number of cells in the k^{th} generation. Observe that the probability that a cell is removed from the cycle before multiplication is $1 - f^*(\lambda)$. If we regard a cell removed from the cycle as a death (giving no birth), then the number of daughter cells per parent is a random variable with probability generating function $H(s) = 1 - f^*(\lambda) + f^*(\lambda)h(s)$. Proceeding the same arguments used in Theorem 5.1 and Theorem 5.2 of Harris [7], pp. 127-128 with minor changes we conclude that the random variable I_k are a Galton-Watson branching process with generating function $H(s)$ and two events $\{Z(t) \to 0$ as $t \to \infty\}$ and $\{I_k$ is equal to zero for some $k\}$ are a.s. equal. The above results, together with the fact $H'(1) = 1$, imply that

$$P(Z(t) \to 0 \quad \text{as} \quad t \to \infty) = 1 .$$

Case 3. $\lambda > c$. Since the loss rate λ is greater than the growth rate c, an immediate question to ask is "Will the population die out eventually?" In order to answer this question we make use of the following decomposition,

$$Z(t) = Z^E(t) - Z^D(t), \tag{6.1}$$

where $Z^E(t)$ is defined to be the total number of cells entering the cycle during the time interval $[0,t]$ and $Z^D(t)$ is defined to be the total number of cells leaving the cycle during the time interval $[0,t]$. Again, using the method suggested in case 1, we obtain immediately the following results.

Theorem 6.1. If we denote $M^E(t) = EZ^E(t)$ and $M^D(t) = EZ^D(t)$, then

(i) $M(t) \to 0$ as $t \to \infty$

(ii) $M^E(t) \to B$ and $M^D(t) \to B$ as $t \to \infty$,

where $B = [1 - A\int_0^\infty e^{-\lambda t} g(t)dt]^{-1}$,

(iii) $Z^E(t) \to W^E$ and $Z^D(t) \to W^D$ as $t \to \infty$,

where W^E and W^D are random variables,

(iv) $EW^E = EW^D = B$ and $P(W^E < \infty) = P(W^D < \infty) = 1$

and

(v) $P(Z(t) \to 0$ as $t \to \infty) = 1$.

Proof. (i) and (ii) are the immediate consequence of applying the standard asymptotic theory of the renewal equation to the equations obtained by using the method suggested in case 1. (iii) follows from the monoticity of the random variables $Z^E(t)$ and $Z^D(t)$. (iv) follows from (iii) and the fact that B is a finite constant. It remains to show (v). Equation (6.1) and (iii) together imply that $Z(t)$ converges to a random variable W, $W = W^E - W^D$, as t approaches infinity. Since $Z(t)$ is a nonnegative function of t we have $w \geq 0$. This, together with the fact that $EW = 0$, establishes (v).

§7. The Case of an Arbitrary Initial Population.

We have assumed from Sections 2-6 that the process starts with a single cell of age zero at epoch zero. In the following we indicate that the corollary to Theorem 4.4 remains valid regardless of the phase and elapsed phase durations of the initial cell and indeed of the number of initial cells.

Case 1. One initial cell. Suppose at epoch zero the population consists of a single cell with age a_0 in phase i (age in phase i being measured from entry into phase i). Assuming random cell removal as defined in Section 6 and letting $\hat{Z}(t)$ denote the product $e^{\lambda t} \cdot Z(t)$ (where $\lambda = 0$ if there is no cell removed) we have

$$(7.1)\quad \hat{Z}(t) = e^{\lambda t} I_{(-\infty,0)}(t-U) I_{(-\infty,0)}(t-L) + e^{\lambda U} I_{(0,\infty)}(t-U) I_{(0,\infty)}(L-U) \sum_{i=1}^{N} \hat{Z}_i(t-U),$$

where U, L and N are independent random variables (U is the epoch at which the initial cell completes its cycle, L is the epoch at which the initial cell leaves the system before completing its cycle and N is the number of daughter cells which it produces) while $\{Z_1,Z_2,Z_3,\ldots\}$ is a sequence of independent copies of the process starting with a single ancestor of age zero. $U,L,N,Z_1,Z_2,\ldots$ are all independent and can be defined on a single product space endowed with the appropriate product measures. From equation (7.1) it follows that, for $\lambda < c$,

$$(7.2)\quad e^{-ct}\, \hat{Z}(t) \underset{m.s.}{\overset{a.s.}{\to}} K\, e^{-(c-\lambda)U} \sum_{i=1}^{N} W_i ,$$

where K was defined in equation (2.12) and W_i is the a.s. (and m.s.) limit of $e^{-ct}\hat{Z}_i(t)$ as $t \to \infty$, and, for $\lambda > c$,

$$(7.3) \qquad Z(t) = I_{(-\infty,0)}(t-U)I_{(-\infty,0)}(t-L) + I_{(0,\infty)}(t-U)I_{(0,\infty)}(L-U)\sum_{i=1}^{N} Z_i(t-U)$$

which implies that

$$(7.4) \qquad Z(t) \overset{a.s.}{\to} 0 \quad \text{as} \quad t \to \infty .$$

We observe that the two events $\{\hat{Z}(t) \to 0\}$ and $\{W_1 = W_2 = \ldots = W_N = 0\}$ are a.s. equal.

The same reasoning shows that Theorem 4.4 and the results of Section 5 remain valid provided the random variable $Z^{(4)}(\underset{\sim}{a}_0,\ldots,\underset{\sim}{a}_n,t)$ is interpreted as $e^{\lambda t}$ times the definition given in Section 1 and the random variable W is replaced by $e^{-(c-\lambda)U}\sum_{i=1}^{N} W_i$. The corollary of Theorem 4.4 remains valid as it stands.

<u>Case 2</u>. <u>K initial cells</u>. If $\lambda < c$ and at epoch zero there are K cells with given phases and ages (measured from entry into the given phase) then Theorem 2.4.4 and its corollary and the results of Section 5 remain valid provided the random variable W is replaced by

$$\sum_{j=1}^{K} e^{-(c-\lambda)U_j} \sum_{i=1}^{N_j} W_i$$

where U_j is the epoch at which the j^{th} initial cell divides ($j=1,\ldots,K$) and N_j is the number of daughter cells it produces. (The probability space on which the W_i's are defined is now a K-fold product of spaces like that considered in case 1 of this section.) Similarly, for the case $\lambda > c$, Theorem 6.1 remains valid.

§8. Application of the Asymptotic n-fold Age Distribution to the Analysis of Labelling Experiments.

To make the results derived in this section more readable, we use the following notation: We denote by $f_{1,2,3,4}(t_1,t_2,t_3,t_4)$ the joint probability density of the four phase durations T_1,T_2,T_3,T_4 for a new born cell. Marginal probability densities will be denoted in the usual way, e.g. $f_{2,3}$ for the density of T_2 and T_3, and the joint densities of sums of phase durations will be indicated by deletion of appropriate commas in the subscripts. Thus for example, the trivariate joint density of T_1, T_2+T_3, T_4 will be denoted $f_{1,23,4}$ and the bivariate density of T_1+T_2, T_3+T_4 will be denoted $f_{12,34}$. We will still use g to denote the probability density of $T_1+T_2+T_3+T_4$. Laplace transforms will be denoted by a superscript $*$. Thus for example

$$(8.1)\quad f^*_{1,23,4}(\alpha,\beta,\gamma) = \int_0^\infty\int_0^\infty\int_0^\infty e^{-\alpha u-\beta v-\gamma w} f_{1,23,4}(u,v,w)\,du\,dv\,dw \ .$$

The Laplace transform of a function of several variables will sometimes be taken with respect to a subset of the variables. The particular subset involved will be indicated by underlining the appropriate subscripts, e.g.

$$(8.2)\quad f^*_{\underline{1},\underline{23},4}(c,c+s,w) = \int_0^\infty\int_0^\infty e^{-cu-(c+s)v} f_{1,23,4}(u,v,w)\,du\,dv \ .$$

Suppose first that the label is administered continuously to a cell population (growing according to the model we have described) from time t_0 onwards. All unlabelled cells entering phase 2 after time t_0 will then become labelled and all offspring of labelled cells will be labelled from birth. Let $C_{t_0}(t)$ be the fraction of cells in phase 4 at time $t_0 + t$ which are labelled. Assuming that t_0 is large and that the observed population is a realization of the process which does not die out, then for each fixed t as $t_0 \to \infty$, $C_{t_0}(t)$ is just the asymptotic probability that a cell selected at random from those in phase 4 entered phase 3 at time t or less prior to selection, i.e.

$$(8.3) \qquad \lim_{t_0 \to \infty} c_{t_0}(t) = \int_{a_{04}+a_{03} \le t} F_0(d\underset{\sim}{a}_0)$$

$$= \frac{c}{f^*_{123}(c) - g^*(c)} \int_0^t e^{-cy} \int_0^y [f^*_{\underline{12},3}(c,u) - f^*_{\underline{12},34}(c,u)]\,du\,dy,$$

where F_0 is the distribution function defined by equation (2.11).

Similarly for pulse labelling (administered at time t_0) the fraction of labelled cells in phase 4 at time $t_0 + t$ satisfies (under the assumptions of the previous paragraph)

$$(8.4) \qquad \lim_{t_0 \to \infty} P_{t_0}(t) = \sum_{\nu=0}^{\infty} P_\nu(t),$$

where

$$(8.5) \qquad P_\nu(t) = \int_{0 \le t - a_{04} - a_{03} - \cdots - a_{\nu 4} - a_{\nu 3} \le a_{\nu 2}} \int F_\nu(d\underset{\sim}{a}_0, d\underset{\sim}{a}_1, \ldots, d\underset{\sim}{a}_\nu) \ .$$

The probabilities $P_\nu(t)$, $\nu = 0,1,2,\ldots$ can be calculated by carrying out the straightforward integration. If we take Laplace transforms in (8.5) we find that

$$(8.6) \qquad P^*_0(s) = \frac{c[f^*_{123}(c) - f^*(c)]^{-1}}{c(s+c)} [f^*_{12,3}(c,s+c) - f^*_{1,23}(c,s+c)$$

$$- f^*_{12,34}(c,s+c) + f^*_{1,234}(c,s+c)]$$

and

$$(8.7)\quad P^*_\nu(s) = \frac{cA[Ag^*(s+c)]^{\nu-1}}{c(s+c)[f^*_{123}(c)-g^*(c)]}[f^*_{123}(s+c)-g^*(s+c)][f^*_{12,34}(c,s+c) - f^*_{1,234}(c,s+c)],\ \nu = 1,2,\ldots .$$

Summing over $\nu = 0,1,2,\ldots$ we obtain

$$(8.8)\quad \int_0^\infty e^{-st} \lim_{t_0\to\infty} P_{t_0}(t) = \frac{c[f^*_{123}(c)-g^*(c)]^{-1}}{s(s+c)[1-Ag^*(s+c)]}\{[1-Ag^*(s+c)][f^*_{12,3}(c,s+c) - f^*_{1,23}(c,s+c)]-[1-Af^*_{123}(c+s)][f^*_{12,34}(c,c+s)-f^*_{1,234}(c,c+s)]\} .$$

It is easy to check that the result (8.8) reduces to the result of Trucco and Brockwell [17] when the phase durations are independent and that both (8.3) and (8.8) reduce to the results of Macdonald [9] when the duration of phase 4 is independent of the durations of phases 1, 2 and 3. (Macdonald considers pulse-labelling with a pulse of non-zero duration δ. The labelling function for such a pulse administered over the interval $[t_0,t_0+\delta]$ is obtained from (8.5) on replacing a_{n2} in the inequalities defining the range of integration by $a_{n2} + \delta$.)

CHAPTER III

AGE DISTRIBUTIONS AND LABELLING FUNCTIONS FOR MULTI-TYPE MULTIPHASE BRANCHING PROCESS

§1. Introduction

The model considered in the previous chapter assumes that every cell has the same phase-duration distribution. In this chapter we shall remove this assumption to make the model more general. Consider a population consisting of m $(m > 1)$ types. We assume that the life-cycle of a cell of the i^{th} type, $i = 1,\dots,m$, is composed of four phases whose durations are random variables with joint probability density function

$$(1.1) \qquad f_i(t_1,t_2,t_3,t_4) = f_i(\underset{\sim}{t}) \;.$$

The distribution of total lifetime for any cell of type i thus has the probability density function

$$(1.2) \qquad g_i(u) = \int\int\int_{t_1+t_2+t_3\le u} f_i(t_1,t_2,t_3,u-t_1-t_2-t_3)dt_1dt_2dt_3 \;.$$

At the end of its life a cell of i^{th} type gives rise to a random vector $\underset{\sim}{N} = (N_1,\dots,N_m)$ of numbers N_j of cells of type j, $j = 1,\dots,m$. Each new-born cell then goes through the cycle appropriate to its type. We denote the probability distribution of the offspring vector $\underset{\sim}{N}$ of a cell of type i by

(1.3) $P_i(\underset{\sim}{n}) = P_i(\underset{\sim}{N} = \underset{\sim}{n})$, where $\underset{\sim}{n} = (n_1,\ldots,n_m)$, $n_j = 0,1,2,\ldots,$
and $j = 1,\ldots,m$,

and the generating function by

$$(1.4) \qquad h_i(\underset{\sim}{s}) = \sum_{\underset{\sim}{n}} P_i(\underset{\sim}{n})(\underset{\sim}{s})^{\underset{\sim}{n}} ,$$

where $\underset{\sim}{s} = (s_1,\ldots,s_m)$, $(\underset{\sim}{s})^{\underset{\sim}{n}} = \prod_{i=1}^{m} s_i^{n_i}$ and s_i are complex numbers such that $|s_i| \le 1$, $i = 1,\ldots,m$.

The daughter cells of any parent are assumed to behave independently of each other. Starting with a new-born cell of i^{th} type, it has been shown that

$$P\{\lim_{t\to\infty} \underset{\sim}{Z}_i(t)e^{-ct} = \underset{\sim}{W}\} = 1$$

where $\underset{\sim}{Z}_i(t) = (Z_i(1,t),\ldots,Z_i(m,t))$, $Z_i(j,t)$ is defined to be the number of cells of type j at time t and $\underset{\sim}{W}$ is a random vector. We will obtain the analogous results which will determine the asymptotic type-phase-age distribution of the population for large values of time t. This distribution will play a very important role in determining the FLM function which will be discussed in Section 8.

§2. Notation

Consider the population at time t descended from a single ancestor of type i having age zero at epoch zero. Define $Z_{ki}^{(4)}(\underset{\sim}{a}_0, i_0, t)$ to be the number of cells of type i_0 in phase 4 at time t which are k^{th} generation descendants of the initial cell and for which the time $Y_{01}^{i_0}, Y_{02}^{i_0}, Y_{03}^{i_0}, Y_{04}^{i_0}$ spent (prior to epoch t) in phase 1,2,3,4 satisfy

$$(2.1) \qquad Y_{0i}^{i_0} \leq a_{0i}^{i_0}, \quad i = 1,2,3,4 .$$

We shall again use the vector notation $\underset{\sim}{Y}_0^{i_0} \leq \underset{\sim}{a}_0$ as an abbreviation for the set of inequalities $Y_{0i}^{i_0} \leq a_{0i}$ between corresponding components of $\underset{\sim}{Y}_0^{i_0}$ and $\underset{\sim}{a}_0$. More generally we define $Z_{ki}^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t)$, $k \geq n$, to be the number of cells of type i_0 in phase 4 at time t which are k^{th} generation descendants of the initial cell and whose j^{th} ancestor was of type i_j, $j = 0,1,\ldots,n$ and spent times $Y_{j1}^{i_j}, \ldots, Y_{j4}^{i_j}$ in phases $1,\ldots,4$ satisfying $\underset{\sim}{Y}_j^{i_j} \leq \underset{\sim}{a}_j$. The quantities $Z_{ki}^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t)$ are random variables defined on a probability space of family trees (whose construction is essentially the same as that in Mode [13]). We shall also make use of the random variables $Z_i(j,t)$, the number of cells of type j at time t, $Z_i^{(4)}(j,t)$, the number of cells of type j in phase 4 at time t, and

$$(2.2) \qquad Z_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t) = \sum_{k=n}^{\infty} Z_{ki}^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t) .$$

We shall also use the following vector random variables.

$$(2.3) \qquad \underset{\sim}{Z}_i^{(4)}(\underset{\sim}{a}_0, \underset{\sim}{a}_1, i_1, \ldots, \underset{\sim}{a}_n, i_n, t) = (Z_i^{(4)}(\underset{\sim}{a}_0, 1, \underset{\sim}{a}_1, i_1, \ldots, \underset{\sim}{a}_n, i_n, t),$$

$$\ldots, Z_i^{(4)}(\underset{\sim}{a}_0, m, \underset{\sim}{a}_1, i_1, \ldots, \underset{\sim}{a}_n, i_n, t)),$$

(2.4) $$\underset{\sim}{Z}_i^{(4)}(t) = (Z_i^{(4)}(1,t),\dots,Z_i^{(4)}(m,t)) ,$$

and

(2.5) $$\underset{\sim}{Z}_i(t) = (Z_i(1,t),\dots,Z_i(m,t)) ,$$

where m is the total number of different types considered in this model. Our main interest is to find the asymptotic joint distribution of the types and the phase durations of a cell selected at random from those in mitosis and its first n ancestors. A corresponding result for sampling from an arbitrary phase is given in Section 7.

We shall use M and S with appropriate subscripts and superscripts to denote first and second moments of the corresponding random variables, e.g.

$$M_{ki}^{(4)}(\underset{\sim}{a}_0,i_1,\dots,\underset{\sim}{a}_n,i_n,t) = E\, Z_{ki}^{(4)}(\underset{\sim}{a}_0,i_0,\dots,\underset{\sim}{a}_n,i_n,t),$$

$$M_i^{(4)}(\underset{\sim}{a}_0,i_0,\dots,\underset{\sim}{a}_n,i_n,t) = E\, Z_i^{(4)}(\underset{\sim}{a}_0,i_0,\dots,\underset{\sim}{a}_n,i_n,t),$$

$$S_{k\ell i}^{(4)}(\underset{\sim}{a}_0,i_0,\dots,\underset{\sim}{a}_n,i_n,t,\tau) = E[Z_{ki}^{(4)}(\underset{\sim}{a}_0,i_0,\dots,\underset{\sim}{a}_n,i_n,t)\, Z_{\ell i}^{(4)}(\underset{\sim}{a}_0,i_0,\dots,\underset{\sim}{a}_n,i_n,t+\tau)],$$

$$M_i^{(4)}(j,t) = E\, Z_i^{(4)}(j,t),\ M_i(j,t) = E\, Z_i(j,t),$$

$$S_i^{(4)}(j,k,t,\tau) = E[Z_i^{(4)}(j,t)Z_i^{(4)}(k,t+\tau)],$$

$$S_i(j,k,t,\tau) = E[Z_i(j,t)Z_i(k,t+\tau)],\ \text{etc.}$$

$$S_i(j,t,\tau) = S_i(j,j,t,\tau),$$

Those moments are known to be finite for all finite t, τ because of the finiteness of $M_i(j,t)$ and $S_i(j,t,\tau)$, provided $\|M\|$ and

$\max_{i,j,k} \mu_{ijk}$ (to be defined later) are finite (see Mode [11]).

We shall also denote $M = (m_{ij})$ and $H(\lambda) = (g_i^*(\lambda)m_{ij})$, two $m \times m$ matrices with ij^{th} element m_{ij} and $g_i^*(\lambda)m_{ij}$ respectively, where the superscript * stands for Laplace transform and

$$m_{ij} = D_j h_i(\underset{\sim}{s})\Big|_{\underset{\sim}{s}=(1,\ldots,1)}$$

with $D_j h_i(\underset{\sim}{s})$ standing for the first derivative of the function $h_i(\underset{\sim}{s})$ with respect to the j^{th} element in the vector $\underset{\sim}{s}$. We set $b_{ij}(\lambda) = (-1)^{i+j}\Delta_{ij}(\lambda)$, where $\Delta_{ij}(\lambda)$ is the complementary minor of the ij^{th} element in $I - H(\lambda)$ and I is an identity matrix, $\Delta(\lambda) = |I - H(\lambda)|$, the determinant of $I - H(\lambda)$, and $[B(\lambda)]_{ij} = b_{ji}(\lambda)$, the adjoint of the matrix $[I - H(\lambda)]$.

We shall use the matrix norm $\|M\|$ defined by

$$\|M\| = \max_i \sum_j |m_{ij}| \tag{2.6}$$

and denote $\mu_{ijk} = D^2_{jk} h_i(\underset{\sim}{s})\Big|_{\underset{\sim}{s}=(1,\ldots,1)}$, where D^2_{jk} stands for the second-order partial derivative of the function $h_i(\underset{\sim}{s})$ with respect to the elements s_j and s_k in the vector $\underset{\sim}{s} = (s_1,\ldots,s_m)$.

§3. Summary of the Known Results in Multitype Single-Phase Branching Process

We summarize some of the known facts in the following.

Lemma 3.1. (a) If M is a matrix of positive elements (every element is positive) and the Perron-Frobenius root (unique positive largest eigenvalue) of M is greater than 1, then there exists a positive number c such that $\Delta(c) = 0$. The number c has the following properties.

(b) The Perron-Frobenius root of the matrix $H(c)$ is 1 and corresponding to this root there are positive left and right eigenvectors $\underset{\sim}{\eta}$ and $\underset{\sim}{\mu}$ such that $\underset{\sim}{\eta} = \underset{\sim}{\eta}H(c)$, $H(c)\underset{\sim}{\mu} = \underset{\sim}{\mu}$ and $\underset{\sim}{\eta}\,\underset{\sim}{\mu} = 1$.

(c) c is the root of the determinantal equation $\Delta(\lambda) = 0$ with largest real part and has multiplicity 1.

(d) $H(c)B(c) = B(c)$ and $B(c)H(c) = B(c)$.

(e) $B(c) = K\underset{\sim}{\mu}\underset{\sim}{\eta} = K(\mu_i \eta_j)$ for some positive constant K and the positive numbers μ_i and η_j are the i^{th} and j^{th} elements in the vectors $\underset{\sim}{\mu}$ and $\underset{\sim}{\eta}$.

Proof: See Theorem (3.1) of Mode [11].

Remark 3.1. The condition that every element in M is positive can be relaxed to some extent. All we need is to assume that M is a matrix of nonnegative elements with at least one positive element and that there exists a positive integer n such that M^n is a matrix of positive elements (see Karlin [8], pp. 469-484).

The above lemma leads to the following result.

Lemma 3.2. If M is a matrix of positive elements such that the Perron-Frobenius root of M is greater than 1 and if $g_i \in L^p$ for some $p > 1$ for all $i = 1,\ldots,m$, then

$$M_i(j,t) = C_{ij}e^{ct}[1 + O(e^{-\epsilon t})],\ t \to \infty \text{ for some } \epsilon > 0, \tag{3.1}$$

where

$$C_{ij} = \frac{b_{ji}(c)}{\Delta'(c)} \int_0^\infty e^{-cy}(1 - G_j(y))dy\ , \tag{3.2}$$

$G_j(y) = \int_0^y g_j(t)dt$, $b_{ij}(c)$ and $\Delta(c)$ were defined in Section 2, c was defined in Lemma 3.1 and $\Delta'(c)$ is the derivative of $\Delta(\lambda)$ evaluated at c.

Proof: See Theorem 3.2 of Mode [11].

It follows from (e) of Lemma 3.1 that

$$C_{ij} = d\mu_i \eta_j \int_0^\infty e^{-cy}(1 - G_j(y))dy \tag{3.3}$$

where $d = \dfrac{k}{\Delta'(c)}$.

Now we let d_{ij} be the ij^{th} element in the matrix $(I-H(2c))^{-1}$ and also use the following notation;

$$d_i = d^2 g_i^*(2c) \sum_{\ell,k=1}^{m} \mu_{i\ell k}\mu_\ell \mu_k\ , \tag{3.4}$$

$$c_i = \sum_{\ell=1}^{m} d_{i\ell} d_\ell\ , \tag{3.5}$$

$$\nu_j = \eta_j \int_0^\infty e^{-cy}(1 - G_j(y))dy, \tag{3.6}$$

$$W_i(j,t) = Z_i(j,t)e^{-ct}, \tag{3.7}$$

and

$$\underset{\sim}{W}_i(t) = (W_i(1,t),\ldots,W_i(m,t))\ . \tag{3.8}$$

Lemma 3.3. If M is a matrix of positive elements such that the Perron-Frobenius root of M is greater than 1, $g_i \in L^p$ for some $p > 1$ for all $i = 1,\ldots,m$ and $\underset{i,j,k}{\text{Max}}\ \mu_{ijk} < \infty$, then

(i) $S_i(j,k,t,\tau) \sim c_i\nu_j\nu_k e^{2ct+c\tau}$, $t \to \infty$.

(ii) As $t \to \infty$, $E[W_i(j,t+\tau) - W_i(j,t)]^2 \to 0$ uniformly in τ, $\tau \geq 0$.

(iii) $W_i(j,t)$ converges in mean square to a random variable $W_i(j)$ as $t \to \infty$.

(iv) $E\, W_i(j) = C_{ij}$.

(v) $\text{Var}\, W_i(j) = (c_i - d^2\mu_i^2)\nu_j^2$.

(vi) $\text{Var}\, W_i(j) > 0$.

Proof: The proof of (i)-(v) can be found in Theorem 3.2, Mode [12]. It remains to show (vi). It has been shown that $M_i(j,t)$ and $S_i(j,k,t,\tau)$ satisfy the following integral equations

$$(3.9)\quad M_i(j,t) = \delta_{ij}(1 - G_i(t)) + \int_0^t \sum_{\nu=1}^m m_{i\nu}M_\nu(j,t-u)g_i(u)du$$

and

$$(3.10)\quad S_i(j,k,t,\tau) = f_{ijk}(t,\tau) + \int_0^t \sum_{\nu=1}^m m_{i\nu}S_\nu(j,k,t-u,\tau)g_i(u)du$$

where

$$(3.11)\quad \lim_{t\to\infty} f_{ijk}(t,\tau)e^{-(2ct+c\tau)} = \sum_{\nu,\nu'=1}^m \mu_{i\nu\nu'}C_{\nu j}C_{\nu' k}g_i^*(2c) .$$

Applying (iii) to equation (3.9) and (3.10) and using the fact that $g_i^*(2c) > (g_i^*(c))^2$ we obtain

$$(3.12)\quad E[W_i(j)]^2 > \sum_{\substack{\nu,\nu'=1 \\ \nu\neq\nu'}}^m \mu_{i\nu\nu'}C_{\nu j}C_{\nu' j}[g_i^*(c)]^2 + \sum_{\nu=1}^m (\mu_{i\nu\nu} + m_{i\nu})E[W_\nu(j)]^2[g_i^*(c)]^2$$

and

$$(3.13)\quad [E\,W_i(j)]^2 = [\sum_{\nu=1}^{m} m_{i\nu}C_{\nu j}g_i^*(c)]^2 = \sum_{\substack{\nu,\nu'=1\\ \nu\neq\nu'}}^{m} m_{i\nu}m_{i\nu'}C_{\nu j}C_{\nu' j}[g_i^*(c)]^2$$

$$+ \sum_{\nu=1}^{m} m_{i\nu}^2\, E[W_\nu(j)]^2[g_i^*(c)]^2 \;.$$

By the assumption that the daughter cells of any parent behave independently we have $\mu_{i\nu\nu'} = m_{i\nu}m_{i\nu'}$ for $\nu \neq \nu'$. Inequality (3.12) and (3.13), together with the fact that $\mu_{i\nu\nu} - m_{i\nu} - m_{i\nu}^2 \geq 0$, establish (vi). This completes the proof of the lemma.

The result (vi) in Lemma 3.3 is used to show that $\{\underset{\sim}{Z}_i(t) \to \underset{\sim}{0}\}$ and $\{\underset{\sim}{W}_i = \underset{\sim}{0}\}$ are a.s. equal, where $\underset{\sim}{W}_i$ is a vector random variable with components $W_i(j)$.

Now let $I_j^{(n)}$ be the number of cells of j^{th} type belonging to the n^{th} generation. It can be shown that $\{\underset{\sim}{I}^{(n)} = (I_1^{(n)},\ldots,I_m^{(n)});$ $n = 1,2,\ldots\}$ forms a multi-dimensional Galton-Watson process with generating function

$$(3.14)\qquad \underset{\sim}{h} = (h_1,\ldots,h_m)$$

and that the following result.

Lemma 3.4. Let A be the event $\underset{\sim}{I}^n = \underset{\sim}{0}$ for some $n \geq 1$, given that at time zero there is a new born cell of type i and let B be the event that $\underset{\sim}{Z}_i(t)$ is equal to zero for some $t > 0$. If $\|M\| < \infty$, then $P(A) = P(B)$ for $i = 1,\ldots,m$.

Proof: See Theorem 4.1 of Mode [13].

Corollary. $P\{\underset{\sim}{Z}_i(t) \to \underset{\sim}{0}$ as $t \to \infty\}$ is the smallest root $\underset{\sim}{q}$ of the equation

(3.15) $$\underset{\sim}{r} = \underset{\sim}{h}(\underset{\sim}{r}) \; .$$

Proof: See Theorem 7.1, p. 41, of Harris [7].

Consider the moment generating function of $\underset{\sim}{W}_i(t)$ and $\underset{\sim}{W}_i$ given by $Q_i(\underset{\sim}{s},t) = E\, e^{-\underset{\sim}{s}\underset{\sim}{W}_i^T(t)}$ and $Q_i(\underset{\sim}{s}) = E\, e^{-\underset{\sim}{s}\underset{\sim}{W}_i^T}$, $\mathrm{Re}(\underset{\sim}{s}) \ge \underset{\sim}{0}$.

Let $Q(\underset{\sim}{s}) = (Q_1(\underset{\sim}{s}),\ldots,Q_m(\underset{\sim}{s}))$. The first jump equation for $Q_i(\underset{\sim}{s},t)$ is obtained by considering the time at which the initial cell first leaves the cycle. Thus

(3.16) $$Q_i(\underset{\sim}{s},t) = e^{-s_i e^{ct}} \int_t^\infty g_i(u)\,du + \int_0^t h_i(Q(\underset{\sim}{s}e^{-cu},t-u))\,g_i(u)\,du \quad .$$

The m.s. convergence of $\underset{\sim}{W}_i(t)$ to $\underset{\sim}{W}_i$ implies that $Q_i(\underset{\sim}{s},t) \to Q_i(\underset{\sim}{s})$ as $t \to \infty$. So we obtain from (3.16) that

(3.17) $$Q_i(\underset{\sim}{s}) = \int_0^\infty h_i(Q(\underset{\sim}{s}\, e^{-cu}))\,g_i(u)\,du \; .$$

Note that equation (3.17) can be found in Theorem 3.2 of Mode [12]. It is obvious that $\underset{\sim}{Z}_i(t) \to \underset{\sim}{0}$ implies that $\underset{\sim}{W}_i = \underset{\sim}{0}$. We want to show that these two events are a.s. equal. This can be done by showing that

(3.18) $$P(\underset{\sim}{W}_i = \underset{\sim}{0}) = q_i$$

where q_i is the probability of extinction of $\underset{\sim}{Z}_i(t)$, i.e.

$$q_i = P(\underset{\sim}{Z}_i(t) \to \underset{\sim}{0} \quad \text{as} \quad t \to \infty).$$

If we let $P(\underset{\sim}{W}_i = \underset{\sim}{0}) = \hat{q}_i$, then, by letting $\underset{\sim}{s} \to \infty$ in (3.17), we obtain

(3.19) $$\hat{q}_i = h_i(\underset{\sim}{\hat{q}}) \quad ,$$

where $\underset{\sim}{\hat{q}} = (\hat{q}_1,\ldots,\hat{q}_m)$.

It follows from (vi) of Lemma 3.3 that $\underset{\sim}{W}_i$ is non-degenerate, so we cannot have $\hat{q}_i = 1$ for all i, and hence from Corollary of Lemma 3.4 we must have $\hat{q}_i = q_i$ for all i. This completes the proof of the following lemma.

Lemma 3.5. If the conditions of Lemma 3.3 are satisfied, then the two events $\{\underset{\sim}{Z}_i(t) \to \underset{\sim}{0} \text{ as } t \to \infty\}$ and $\{\underset{\sim}{W}_i = \underset{\sim}{0}\}$ are a.s. equal.

This lemma will be used to obtain the asymptotic joint type-phase durations distribution. Furthermore it can be shown from equation (3.17) (see Theorem 3,2, Mode [12]) that

Lemma 3.6. If the conditions of Lemma 3.3 are satisfied, then there exists a scalar random variable $\hat{W}_i$ and a positive constant K_1 such that

$$(3.20) \qquad W_i(j) = K_1 \hat{W}_i \nu_j \quad \text{for all} \quad j = 1,\ldots,m \,.$$

Now we state the last lemma of this section which has been shown in Theorem 4.1 of Mode [12].

Lemma 3.7. If the conditions of Lemma 3.3 are satisfied, then $\underset{\sim}{W}_i(t) \to \underset{\sim}{W}_i$, with probability one, as $t \to \infty$.

§4. Asymptotic Behavior of $M_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t)$

The equation (2.2) leads at once to an integral equation for $M_i^{(4)}$,

$$(4.1)\quad M_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t) = M_{ni}^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t) + \int_0^t \sum_{\nu=1}^m m_{i\nu} M_\nu^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t-u) g_i(u)\, du,$$

where $M_{ni}^{(4)}$ is easily shown (by induction) to be

$$(4.2)\quad M_{ni}^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t) =$$

$$\delta_{ii_n} m_{i_n i_{n-1}} \cdots m_{i_1 i_0} \underset{\substack{\underset{\sim}{y}_j \le \underset{\sim}{a}_j,\, j=1,\ldots,n \\ y_{0j} \le a_{0j},\, j=1,2,3}}{\int \cdots\cdots\cdots \int} f_{i_n}(\underset{\sim}{y}_n) \cdots f_{i_0}(\underset{\sim}{y}_0) \times$$

$$I_{[0, y_{04} \wedge a_{04}]}(t - y_{03} - y_{02} - \cdots - y_{n1})\, d\underset{\sim}{y}_0 \cdots d\underset{\sim}{y}_n \ .$$

If we use exactly the same arguments as Mode [11], pp. 11-14, we obtain that as $t \to \infty$

$$(4.3)\quad M_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t) \sim C_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n) e^{ct} ,$$

where

$$(4.4)\quad C_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n) = K_i^{(4)} F_n(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n),$$

$$(4.5)\quad F_n(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n) = \frac{\eta_{i_n} m_{i_n i_{n-1}} \cdots m_{i_1 i_0}}{\sum_{k=1}^m \eta_k \int \psi_{k,0}(\underset{\sim}{y})\, d\underset{\sim}{y}} \prod_{k=0}^n \Psi_{i_k,k}(\underset{\sim}{a}_k) ,$$

$$(4.6) \qquad \Psi_{\ell,k}(\underset{\sim}{a}_k) = \int_{\underset{\sim}{y} \le \underset{\sim}{a}_k} \psi_{\ell,k}(\underset{\sim}{y})\, d\underset{\sim}{y} \,,$$

$$(4.7) \qquad \psi_{\ell,0}(\underset{\sim}{y}) = e^{-c(y_1+\ldots+y_4)} \int_{y_4}^{\infty} f_\ell(y_1,y_2,y_3,u)\, du \,,$$

$$(4.8) \qquad \psi_{\ell,k}(\underset{\sim}{y}) = e^{-c(y_1+\ldots+y_4)} f_\ell(\underset{\sim}{y}),\ k = 1,2,\ldots,$$

$$(4.9) \qquad K_i^{(4)} = K\, u_i \sum_{k=1}^{m} \eta_k \int \psi_{k,0}(\underset{\sim}{y})\, d\underset{\sim}{y} / \Delta'(c) \,,$$

and c is the unique number determined in Lemma 3.1.

It is easy to show, using precisely the same argument, that as $t \to \infty$

$$(4.10) \qquad M_i^{(4)}(j,t) \sim \frac{K\mu_i \eta_j}{\Delta'(c)} \int \psi_{j,0}(\underset{\sim}{y})\, d\underset{\sim}{y}\; e^{ct} \,.$$

Hence we obtain the asymptotic behavior of the expected number of cells in phase 4 as $t \to \infty$

$$(4.11) \qquad M_i^{(4)}(t) \sim K_i^{(4)}\, e^{ct} \,.$$

Equations (4.3) and (4.11) immediately suggest the following. If a cell is selected at random from those in phase 4 at time t, then the joint distribution of the type I_k and the times $Y_{k,j}^{I_k}$, $j = 1,\ldots,4$, $k = 0,\ldots,n$, I_k is an integer between 1 and m, spent by the k^{th} ancestor of the cell of type I_k in phase j prior to time t has the following asymptotic form as $t \to \infty$:

$$(4.12) \qquad F_n(\underset{\sim}{y}_0,i_0,\ldots,\underset{\sim}{y}_n,i_n) = P(I_0 = i_0,\ \underset{\sim}{Y}_0^{i_0} \le \underset{\sim}{y}_0,\ldots,I_n = i_n,\ \underset{\sim}{Y}_n^{i_n} \le \underset{\sim}{y}_n) \,.$$

We shall again refer to the function F_n defined by equation (4.4) as the asymptotic <u>n-fold age distribution</u> of cells in phase 4. This distribution is essential in deriving the FLM function which will be discussed in Section 8.

§5. Mean Square Convergence of $Z_i^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n,t)e^{-ct}$

It is convenient to decompose the second moment $S_i^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n,t,\tau) = E[Z_i^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n,t)Z_i^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n,t+\tau)]$ in the following manner:

$$(5.1)\quad S_i^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n,t,\tau) = Q_i(t,\tau) + \sum_{k=n+1}^{\infty}\sum_{\ell=n+1}^{\infty} S_{k\ell i}^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n,t,\tau),$$

where $S_{k\ell i}^{(4)}$ was defined in Section 2 and

$$(5.2)\quad 0 \le Q_i(t,\tau) \le E[Z_{ni}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n,t)Z_i(i_0,t+\tau) Z_i(i_0,t)Z_{ni}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n,t+\tau)] .$$

From equation (5.1) we obtain the following integral equation

$$(5.3)\quad S_i^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n,t,\tau) = Q_i(t,\tau) + \int_0^t \sum_{\nu,\nu'=1}^{m} \mu_{i\nu\nu'} M_{\nu}^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n,t-u)M_{\nu'}^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n,t+\tau-u) g_i(u)\,du + \int_0^t \sum_{\nu=1}^{m} m_{i\nu} S_{\nu}^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n,t-u,\tau)g_i(u)\,du .$$

Writing $R_i(t,\tau)$ for the sum of the first two terms on the right side of equation (5.3) and letting $\overline{R}_i(t,\tau) = R_i(t,\tau)e^{-2ct-c\tau}$ it is not difficult to show that

$$(5.4)\quad \lim_{t\to\infty} \overline{R}_i(t,\tau) = \sum_{\nu,\nu'=1}^{m} \mu_{i\nu\nu'} C_{\nu}^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n)C_{\nu'}^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n) g_i^*(2c) .$$

The convergence being uniformly for $\tau \ge 0$.

Now letting $\overline{S}_i^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n,t,\tau) = S_i^{(4)}(a_0,i_0,\ldots,a_n,i_n,t,\tau)e^{-2ct-c\tau}$ and applying Theorem 2.2 of Mode [12] to equation (5.3) we obtain under the conditions of Lemma 3.3 that

$$(5.5) \qquad \lim_{t\to\infty} \underset{\sim}{\overline{S}}^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n,t,\tau) = (I - H(2c))^{-1}\underset{\sim}{\overline{R}}(t,\tau),$$

where $\underset{\sim}{\overline{S}}^{(4)}$ and $\underset{\sim}{\overline{R}}$ are column vectors with i^{th} component $\overline{S}_i^{(4)}$ and $\overline{R}_i$ respectively. If we let

$$(5.6) \qquad \nu^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n) = \eta_{i_n} m_{i_n i_{n-1}} \cdots m_{i_1 i_0} \prod_{k=0}^{n} \Psi_{i_k,k}(\underset{\sim}{a}_k),$$

where $\Psi_{i_k,k}$ was defined in equations (4.6)-(4.8), then it follows from equation (4.4) that

$$(5.7) \qquad C_i^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n) = d\mu_i\nu^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n)$$

and from (5.4) that

$$(5.8) \qquad \lim_{t\to\infty} \overline{R}_i(t,\tau) = d_i[\nu^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n)]^2 ,$$

where d_i was defined in equation (3.4).

So we see from (5.8) that the i^{th} element in the column vector in (5.5) is $c_i[\nu^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n)]^2$, where c_i was defined in equation (3.5). Thus we have

$$(5.9) \qquad \lim_{t\to\infty} \overline{S}_i^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n,t,\tau) = c_i[\nu^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n)]^2 .$$

The convergence is uniform in τ, $\tau \geq 0$.

It follows from (5.9) that

<u>Theorem 5.1</u>. If M is a matrix of positive elements such that the Perron-Frobenius root of M is greater than 1, $g_i(t) \in L^p$ for some

$p > 1$ for all $i = 1,\ldots,m$ and $\max\limits_{i,j,k} \mu_{ijk} < \infty$, and if we define

$$W_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t) = Z_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t)e^{-ct}, \text{ then}$$

(i) As $t \to \infty$, $E[W_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t+\tau) - W_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t)]^2 \to 0$ uniformly in $\tau \geq 0$,

(ii) $W_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t)$ converges in mean square to a random variable $W_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n)$ as $t \to \infty$, and

(iii) $W_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n) = \dfrac{K_i^{(4)}}{C_{i\,i_0}} F_n(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n) W_i(i_0)$ a.s.,

where $W_i(i_0)$ was defined in Lemma 3.3.

<u>Proof</u>: (i) and (ii) follow from (5.9) and completeness of L^p space. (iii) can be established by writing down the integral equation for $E[Z_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t) Z_i(i_0, t)]$ and using the same arguments as in this section to deduce from these that

$$E[C_{ii_0} W_i(i_0,t) - K_i^{(4)} F_n(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n) W_i(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t)]^2 \to 0 \text{ as } t \to \infty.$$

Since the procedures are quite similar and lengthy, the proof is omitted.

<u>Remark 5.1</u>. By using (3.20) we can rewrite (iii) of Theorem 5.1 as follows.

$$(5.10) \quad W_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n) = \frac{K_1 K_i^{(4)}}{d\mu_i} F_n(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n)\hat{W}_i \quad \text{a.s.}$$

§6. The Almost Sure Convergence of $Z_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t)e^{-ct}$

As in Section 4, Chapter 2 we decompose $Z_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t)$ as follows:

$$(6.1)\quad Z_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t) = Z_i^{E}(\underset{\sim}{a}_0', i_0, \underset{\sim}{a}_1, i_1, \ldots, \underset{\sim}{a}_n, i_n, t) -$$

$$Z_i^{e}(\underset{\sim}{a}_0', i_0, \underset{\sim}{a}_1, i_1, \ldots, \underset{\sim}{a}_n, i_n, t-a_{04}) - X_i(\underset{\sim}{a}_0', i_0, \underset{\sim}{a}_1, i_1, \ldots, \underset{\sim}{a}_n, i_n, t, a_{04}),$$

where $\underset{\sim}{a}_0'$ denotes a three-component vector (a_{01}, a_{02}, a_{03}) and $Z_i^{E}(\underset{\sim}{a}_0', i_0, \ldots, \underset{\sim}{a}_n, i_n, t)$, $Z_i^{e}(\underset{\sim}{a}_0', i_0, \ldots, \underset{\sim}{a}_n, i_n, t)$ are respectively the number of cells entering phase 4 in $[0,t]$, $[0,t)$ with ancestral phase durations $Y_{03}^{i_0}, Y_{02}^{i_0}, \ldots, Y_{n3}^{i_n}, Y_{n2}^{i_n}, Y_{n1}^{i_n}$ satisfying $\underset{\sim}{Y}_0'^{\,i_0} \le \underset{\sim}{a}_0'$, $\underset{\sim}{Y}_1^{i_1} \le \underset{\sim}{a}_1, \ldots, \underset{\sim}{Y}_n^{i_n} \le \underset{\sim}{a}_n$. Similarly $X(\underset{\sim}{a}_0', \ldots, \underset{\sim}{a}_n, i_n, t, a)$ is the number of cells entering and leaving phase 4 in $[t-a,t]$ with $\underset{\sim}{Y}_0'^{\,i_0} \le \underset{\sim}{a}_0', \ldots, \underset{\sim}{Y}_n^{i_n} \le \underset{\sim}{a}_n$.

Derivation of the asymptotic behavior of the first and second moments of these three random variables follows the pattern of Sections 4 and 5 (also see Section 4, Chapter 2) so we shall omit the procedures and simply state the following result.

Theorem 6.1. If the conditions of Theorem 5.1 are satisfied, then

$$P\{\lim_{t\to\infty} Z_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t)e^{-ct} = \frac{K_1 K_i^{(4)}}{d\mu_i} F_n(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n)\hat{W}_i \text{ for all } \underset{\sim}{a}_i \ge 0,\ i = 0,1,\ldots,n \text{ and for all } n = 0,1,2,\ldots\} = 1,$$

where K_1 and F_n were defined in (3.20) and F_n was defined in (4.5).

Similarly, under the conditions of Theorem 5.1,

$$(6.2)\quad \lim_{t\to\infty} Z_i^{(4)}(j,t)e^{-ct} = \frac{K_1 K\mu_i}{\Delta'(c)\, d\mu_i} \eta_j \int \psi_{j,0}(\underset{\sim}{y})\, d\underset{\sim}{y}\ \hat{W}_i \quad \text{a.s.}$$

(see equation (4.10)).

If we let $Z_i^{(4)}(t) = \sum_{j=1}^{m} Z_i^{(4)}(j,t)$, then it follows from (6.2) that

$$\lim_{t\to\infty} Z_i^{(4)}(t)e^{-ct} = \frac{K_1 K_i^{(4)}}{d\mu_i}\hat{W}_i \quad \text{a.s.} \tag{6.3}$$

Hence we obtain the following result.

<u>Corollary</u>. Given that $\underset{\sim}{Z}_i(t) \not\to \underset{\sim}{0}$ as $t \to \infty$, the ratio $\dfrac{Z_i^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n,t)}{Z_i^{(4)}(t)}$ converges to the joint distribution $F_n(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n)$ for all $\underset{\sim}{a}_0 \geq 0,\ldots,\underset{\sim}{a}_n \geq 0$, $i_n = 1,\ldots,m$ and for all $n = 0,1,2,\ldots$ with probability one.

<u>Proof</u>: It follows from (6.3), Theorem 6.1, Lemma 3.5 and Lemma 3.6.

<u>Remark 6.1</u>. We have assumed throughout that the process starts with a single cell of type i with age zero at epoch zero. It is not difficult to see that the preceeding Corollary remains valid regardless of the type, the phase and elapsed phase durations of the initial cell and indeed of the number of initial cells (see Section 7, Chapter 2).

§7. Random Selection of a Cell from the Population as a Whole

The Corollary of Theorem 6.1 gives the asymptotic joint distribution of $I_0 = i_0, \underset{\sim}{Y}_0^{i_0}, \dots, I_n = i_n, \underset{\sim}{Y}_n^{i_n}$ for a cell selected randomly from those in phase 4 at time t. Analogous results could clearly be derived for sampling from any particular phase. We now indicate the corresponding result for sampling from the entire cell population.

Let $Z_i^{(j)}(\underset{\sim}{a}_0^{(j)}, i_0, \underset{\sim}{a}_1, i_1, \dots, \underset{\sim}{a}_n, i_n, t)$ be the number of cells of type i_0 in phase j at time t for which the elapsed phase durations $Y_{0j}^{i_0}, \dots, Y_{01}^{i_0}, \underset{\sim}{Y}_1^{i_1}, \dots, \underset{\sim}{Y}_n^{i_n}$ of the cell itself and its first n ancestors of type i_n satisfy $Y_{0j}^{i_0} \le a_{0j}, \dots, Y_{01}^{i_0} \le a_{01}$, $\underset{\sim}{Y}_1^{i_1} \le \underset{\sim}{a}_1, \dots, \underset{\sim}{Y}_n^{i_n} \le \underset{\sim}{a}_n$. Then by the same arguments as we have used for the case $j = 4$ it can be shown that (using (3.20)), conditional on $\underset{\sim}{Z}_i(t) \not\to \underset{\sim}{0}$,

$$(7.1)\quad \frac{Z_i^{(j)}(\underset{\sim}{a}_0^{(j)}, i_0, \dots, \underset{\sim}{a}_n, i_n, t)}{Z_i(t)} \to \frac{K_i^{(j)}}{\sum_{\ell=1}^{m} C_{i\ell}} F_n^{(j)}(\underset{\sim}{a}_0^{(j)}, i_0, \dots, \underset{\sim}{a}_n, i_n)$$

for all $\underset{\sim}{a}_0^{(j)} \ge 0, \dots, \underset{\sim}{a}_n \ge 0$, $i_n = 1, \dots, m$ and for all $n = 0,1,2,\dots$ with probability 1, where

$$(7.2)\quad K_i^{(j)} = K\mu_i \sum_{k=1}^{m} \eta_k \int e^{-c(y_1+\dots+y_j)} \Big[\int_{y_j}^{\infty} [f(\underset{\sim}{y})]_{y_j=u} du\Big] d\underset{\sim}{y} / \Delta'(c),$$

$$(7.3)\quad F_n^{(j)}(\underset{\sim}{a}_0^{(j)}, i_0, \dots, \underset{\sim}{a}_n, i_n) =$$

$$\frac{\eta_{i_n} m_{i_n i_{n-1}} \cdots m_{i_1 i_0} \int_{(y_1,\dots,y_j) \le \underset{\sim}{a}_0^{(j)}} e^{-c(y_1+\dots+y_j)} \Big[\int_{y_j}^{\infty} [f(\underset{\sim}{y})]_{y_j=u} du\Big] d\underset{\sim}{y}}{\sum_{k=1}^{m} \eta_k \int e^{-c(y_1+\dots+y_j)} \Big[\int_{y_j}^{\infty} [f(\underset{\sim}{y})]_{y_j=u} du\Big] d\underset{\sim}{y}}$$

$$\times \prod_{k=1}^{n} \Psi_{i_k,k}(\underset{\sim}{a}_k),$$

and $Z_i(t) = \sum_{k=1}^{m} Z_i(j,t)$. (The function $\Psi_{i_k,k}$ was defined in Section 4.)

Equation (7.1) defines in considerable detail the asymptotic age-type-phase structure of the population as $t \to \infty$, conditional upon non-extinction. We note that $F_n^{(j)}(\underset{\sim}{a}_0^{(j)}, i_0, \ldots, \underset{\sim}{a}_n, i_n)$ defined in (7.3) gives the asymptotic joint distribution of $I_0 = i_0$, $Y_{0j}^{i_0}, \ldots, Y_{01}^{i_0}$, $I_1 = i_1$, $\underset{\sim}{Y}_1^{i_1}, \ldots, I_n = i_n$, $\underset{\sim}{Y}_n^{i_n}$ for a cell selected randomly from those in phase i at time t.

§8. Application of the Asymptotic n-fold Age Distribution to the Analysis of Labelling Experiments

For a better representation we shall use the following notation in this section. The joint probability density of the four phase durations for a new born cell of type i will be denoted $f_{i,1,2,3,4}(t_1,t_2,t_3,t_4)$. Marginal probability densities will be denoted in the usual way and the joint densities of sums of phase durations (type i is fixed) will be indicated by deletion of appropriate commas in the subscripts (see Section 8, Chapter 2). Laplace transforms will be denoted by a superscript *, e.g.

$$f^*_{i,1,23,4}(\alpha,\beta,\gamma) = \int_0^\infty\int_0^\infty\int_0^\infty e^{-\alpha\mu-\beta\nu-\gamma\omega} f_{i,1,23,4}(\mu,\nu,\omega)\,d\mu\,d\nu\,d\omega \ .$$

The Laplace transform of a function of several variables will sometimes be taken with respect to a subset of the variables. The particular subset involved will be indicated by underlining the appropriate subscripts (see Section 8, Chapter 2).

Suppose first that the label is administered continuously to a cell population (growing according to the model we have described) from time t_0 onwards. Let $C_{t_0}(t)$ be the fraction of cells in phase 4 at time $t_0 + t$ which are labelled. Assuming that t_0 is large and that the observed population is a realization of the process which does not die out, then for each fixed t as $t_0 \to \infty$, $C_{t_0}(t)$ is just the asymptotic probability that a cell selected at random from those in phase 4 entered phase 3 at time t or less prior to selection, i.e.

$$(8.1)\quad \lim_{t_0\to-\infty} C_{t_0}(t) = \sum_{i=1}^{m} \int_{a_{04}+a_{03}\le t} F_0(i, d\underset{\sim}{a}_0)$$

$$= \sum_{i=1}^{m} \frac{c\eta_i}{\sum_{k=1}^{m} \eta_k[f^*_{k,123}(c)-g^*_k(c)]} \int_0^t e^{-cy} \int_0^y [f^*_{i,\underline{12},3}(c,u)-f^*_{i,\underline{12},34}(c,u)]\,du\,dy,$$

where F_0 is the distribution function defined by equation (4.5).

Similarly for pulse labelling (administered at time t_0) the fraction of labelled cells in phase 4 at time $t_0 + t$ satisfies (under the assumptions of the previous paragraph)

$$(8.2)\qquad \lim_{t_0\to-\infty} P_{t_0}(t) = \sum_{n=0}^{\infty} P_n(t)\ ,$$

where

$$P_n(t) = \sum_{i_0=1}^{m} \cdots \sum_{i_n=1}^{m} \int_{0\le t-a_{04}-a_{03}-\cdots-a_{n4}-a_{n3}\le a_{n2}} \cdots\cdots\cdots\cdots\cdots \int F_\nu(d\underset{\sim}{a}_0, i_0, \ldots, d\underset{\sim}{a}_n, i_n)\ .$$

The probabilities $P_n(t)$, $n = 0,1,2,\ldots$ can be calculated by carrying out the straightforward integration. If we take Laplace transforms in (8.2) we find that

$$(8.3)\quad P^*_0(s) =$$

$$\frac{c\sum_{j=1}^{m} \eta_j[f^*_{j,12,3}(c,c+s)-f^*_{j,12,34}(c,c+s)-f^*_{j,1,23}(c,c+s)+f^*_{j,1,234}(c,c+s)]}{s(s+c)\sum_{k=1}^{m}\eta_k[f^*_{k,123}(c) - g^*_k(c)]}\ ,$$

$$(8.4)\quad P^*_1(s) =$$

$$\frac{c\sum_{i=1}^{m}\sum_{j=1}^{m} \eta_i m_{ij}[f^*_{i,12,34}(c,c+s)-f^*_{i,1,234}(c,c+s)][f^*_{j,123}(c+s)-g^*_j(c+s)]}{s(s+c)\sum_{k=1}^{m}\eta_k[f^*_{k,123}(c) - g^*_k(c)]}\ ,$$

and, for $n \geq 2$,

$$(8.5) \quad P_n^*(s) =$$

$$\frac{c \sum_{i=1}^{m} \sum_{i_1=1}^{m} \cdots \sum_{i_{n-1}=1}^{m} \sum_{j=1}^{m} \eta_i m_{ii_1} \cdots m_{i_{n-1}j} \, g^*_{i_1}(c+s) --- g^*_{i_{n-1}}(c+s)}{s(s+c) \sum_{k=1}^{m} \eta_k [f^*_{k,123}(c) - g^*_k(c)]} \times$$

$$[f^*_{i,12,34}(c,c+s) - f^*_{i,1,234}(c,c+s)][f^*_{j,123}(c+s) - g^*_j(c+s)].$$

According to a remark of Bellman and Cooke [3], p. 259, the Perron-Frobenius root of the matrix $H(c+s)$, $s > 0$, is strictly less than the corresponding root of the matrix $H(c)$. But the Perron-Frobenius root of the matrix $H(c)$ is 1, which means that the largest positive root of the matrix $H(c+s)$, $s > 0$, is less than 1, which implies that $H^n(c+s) \to 0$, the $m \times m$ zero matrix, as $n \to \infty$, which implies that

$$(8.6) \quad [I - H(c+s)]^{-1} = I + H(c+s) + H^2(c+s) + \ldots, \; s > 0.$$

If we denote $h^{(n)}_{ij}(c+s)$ the ij^{th} element of the matrix $H^n(c+s)$, then (8.4) and (8.5) can be rewritten in the form

$$(8.7) \quad P_n^*(s) =$$

$$\frac{c \sum_{i=1}^{m} \sum_{j=1}^{m} \eta_i h^{(n)}_{ij}(c+s)[g^*_i(c+s)]^{-1}[f^*_{i,12,34}(c,c+s) - f^*_{i,1,234}(c,c+s)][f^*_{j,123}(c+s) - g^*_j(c+s)]}{s(s+c) \sum_{k=1}^{m} \eta_k [f^*_{k,123}(c) - g^*_k(c)]}$$

Summing over $n = 0,1,2,\ldots$, and using the fact that $[I - H(c+s)]^{-1} = \frac{B(c+s)}{\Delta(c+s)}$ we obtain

$$(8.8)\quad \int_0^\infty e^{-st}\lim_{t_0\to\infty} P_{t_0}(t)dt = \frac{c(\sum_{k=1}^m \eta_k[f^*_{k,123}(c) - g^*_k(c)])^{-1}}{\Delta(s+c)s(s+c)}$$

$$\{\sum_{i=1}^m \sum_{j=1}^m \eta_i b_{ji}(s+c)[g^*_i(s+c)]^{-1}[f^*_{i,12,34}(c,s+c)-f^*_{i,1,234}(c,s+c)][f^*_{j,123}(s+c)-g^*_j(s+c)]$$

$$-\Delta(s+c)\sum_{i=1}^m \sum_{j=1}^m \eta_i[g^*_i(s+c)]^{-1}[f^*_{i,123}(c,s+c)-f^*_{i,1,234}(c,s+c)][f^*_{j,123}(s+c)-g^*_j(s+c)]$$

$$-\Delta(s+c)\sum_{j=1}^m \eta_j[f^*_{j,123}(c,c+s)-f^*_{j,12,34}(c,c+s)-f^*_{j,123}(c,c+s)+f^*_{j,1,234}(c,c+s)]\ \} .$$

<u>Example</u>. Consider the case where $M = (m_{ij})$ is a $m \times m$ matrix such that $m_{ij} = \hat{m} > \frac{1}{m}$ for every $i,j = 1,\ldots,m$, and $f_{i,1,2,3,4} = f_{1,2,3,4}$ for each $i = 1,\ldots,m$. Let $A = \hat{m}\, m > 1$. Then for every m-dimensional non-negative vector $\underset{\sim}{x} = (x_1,\ldots,x_m)$,

$$\rho_{\underset{\sim}{x}}(M) \overset{\text{def}}{=} \underset{i=1,\ldots,m}{\text{Min}} \frac{\sum_{j=1}^m m_{ij}x_j}{x_i} = \frac{\hat{m}\sum_{j=1}^m x_j}{\underset{i=1,\ldots,m}{\text{Max}} x_i} \le \hat{m}\, m = A .$$

But, for $\underset{\sim}{x} \ni x_1 = \ldots = x_m$ we have $\sum_{j=1}^m x_j / \underset{i=1,\ldots,m}{\text{Max}} x_i = m$. Thus the Perron-Frobenius root of M, $\rho(M)$, is

$$\rho(M) = \underset{\underset{\sim}{x}}{\text{Max}}\ \underset{i=1,\ldots,m}{\text{Min}} \frac{\sum_{j=1}^m m_{ij}x_j}{x_i} = \hat{m}\, m = A > 1 .$$

It is easy to check that $\Delta(\lambda) = |I - H(\lambda)| = 1 - Ag^*(\lambda)$. Thus the unique positive root c of $\Delta(\lambda) = 0$ is the unique positive root $c \ni Ag^*(c) = 1$. Now we see that $P^*_0(s)$ in (8.3) is the same as $P^*_0(s)$ in (8.6), Chapter 2 and $P^*_n(s)$, $n = 1,2,\ldots$, in (8.4) and (8.5) are the same as $P^*_n(s)$, $n = 1,2,\ldots$, in (8.7), Chapter 2. This is not a surprising result but provides a useful check on the computations.

§9. Generalization to Include Random Cell Removal

We now consider the following modification of the process considered in Sections 1-7. Suppose that in addition to elimination of cells by death at the end of the cycle there is "removal" of cells in the following manner: for any cell of type i, independently of its age or phase and independently of other cells, the probability of removal is any small time interval $(t, t+\delta t)$ is $\lambda_i \delta t + o(\delta t)$; the probability that more than one cell is removed in $(t, t+\delta t)$ is $o(\delta t)$. Suppose at epoch zero the population consists of a new born cell of type i. We consider the following two cases.

Case 1. λ_i are all equal. Let $\lambda = \lambda_i$, $i = 1,\dots,m$. It is not difficult to check that if we assume that the conditions of Theorem 5.1 are satisfied and $\lambda < c$ and if we define $\hat{Z}_i(j,t)$ to be

$$\hat{Z}_i(j,t) = e^{\lambda t} Z_i(j,t),$$

where c is the number determined in Lemma 3.1, $Z_i(j,t)$ is the number of cells of type j in the population at time t, and likewise $\hat{Z}_i^{(4)}(\underset{\sim}{a}_0, i_0, \dots, \underset{\sim}{a}_n, i_n, t)$, etc. each to be $e^{\lambda t}$ times their definitions given in Section 2, then the whole of Sections 3-8 carries over verbatim to the new process. We remark that the condition $\lambda < c$ is used at the place where we establish the m.s. convergence. Similarly, if $\lambda > c$, then the population will eventually die out (see Section 6, Chapter 2). For the critical case $\lambda = c$ the proof given in Section 3 and Sections 5-7 do not hold.

Case 2. λ_i are not all equal. The equation (2.2) again leads at once to an integral equation for $M_i^{(4)}$,

$$(9.1)\quad M_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t) = M_{ni}^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t) +$$

$$\int_0^t \sum_{\nu=1}^m m_{i\nu} M_\nu^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t-u) e^{-\lambda_i u} g_i(u)\,du \ ,$$

where

$$(9.2)\quad M_{ni}^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t) = \delta_{i i_n} m_{i_n i_{n-1}} \cdots m_{i, i_0} \int \cdots \int_A \ell_{i_n}(x_n, y_n) \cdots \ell_{i_0}(x_0, \underset{\sim}{y}_0)$$

$$I_{[0, y_{04} \wedge a_{04}]}(t - y_{03} - y_{02} - \cdots - y_{n1})\, dx_n\, d\underset{\sim}{y}_n \cdots dx_0\, d\underset{\sim}{y}_0 \ ,$$

$A = [\underset{\sim}{y}_j \le \underset{\sim}{a}_j,\ j = 1, \ldots, n,\ y_{0j} \le a_{0j},\ j = 1,2,3,$

$x_j > \sum_{k=1}^{4} y_{j,k},\ j = 1, \ldots, n,\ x_0 > t - \sum_{j=1}^{n} \sum_{k=1}^{4} y_j\ ],$

and $\ell_i(x, \underset{\sim}{y}) = \lambda_i e^{-\lambda_i x} f_i(\underset{\sim}{y}),\ i = 1, \ldots, m.$

Let $\hat{g}_i(t) = e^{-\lambda_i t} g_i(t)/g_i^*(\lambda_i)$, $\hat{m}_{ij} = g_i^*(\lambda_i) m_{ij}$, $\hat{M} = (\hat{m}_{ij})$ and $\hat{H}(s) = (\hat{g}_i^*(s)\hat{m}_{ij})$.

Remark 9.1. It is easy to check that, for some positive integer n, M^n is a positive matrix if and only if $\hat{M}^n$ is a positive matrix. Again we set $\hat{b}_{ij}(s) = (-1)^{i+j} D_{ij}(s)$, where $D_{ij}(s)$ is the complementary minor of the ij^{th} element in $I - \hat{H}(s)$, $D(s) = |I - \hat{H}(s)|$, the determinant of $I - \hat{H}(s)$, and $\hat{B}(s) = (\hat{b}_{ji}(s))$, the adjoint of the matrix $I - \hat{H}(s)$. As before the asymptotic behavior of $M_i^{(4)}$ will be associated with the roots of the determinantal equation $D(s) = 0$.

Assumption 9.1. M is a positive matrix such that the Perron-Frobenius root of $\hat{M}$ is greater than 1, $g_i(t) \in L^p$ for some $p > 1$ and for all i, and

$\max\limits_{i,j,k} \mu_{ijk} < \infty$.

If Assumption 9.1 is satisfied, then there exists a positive number α such that $D(\alpha) = 0$ (see Lemma 3.1). By using exactly the same arguments as Mode [11], pp. 11-14, we obtain that as $t \to \infty$

$$(9.3)\qquad M_i^{(4)}(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n, t) \sim C_i^{(4)} \hat{F}_n(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n) e^{\alpha t},$$

where

$$(9.4)\qquad \hat{F}_n(\underset{\sim}{a}_0, i_0, \ldots, \underset{\sim}{a}_n, i_n) = \frac{\hat{\eta}_{i_n} m_{i_n i_{n-1}} \cdots m_{i,i_0}}{\sum\limits_{k=1}^{n} \hat{\eta}_k \int \phi_{k,0}(\underset{\sim}{y}) d\underset{\sim}{y}} \prod_{k=0}^{n} \Phi_{i_k,k}(\underset{\sim}{a}_k),$$

$$(9.5)\qquad \Phi_{\ell,k}(\underset{\sim}{a}_k) = \int_{\underset{\sim}{y} \le \underset{\sim}{a}_k} \phi_{\ell,k}(\underset{\sim}{y}) d\underset{\sim}{y},$$

$$(9.6)\qquad C_i^{(4)} = k\, \hat{\mu}_i \sum_{k=1}^{m} \hat{\eta}_k \int \phi_{k,0}(\underset{\sim}{y}) d\underset{\sim}{y} / D'(\alpha),$$

and $\phi_{\ell,k}(\underset{\sim}{y})$ is the same as $\psi_{\ell,k}(\underset{\sim}{y})$ defined in equations (4.7) and (4.8) with c replaced by $\alpha + \lambda_\ell$ where α is the unique number determined under Assumption 9.1 and $\hat{\mu}_i$ and $\hat{\eta}_j$ are determined by the relation $\hat{B}(\alpha) = k(\hat{\mu}_i \hat{\eta}_j)$ (see Lemma 3.1).

We note that Assumption 9.1 can be achieved in many ways. One of those is

<u>Lemma 9.1</u>. The conditions stated in Theorem 5.1 and $c > \max\limits_{i=1,\ldots,m} \lambda_i$ imply that Assumption 9.1 is satisfied, where c is the unique number determined in Lemma 3.1.

<u>Proof</u>: We only need to show that the Perron-Frobenius root of $\hat{M}$ is greater than 1. The condition $c > \max\limits_{i=1,\ldots,m} \lambda_i$ implies that $(\hat{M})_{ij} > (H(c))_{ij}$ for every (i,j). Using the fact that the Perron-

Frobenius root of $H(c)$ is 1 and applying the theory of positive matrices (Bellman and Cooke [3], p. 259) we complete the proof of the lemma.

Using the above method to repeat the whole of Sections 5 and 6 we obtain the following result.

Theorem 9.1. If Assumption 9.1 is satisfied, then, given that $\underset{\sim}{Z}_i(t) \not\to \underset{\sim}{0}$ as $t \to \infty$, the ratio

$$\frac{Z_i^{(4)}(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n,t)}{Z_i^{(4)}(t)}$$ converges to the joint distribution

$\hat{F}_n(\underset{\sim}{a}_0,i_0,\ldots,\underset{\sim}{a}_n,i_n)$ for all $\underset{\sim}{a}_0 \ge 0,\ldots,\underset{\sim}{a}_n \ge 0$, $i_n = 1,\ldots,m$ and for all $n = 0,1,2,\ldots$ with probability one.

Using this $\hat{F}_n$ to repeat the whole of Section 8 we find that

$$(9.7)\quad \lim_{t_0 \to \infty} C_{t_0}(t) =$$

$$\sum_{i=1}^{m} \frac{\hat{\eta}_i}{\sum_{k=1}^{m} \hat{\eta}_k [f^*_{k,123}(\alpha+\lambda_k) - g^*_k(\alpha+\lambda_k)](\alpha+\lambda_k)^{-1}} \int_0^t e^{-(\alpha+\lambda_i)y} \int_0^y [f^*_{i,\underline{12},3}(\alpha+\lambda_i,u) -$$

$$f^*_{i,\underline{12},34}(\alpha+\lambda_i,u)]dudy$$

and

$$(9.8)\qquad \lim_{t_0 \to \infty} P_{t_0}(t) = \sum_{n=0}^{\infty} P_n(t),$$

where

$$(9.9)\quad P^*_0(s) =$$

$$\frac{1}{s \sum_{k=1}^{m} \hat{\eta}_k \ell^*_{k,4}(\alpha+\lambda_k)} \sum_{j=1}^{m} \frac{\hat{\eta}_j}{s+\alpha+\lambda_j} [f^*_{j,12,3}(\alpha+\lambda_j,\alpha+\lambda_j+s) - f^*_{j,12,34}(\alpha+\lambda_j,\alpha+\lambda_j+s)$$

$$- f^*_{j,1,23}(\alpha+\lambda_j,\alpha+\lambda_j+s) + f^*_{j,1,234}(\alpha+\lambda_j,\alpha+\lambda_j+s)],$$

(9.10) $P_1^*(s) =$

$$\frac{1}{s\sum_{k=1}^{m}\hat{\eta}_k\ell_{k,4}^*(\alpha+\lambda_k)}\sum_{i=1}^{m}\sum_{j=1}^{m}\hat{\eta}_i m_{ij}\ell_{j,4}^*(\alpha+\lambda_j+s)[f_{i,12,34}^*(\alpha+\lambda_i,\alpha+\lambda_i+s)$$

$$- f_{i,1,234}^*(\alpha+\lambda_i,\alpha+\lambda_i+s)] ,$$

for $n \geq 2$,

(9.11) $P_n^*(s) =$

$$\frac{}{s\sum_{k=1}^{m}\hat{\eta}_k\ell_{k,4}^*(\alpha+\lambda_k)}\sum_{i_0=1}^{m}\cdots\sum_{i_n=1}^{m}\hat{\eta}_{i_0}\Big[\prod_{\nu=0}^{n-1}m_{i_\nu i_{\nu+1}}g_{i_{\nu+1}}^*(\alpha+\lambda_{i_{\nu+1}}+s)\Big]$$

$$[g_{i_n}^*(\alpha+\lambda_{i_n}+s)]^{-1}\ell_{i_n,4}^*(\alpha+\lambda_{i_n}+s)[f_{i_0,12,34}^*(\alpha+\lambda_{i_0},\alpha+\lambda_{i_0}+s)$$

$$- f_{i_0,1,234}^*(\alpha+\lambda_{i_0},\alpha+\lambda_{i_0}+s)]$$

and

(9.12) $$\ell_{k,4}^*(s) = [f_{k,123}^*(s) - g_k^*(s)]s^{-1} .$$

We end this section by making the following remark. If λ_i are all equal to λ and $\lambda < c$, then it is not difficult to see that $c = \alpha+\lambda$ and $\hat{F}_n$ defined in (9.4) is the same as F_n defined in (4.5).

CHAPTER IV

A MODEL FOR CELLS OF THE CORNEAL EPITHELIUM

§1. Introduction

In the corneal epithelium of mice the cell population can be divided into two subpopulations, a proliferative population occupying the inner or basal layer, and a non-proliferative population occupying the outer or super layer. When a cell in the basal layer multiplies (giving birth to two new cells) one of the daughter cells remains in the basal layer while the other moves up into the super layer with probability $(1-p)$ and remains in the basal layer with probability p. Cells in the super layer do not reproduce, but cells in the basal layer do reproduce and we shall suppose that the reproductive cycle is as described in Chapter II with independent phase durations, i.e. the four phase durations T_1, T_2, T_3 and T_4 for any new-born cell are independently distributed with probability density function f_1, f_2, f_3 and f_4 respectively, and mean number of proliferative daughter cells per parent, $A = 1 + p$. The distribution of total lifetime for any cell thus has the probability density

$$g(u) = f_1 * f_2 * f_3 * f_4(u),$$

where $*$ stands for the convolution.

The departure of cells from the basal layer into the super layer appears to be a mixture of departures of daughter cells at birth and departures occurring randomly throughout the proliferative cycle (see Section 6, Chapter II). The departure mechanism is such that the population of cells in the basal layer is approximately constant. Thus the rate λ of random loss from the basal layer must be related to p in such a way as to maintain the expected number of cells in the basal layer approximately constant. In fact we shall assume that the basal layer population is a very slightly supercritical branching process and use the results derived in Chapter II to analyze an experiment used by R.J.M. Fry [5] to estimate the value of p.

A radioactive substance such as tritiated thymidine will be injected into the population at epoch t_0 to label all cells which are in phase 2 and in the basal layer at epoch t_0 and no others. Subsequently when the labelled cells divide they pass the label on to their two offspring. The cells which have migrated from the basal layer do not pick up the label, but they retain any label they had before leaving the basal layer. We are interested in the FLU function defined to be the ratio of the number of labelled cells in the super layer to the total number of labelled cells (in the basal and super layer).

The purpose of studying FLU functions is to try to estimate the parameter p. The means and variances of the phase durations for the proliferative cells in the corneal epithelium of the CF_1 Anl mouse have been estimated previously using FLM data by Brockwell, Trucco and Fry [5]. However, the FLM function is extremely insensitive to change in the value of p so that additional information, such as provided by the FLU function, is necessary for the estimation of p.

§2. FLU Function and Some Numerical Results

Again at time zero a pulse of tritiated thymidine is administered into a cell population to label the cells which are in phase 2. The experimental FLU(t) is the fraction of labelled cells in the super layer at time t. In this section we shall first derive a theoretical FLU function which is defined to be the ratio of the expected labelled cells in the super layer to the total expected number of labelled cells at time t. Then we use this function to fit the experimental data for the corneal epithelium of the CF_1 Anl mice which was prepared by R.J.M. Fry. If we let LB(t) and LU(t) be the expected number of labelled cells in the basal and super layer at time t respectively and let $LT(t) = LB(t) + LU(t)$, then

$$(2.1) \qquad FLU(t) = \frac{LU(t)}{LT(t)} = 1 - \frac{LB(t)}{LT(t)} \quad .$$

We shall also define $M(t|i,a)$ and $U(t|i,a)$ to be the expected number of cells in the basal layer and in the super layer at time t, respectively, given that at time zero there is a single cell (i,a) in phase i with age a in the basal layer. If the initial joint phase-age distribution $\xi(i,a)$, $i = 1,2,3,4$ and $a \geq 0$, is known, then it is clear that

$$(2.2) \qquad LB(t) = \int_0^\infty M(t|2,a)\xi(2,da)$$

and

$$(2.3) \qquad LU(t) = \int_0^\infty U(t|2,a)\xi(2,da) \quad .$$

The mean functions $M(t|i,a)$ and $U(t|i,a)$ are determined in Appendix A.

Since the initial phase-age distribution cannot be measured it is reasonable to take as initial distribution the limiting distribution as $t \to \infty$ if such a limiting distribution exists. When $\lambda > c$ and $p > 0$ the limiting phase-age distribution $\xi(i,a)$ exists (see Chapter II) and takes the form

$$(2.4) \qquad \xi(i,a_{0i}) = \lim_{\substack{\underset{\sim}{a}_k\uparrow\infty,k=1,\dots,n\\ a_{0j}\uparrow\infty,j=1,\dots,i-1}} \frac{K^{(i)}}{K} F_n^{(i)}(\underset{\sim}{a}_0^{(i)},\dots,\underset{\sim}{a}_n)$$

$$= \frac{c}{(A-1)\prod_{m=i}^{4} f_m^*(c)} \int_0^{a_{0i}} e^{-cy}\int_y^{\infty} f_i(u)\,du\,dy ,$$

where $K^{(i)}$ and $F_n^{(i)}$ were defined in equations (5.2)-(5.3), Chapter II.

We recall here that c is the unique positive root of the equation $Ag^*(c) = 1$ and superscript * stands for Laplace transform.

What we are going to do is to obtain FLU(t) under the assumption that λ is less than c by a very small positive number and that the initial population of cells in the basal layer has joint phase-age distribution $\xi(i,a)$ given by (2.4). Then we take $\lim_{\lambda\uparrow c} \text{FLU}(t)$ as the theoretical FLU function to analyze the experimental data. We proceed as follows.

If we denote $\overline{\text{FLU}}(t) = \lim_{\lambda\uparrow c} \text{FLU}(t)$, $\overline{\text{LB}}(t) = \lim_{\lambda\uparrow c} \text{LB}(t)$, $\overline{\text{LU}}(t) = \lim_{\lambda\uparrow c} \text{LU}(t)$ and $\overline{\text{LT}}(t) = \overline{\text{LB}}(t) + \overline{\text{LU}}(t)$, then it follows from (2.1) and the monotonicity of $M(t|2,a)$ and $U(t|2,a)$ in λ that

$$(2.5) \qquad \overline{\text{FLU}}(t) = 1 - \frac{\overline{\text{LB}}(t)}{\overline{\text{LT}}(t)} ,$$

where

$$(2.6) \qquad \overline{LB}(t) = \int_0^\infty \lim_{\lambda\uparrow c} M(t|2,a)\xi(2,da)$$

and

$$(2.7) \qquad \overline{LU}(t) = \int_0^\infty \lim_{\lambda\uparrow c} U(t|2,a)\xi(2,da) \ .$$

We note that the functions $M(t|2,a)$ and $U(t|2,a)$ actually depend on λ although not shown explicitly. It follows from equations (A.6), (A.11) and (A.12) of Appendix A and a straightforward calculation that

$$(2.8) \qquad \overline{LB}^*(s) = N_1 f_1^*(c)\left[\frac{1 - f_2^*(c)}{c(s+c)} + \frac{p}{s+c} H^*(s)\right] ,$$

$$(2.9) \qquad \overline{LU}^*(s) = N_1 f_1^*(c)\left[\frac{1 - f_2^*(c)}{s(s+c)} + \frac{s+c-ps}{s(s+c)} H^*(s)\right] ,$$

$$(2.10) \qquad \overline{LT}^*(s) = N_1 f_1^*(c) \left[\frac{1 - f_2^*(c)}{sc} + \frac{1}{s} H^*(s)\right] ,$$

$$(2.11) \quad H^*(s) = \left[f_3^*(s+c) f_4^*(s+c) \frac{f_2^*(c) - f_2^*(s+c)}{s}\right] \frac{1}{1 - Af^*(s+c)} , \ s > 0,$$

and N_1 is a fixed constant depending on the size of the population in the basal layer at time zero.

If we let T_{kj}, $j = 1,2,3,4;\ k = 0,1,2,3,\ldots$ be a sequence of independent random variables with probability densities $g_{k,j}$ defined by

$$(2.12) \quad g_{kj}(t) = \begin{cases} \dfrac{e^{-ct}\int_0^t f_4(u)du}{\int_0^\infty e^{-ct}\int_0^t f_4(u)dudt} , & \text{if } k = 0,\ j = 4, \\[2ex] \dfrac{e^{-ct}f_j(t)}{\int_0^\infty e^{-ct} f_j(t)dt} , & \text{otherwise,} \end{cases}$$

and let $E_\nu(t)$, $\nu = 0,1,2,\ldots$ be a sequence of mutually exclusive events defined by

$$E_0(t) = \{T_{04} + T_{03} \le t \text{ and } T_{04} + T_{03} + T_{02} > t\} , \tag{2.13}$$

while for $\nu \ge 1$,

$$\begin{aligned} E_\nu(t) = \{ & (\sum_{j=0}^{\nu-1} \sum_{k=1}^{4} T_{jk}) + T_{\nu 4} + T_{\nu 3} \le t \quad \text{and} \\ & (\sum_{j=0}^{\nu-1} \sum_{k=1}^{4} T_{jk}) + T_{\nu 4} + T_{\nu 3} + T_{\nu 2} > t\} , \end{aligned} \tag{2.14}$$

then the inverse functions $\overline{LB}(t)$ of $\overline{LB}^*(s)$ and $\overline{LT}(t)$ of $\overline{LT}^*(s)$ can be written as follows;

$$\overline{LB}(t) = N_1[f_1^*(c)f_2^*(c)e^{-ct} + \frac{p}{c(1+p)} \sum_{\nu=0}^{\infty} P_\nu(t)] , \tag{2.15}$$

$$\overline{LT}(t) = N_1 [f_1^*(c)f_2^*(c) + \frac{1}{c(1+p)} \sum_{\nu=0}^{\infty} P_\nu(t) + \frac{1}{1+p} \int_0^t \sum_{\nu=0}^{\infty} P_\nu(u)du] , \tag{2.16}$$

where

(2.17) $P_\nu(t)$ is the probability of occurrence of the event $E_\nu(t)$.

Now we include some numerical results for this FLU function.

Example 1. $f_j(t)$, $j = 1,\ldots,4$ are gamma type which are characterized by their means m_j and coefficients of variation c_j, $j = 1,2,3,4$. Assume that $m_1 = 87.3$ hr., $m_2 = 10.9$ hr., $m_3 = 3.5$ hr., $m_4 = 0.5$ hr., $c_1 = 0.37$, $c_2 = 0.14$, $c_3 = 0.3$, and $c_4 = 0.00$. These figures are taken from [5]. A fairly complicated Fortran program produces the following numerical results for varied values of p.

TABLE 1

p	p=0.05	p=0.2	p=0.4	p=0.5	p=0.60337
TIME (hr.)	$\overline{\text{FLU}}$	$\overline{\text{FLU}}$	$\overline{\text{FLU}}$	$\overline{\text{FLU}}$	$\overline{\text{FLU}}$
1.0	0.00048	0.00180	0.00334	0.00404	0.00472
2.0	0.00108	0.00370	0.00676	0.00814	0.00947
3.0	0.00641	0.00956	0.01339	0.01489	0.01630
4.0	0.03498	0.03484	0.03441	0.03347	0.03235
5.0	0.09025	0.08200	0.07066	0.06443	0.05790
6.0	0.15272	0.13471	0.11053	0.09824	0.08551
7.0	0.20987	0.18296	0.14721	0.12936	0.11095
8.0	0.25964	0.22513	0.17950	0.15683	0.13348
9.0	0.30320	0.26214	0.20797	0.18111	0.15347
10.0	0.34159	0.29482	0.23321	0.20269	0.17131
11.0	0.37563	0.32383	0.25569	0.22197	0.18731
12.0	0.40563	0.34943	0.27559	0.23908	0.20159
13.0	0.43092	0.37109	0.29253	0.25374	0.21394
14.0	0.45032	0.38786	0.30592	0.26552	0.22409
15.0	0.46344	0.39950	0.31567	0.27438	0.23206
16.0	0.47237	0.40679	0.32235	0.28080	0.23822
17.0	0.47544	0.41105	0.32689	0.28550	0.24312
18.0	0.47745	0.41353	0.33016	0.28919	0.24727
19.0	0.47837	0.41513	0.33279	0.29237	0.25104
20.0	0.47882	0.4634	0.33514	0.29532	0.25464
24.0	0.47979	0.42052	0.34399	0.30665	0.26860
48.0	0.48619	0.44549	0.39493	0.37104	0.34713
72.0	0.50854	0.48187	0.44879	0.43312	0.41735
73.0	0.51042	0.48411	0.45131	0.43572	0.41999
101.0	0.57598	0.55024	0.51678	0.50042	0.48367

We remark that $\overline{FLU}(t)$ is approximately linear between $t = 20$ and $t = 101$. We plot this table for some values of p in Figure 1. The circles represent the experimental data.

Example 2. Four phase durations are deterministic with $d_1 = 87.3$ hr., $d_2 = 10.9$ hr., $d_3 = 3.5$ hr. and $d_4 = 0.5$ hr. A quite simple program produces the following numerical results for varied values of p (Table 2).

We plot this table for some values of p in Figure 2.

Both Table 1 and Table 2 indicate that the FLU functions obtained from gamma phase durations and deterministic phase durations are about the same for each fixed p, $p > 0$. This is a very interesting and surprising result. Due to this fact we can, in fitting the case with gamma distributions, assume that the four phase durations are determinisitic with $d_1 = 87.3$ hr., $d_2 = 10.9$ hr., $d_3 = 3.5$ hr. and $d_4 = 0.5$ hr. and vary the value of p to obtain the best agreement between the theoretical and experimental FLU curves in the sense that the sum of aquares of the residuals is minimized. A non-linear minimization procedure program produces the following results.

Assumed parameters: $d_1 = 87.3$, $d_2 = 10.9$, $d_3 = 3.5$, $d_4 = 0.5$

Fitted parameter: $p = 0.60337$

TABLE 3

TIME (hr.)	Experimental FLU	Fitting FLU	Residual
5.0	0.025	0.05790	0.0229
17.5	0.185	0.2453	0.0603
24.0	0.262	0.2686	0.0066
48.0	0.418	0.3471	-0.0709
72.8	0.499	0.419	-0.08

For details see Table 1 and Table 2.

TABLE 2

p	p=0.05	p=0.2	p=0.4	p=0.5	p=0.60337	p=0.7
TIME (hr.)	$\overline{\text{FLU}}$	$\overline{\text{FLU}}$	$\overline{\text{FLU}}$	$\overline{\text{FLU}}$	$\overline{\text{FLU}}$	$\overline{\text{FLU}}$
1.0	0.00048	0.00178	0.00329	0.00396	0.00461	0.00518
2.0	0.00095	0.00356	0.00656	0.00790	0.00920	0.01033
3.0	0.00143	0.00534	0.00983	0.01183	0.01376	0.01546
4.0	0.00191	0.00711	0.01308	0.01574	0.01831	0.02055
5.0	0.08169	0.07436	0.06405	0.05872	0.05312	0.04781
6.0	0.14915	0.13135	0.10740	0.09538	0.08292	0.07126
7.0	0.20694	0.18027	0.14476	0.12704	0.10875	0.09169
8.0	0.25700	0.22273	0.17729	0.15468	0.13139	0.10969
9.0	0.30078	0.25992	0.20589	0.17905	0.15142	0.12570
10.0	0.33940	0.29279	0.23124	0.20070	0.16929	0.14006
11.0	0.37372	0.32203	0.25388	0.22009	0.18535	0.15304
12.0	0.40441	0.34823	0.27422	0.23756	0.19988	0.16485
13.0	0.43203	0.37184	0.29261	0.25339	0.21311	0.17566
14.0	0.45701	0.39322	0.30933	0.26783	0.22521	0.18561
15.0	0.47757	0.41091	0.32335	0.28006	0.23562	0.19435
16.0	0.47782	0.41196	0.32557	0.28291	0.23914	0.19852
17.0	0.47807	0.41301	0.32779	0.28575	0.24265	0.20267
18.0	0.47832	0.41406	0.33000	0.28858	0.24614	0.20680
19.0	0.47857	0.41510	0.33220	0.29139	0.24961	0.21091
20.0	0.47882	0.41615	0.33439	0.29420	0.25307	0.21499
24.0	0.47981	0.42030	0.34310	0.30531	0.26675	0.23113
48.0	0.48574	0.44460	0.39301	0.36841	0.34370	0.32121
72.0	0.49159	0.46787	0.43913	0.42577	0.41257	0.40073
73.0	0.49184	0.46882	0.44097	0.42804	0.41528	0.40384
101.0	0.49858	0.49470	0.49020	0.48818	0.48622	0.48450

FIGURE 1

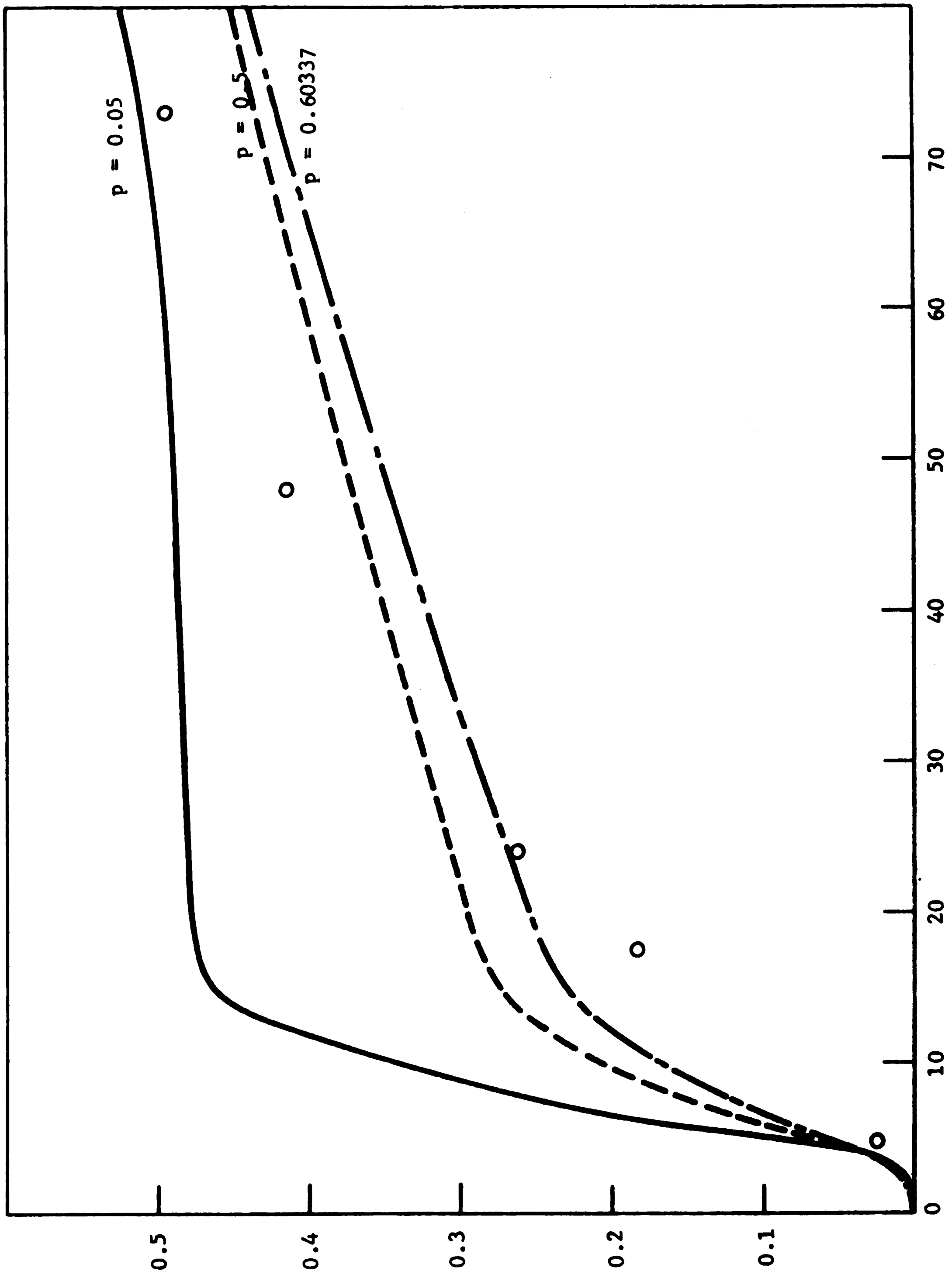
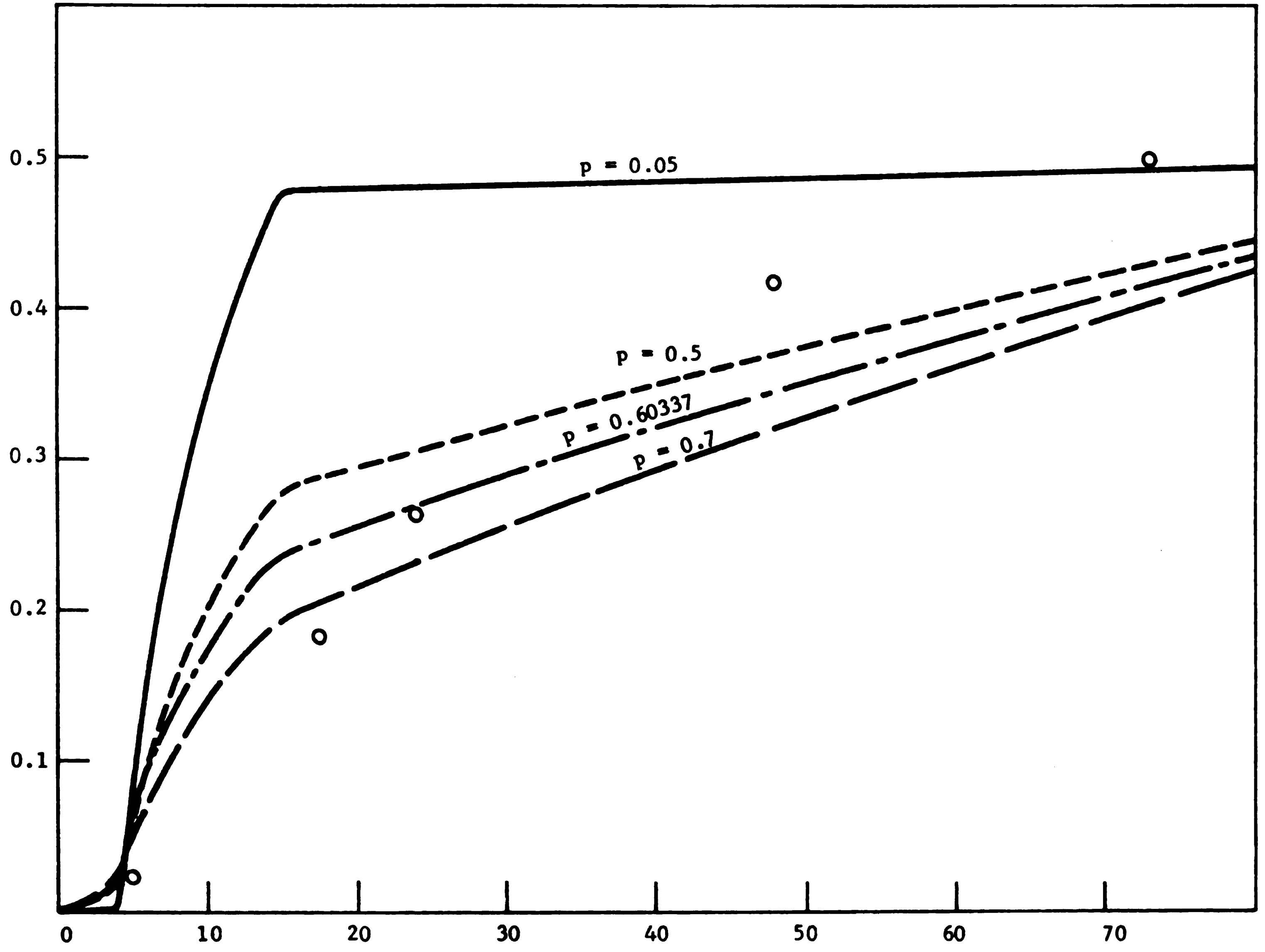


FIGURE 2



We summarize the results obtained in this section as follows:

1. The FLU functions obtained by using deterministic phase durations and by using gamma phase durations are almost identical over the range of parameters relevant to our problem. The deterministic FLU function (which is much easier to compute) was therefore used to simplify the fitting problem.
2. The FLU function is very sensitive with respect to the value of p.
3. The fit is not as good as could be hoped but it seems to be compatible with experimental errors and indicates that p is not either 0 or 1 as has been assumed by various authors.

For the experiment in which the label is administered continuously to a cell population, we are also able to calculate the CLU(t) defined to be the ratio of the expected number of labelled cells in the super layer to the total expected number of labelled cells at time t. This CLU function is presented in Appendix B. We devote Appendix C to the discussion of the asymptotic behavior of $M(t|i,a_0)$ and the random variable $Z^s(t|i,a_0)$.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Barrett, J.C. A mathematical model of the mitotic cycle and its application to the interpretation of percentage labelled mitoses data. J. Nat'l. Cancer Inst., 37, 443-450, 1966.

2. Barrett, J.C. Optimized parameters for the mitotic cycle. Cell and Tissue Kinetics 3, 349-353, 1970.

3. Bellman, R.E. and Cooke, K.L. "Differential-difference equations", Princeton Univ. Press, Princeton, New Jersey, 1963.

4. Brockwell, P.J. and Trucco, E. On the decomposition by generations of the PLM-function. J. Theoret. Biol., 26, 149-179, 1970.

5. Brockwell, P.J., Trucco, E. and Fry, R.J.M. The determination of cell-cycle parameters from measurements of the fraction of labelled mitoses. Bull. Math. Biophys., 34, 1-12, 1972.

6. Feller, W. "An Introduction to Probability Theory and its Applications", 2nd ed., Vol. II, John Wiley and Sons, Inc., New York, 1971.

7. Harris, T.E. "The Theory of Branching Processes", Springer Verlag, Berlin, 1963.

8. Karlin, S. "A First Course in Stochastic Processes", Academic Press, New York and London, 1968.

9. Macdonald, P.D.M. Statistical inference from the fraction labelled mitoses curve. Biometrika, 57, 489-503, 1970.

10. Mendelsohn, M.L. and Takahashi, M. A critical evaluation of the fraction of labelled mitoses method as applied to the analysis of tumor and other cycles. Chapter III of The Cell Cycle and Cancer (R. Baserga, ed.), Marcel Dekker, Inc., New York, 1972.

11. Mode, C.J. A multidimensional age-dependent branching process with applications to natural selection I. Math. Biosci. 3, 1-18, 1968.

12. Mode, C.J. A multidimensional age-dependent branching process with applications to natural selection II. Math. Biosci. 3, 231-247, 1968.

13. Mode, C.J. A generalized multidimensional age-dependent branching process. ORNL-4331, Oak Ridge National Laboratory.

14. Powell, E.P. Growth rate and generation time of bacteria, with special reference to continuous culture. J. Gen. Microbiol., 15, 492-511, 1955.

15. Steel, G.G. and Hanes, S. The technique of labelled mitoses: analysis by automatic curve fittings. Cell Tissue Kinet., 4, 93-105, 1970.

16. Takahashi, M. Theoretical basis for cell cycle analysis II. J. Theoret. Biol., 18, 195-209, 1968.

17. Trucco, E. and Brockwell, P.J. Percentage labelled mitoses curves in exponentially growing cell populations. J. Theoret. Biol., 20, 321-337, 1968.

APPENDICES

APPENDIX A

INTEGRAL EQUATIONS FOR THE MEAN FUNCTIONS $M(t|i,a)$ AND $U(t|i,a)$

Define $Z(t|i,a)$ to be the number of cells in the basal layer at time t, given that at time zero there is a single cell (i,a) in phase i with age a in the basal layer. Let $Z_\nu(t|i,a)$ be the number of ν^{th} generation of the initial cell (i,a) in the basal layer at time t. We shall denote $M(t|i,a) = E\,Z(t|i,a)$ and $M_\nu(t|i,a) = E\,Z_\nu(t|i,a)$.

Consideration of the time at which the initial cell first leaves the cycle leads to the following integral equations for the function $M_\nu(t|i,a)$.

(A.1) $M_0(t|i,a) = P(\text{the initial cell } (i,a) \text{ is itself in the basal layer at time } t)$

$$= e^{-\lambda t}\int_t^\infty y(i,a,u)\,du$$

and, for $\nu \geq 1$,

$$M_\nu(t|i,a) = A\int_0^t M_{\nu-1}(t-u|1,0)\,e^{-\lambda u}\,y(i,a,u)\,du\ , \tag{A.2}$$

where

$$y(i,a,u) = r(i,a,\cdot) * f_{i+1} * \ldots * f_4(u)\ , \tag{A.3}$$

$$r(i,a,t) = f_i(a+t)\Big/\int_a^\infty f_i(u)\,du,\ t \geq 0, \tag{A.4}$$

and A and f_i, $i = 1,2,3,4$ are defined in Section 1, Chapter IV.

Taking Laplace transform on both sides of equations (A.1) and (A.2) and using the equality

$$\text{(A.5)} \qquad M(t|i,a) = \sum_{\nu=0}^{\infty} M_{\nu}(t|i,a)$$

we obtain

$$\text{(A.6)} \quad M^*(s|i,a) = \frac{1 - y^*(i,a,s+\lambda)}{s+\lambda} + \frac{1 - g^*(s+\lambda)}{s+\lambda}\,\frac{y^*(i,a,s+\lambda)}{g^*(s+\lambda)} \sum_{\nu=1}^{\infty} (Ag^*(s+\lambda))^{\nu},$$

where superscript * stands for Laplace transform and $g^*(s) = \prod_{i=1}^{4} f_i^*(s)$.

Now define $Z^s(t|i,a)$ to be the number of cells in the super layer at time t, given that at time zero there is a single cell (i,a) in phase i with age a in the basal layer. Let $Z_{\nu}^s(t|i,a)$ be the number of ν^{th} generation of the initial cell (i,a) in the super layer at time t. We shall also denote $U(t|i,a) = E\, Z^s(t|i,a)$ and $U_{\nu}(t|i,a) = E\, Z_{\nu}^s(t|i,a)$. Consideration of the time at which the initial cell first leaves the basal layer leads to the following integral equations.

$$\text{(A.7)} \quad U_0(t|i,a) = P(\text{the initial cell } (i,a) \text{ is itself in the super layer at time } t)$$

$$= \int_0^t \lambda e^{-\lambda u} du \int_u^{\infty} y(i,a,\theta)\, d\theta \;,$$

$$\text{(A.8)} \quad U_1(t|i,a) = (1-p)\int_0^t e^{-\lambda u} y(i,a,u)\, du + A\int_0^t U_0(t-u|1,0)e^{-\lambda u} y(i,a,u)\, du,$$

and, for $\nu \geq 2$,

$$\text{(A.9)} \qquad U_{\nu}(t|i,a) = A\int_0^t U_{\nu-1}(t-u|1,0)e^{-\lambda u} y(i,a,u)\, du \;.$$

Again taking Laplace transform on both sides of equations (A.7) - (A.9) and using the relation

(A.10) $$U(t|i,a) = \sum_{\nu=0}^{\infty} U_{\nu}(t|i,a)$$

we obtain

(A.11) $$U^{*}(s|i,a) = \frac{\lambda}{s(s+\lambda)} + [\frac{1-p}{s} + A\,U^{*}(s|1,0) - \frac{\lambda}{s(s+\lambda)}]y^{*}(i,a,s+\lambda)$$

and

(A.12) $$U^{*}(s|1,0) = (\frac{1-p}{s}\,g^{*}(s+\lambda) + \frac{\lambda}{s(s+\lambda)}\,[1 - g^{*}(s+\lambda)])\sum_{\nu=0}^{\infty}[Ag^{*}(s+\lambda)]^{\nu}.$$

APPENDIX B

CLU FUNCTION

Suppose that the label is administered continuously to a cell population. Let CLU(t) be the ratio of the expected number of labelled cells in the super layer to the total expected number of labelled cells at time t. As in Section 2, Chapter IV we shall assume that λ is less than c by a very small number and that the initial population of cells in the basal layer has joint phase-age distribution $\xi(i,a)$ given by equation (2.4), Chapter IV. If we define $LB(t|i,a)$ and $LU(t|i,a)$ to be the expected number of labelled cells in the basal layer and in the super layer at time t respectively, given that at time zero there is a single cell (i,a) in phase i with age a in the basal layer, then it is clear that

$$\text{(B.1)} \qquad CLU(t) = \frac{LU(t)}{LT(t)} = 1 - \frac{LB(t)}{LT(t)} \quad ,$$

where

$$\text{(B.2)} \qquad LB(t) = \sum_{i=1}^{4} \int_{0}^{\infty} LB(t|i,a)\xi(i,a)$$

and

$$\text{(B.3)} \qquad LU(t) = \sum_{i=1}^{4} \int_{0}^{\infty} LU(t|i,a)\xi(i,a) \; .$$

Again we are interested in the function $\overline{CLU}(t) = \lim_{\lambda \uparrow c} CLU(t)$. It is clear that

$$\overline{CLU}(t) = 1 - \frac{\overline{LB}(t)}{LT(t)} , \tag{B.4}$$

where

$$\overline{LB}(t) = \sum_{i=1}^{4} \int_0^\infty \lim_{\lambda \uparrow c} LB(t|i,a)\xi(i,da) \tag{B.5}$$

and

$$\overline{LU}(t) = \sum_{i=1}^{4} \int_0^\infty \lim_{\lambda \uparrow c} LU(t|i,a)\xi(i,da) \quad . \tag{B.6}$$

It remains to determine $LU(t|i,a)$ and $LB(t|i,a)$. It is obvious that

$$LB(t|2,a) = M(t|2,a) \tag{B.7}$$

and

$$LU(t|2,a) = U(t|2,a) , \tag{B.8}$$

where M and U were defined in Chapter IV (also in Appendix A).

Consideration of the time at which the initial cell first leaves the phase i leads to the following integral equations.

$$LB(t|1,a) = \int_0^t M(t-u|2,0)e^{-\lambda u}r(1,a,u)\,du, \tag{B.9}$$

$$LB(t|3,1) = \int_0^t LB(t-u|4,0)e^{-\lambda u}r(3,a,u)\,du, \tag{B.10}$$

$$LB(t|4,a) = A\int_0^t LB(t-u|1,0)e^{-\lambda u}r(4,a,u)\,du, \tag{B.11}$$

$$LU(t|1,a) = \int_0^t U(t-u|2,0)e^{-\lambda u}r(1,a,u)\,du, \tag{B.12}$$

$$LU(t|3,a) = \int_0^t LU(t-u|4,0)e^{-\lambda u}r(3,a,u)\,du \tag{B.13}$$

and

$$(B.14) \qquad LU(t|4,a) = A\int_0^t LU(t-u|1,0)e^{-\lambda u}r(4,a,u)du,$$

where $r(i,a,u)$ was defined in (A.4), Appendix A.

Taking Laplace transform on both sides of equations (B.7), (B.9), (B.10) and (B.11) we obtain

$$(B.15) \qquad LB^*(s|1,a) = M^*(s|2,0)r^*(1,a,s+\lambda),$$

$$(B.16) \qquad LB^*(s|2,a) = M^*(s|2,a),$$

$$(B.17) \qquad LB^*(s|3,a) = A\,M^*(s|2,0)f_1^*(s+\lambda)f_4^*(s+\lambda)r^*(3,a,s+\lambda)$$

and

$$(B.18) \qquad LB^*(s|4,a) = A\,M^*(s|2,0)f_1^*(s+\lambda)r^*(4,a,s+\lambda) \quad .$$

It follows from equation (A.6), Appendix A and a straightforward calculation that

$$(B.19) \quad N_2\overrightarrow{LB}^*(s) = \frac{f_1^*(c)(1 - f_2^*(c))}{c(s+c)} + \frac{f_1^*(c)}{s(s+c)}[1-Af_1^*(s+c)f_2^*(c)f_3^*(s+c)f_4^*(s+c)$$
$$+P\ f_2^*(c)f_3^*(s+c)f_4^*(s+c)],$$

where N_2 is a fixed constant depending on the initial population size.

Similarly, equations (B.8), (B.12), (B.13), (B.14) and equations (A.11) and (A.12) of Appendix A lead to the explicit form of $\overline{LU}(t)$ in terms of Laplace transform.

$$(B.20) \quad N_2\overrightarrow{LU}^*(s) = \frac{f_1^*(c)(1-f_2^*(c))}{s(s+c)} + \frac{f_1^*(c)}{s^2(s+c)}\{C-ACf_1^*(s+c)f_3^*(s+c)f_4^*(s+c)$$
$$+ [(s+c)-ps]f_2^*(c)f_3^*(s+c)f_4^*(s+c)\} \quad .$$

Thus we have

$$(B.21)\quad N_2\overline{TL}^*(s) = \frac{f_1^*(c)(1-f_2^*(c))}{cs} + \frac{f_1^*(c)}{c}[1-Af_1^*(s+c)f_2^*(c)f_3^*(s+c)f_4^*(s+c) + f_2^*(c)f_3^*(s+c)f_4^*(s+c)]\ .$$

We note that it is very easy to invert $\overline{LB}^*(s)$ and $\overline{LT}^*(s)$ into $LB(t)$ and $LT(t)$ respectively.

APPENDIX C

ASYMPTOTIC BEHAVIOR OF THE FUNCTIONS $M(t|i,a)$ AND $U(t|i,a)$

The notation to be used in this Appendix is the same as those in Chapter IV and in Appendix A. Observe that the number of cells in the basal layer in the time interval $[0,t]$ is non-decreasing in t and finite for every finite t. Thus $\sup_{0\le\tau\le t} M(\tau|i,a) < \infty$ for every finite t. This leads to the following obvious lemma.

Lemma C.1. If $M(t|i,a) \to C_1$ asd $t \to \infty$ and C_1 is a positive constant then

(i) $\int_0^\infty M(t|i,a)dt = \infty$

(ii) $\exists\ C_2$ independent of $t \ni M(t|i,a) \le C_2$.

Now we are ready to state the following result.

Theorem C.1. A necessary and sufficient condition for the function $M(t|i,a)$ converge to a constant $K(i,a)$ as $t \to \infty$ is that $\lambda = c$. Then $K(i,a) > 0$ and

$$(C.1) \qquad K(i,a) = \frac{c^{-1}(A-1)r^*(i,a,c) \prod_{m=i+1}^{4} f_m^*(c)}{A \int_0^\infty ye^{-cy}g(y)\,dy} .$$

Proof: (i) It follows from equations (A.1), (A.2) and (A.5), Appendix A that

$$(C.2) \quad M(t|i,a) = e^{-\lambda t}\int_t^\infty y(i,a,u)du + A\int_0^t M(t-u|1,0)e^{-\lambda u}y(i,a,u)du .$$

Setting $i = 1$ and $a = 0$ in equation (C.2), multiplying both sides of equation (C.2) by $e^{\lambda t}$ and letting $\hat{M}(t|1,0) = M(t|1,0)e^{\lambda t}$ (this is the method suggested in Section 6, Chapter II) we obtain a renewal equation for the function $\hat{M}(t|1,0)$. By applying the asymptotic theory of the renewal equation we establish the result that $\lambda = c$ is a sufficient condition for $M(t|i,a)$ converging to $K(i,a)$ as $t \to \infty$. (ii) Suppose that $\lim_{t\to\infty} M(t|i,a) = K(i,a)$. If $\lambda > c$, then it follows from (A.6), Appendix A that

$$M^*(s|i,a_0) \overset{s\to0}{\to} \frac{1}{\lambda}[1-y^*(i,a,\lambda)] + \frac{1-g^*(\lambda)}{\lambda} y^*(i,a_0,\lambda) \frac{A}{1-Ag^*(s)} < \infty .$$

This contradicts (i) of Lemma C.1 (see Feller [6], pp. 418-420). Thus λ must be less than or equal to c. If $p = 0$, i.e. $A = 1$ and $c = 0$, then we have completed the proof of the theorem. If $p > 0$, i.e. $c > 0$, it remains to rule out the possibility $\lambda > c$. Suppose $\lambda > c$. We can find a positive s_0 such that $\lambda+s_0 < c$. For such a s_0 we have that $\sum_{\nu=0}^{\infty} [Ag^*(s_0+\lambda)]^\nu = \infty$, i.e. $M^*(s_0|i,a_0) = \infty$. But for any $s > 0$,

$$\begin{aligned} M^*(s|i,a_0) &= \int_0^T e^{-st}M(t|i,a_0)dt + \int_T^\infty e^{-st}M(t|i,a_0)dt \\ &< \int_0^T e^{-st}M(t|i,a_0)dt + (K(i,a_0) + \epsilon) \int_T^\infty e^{-st}dt < \infty, \end{aligned}$$

where ϵ, T are chosen properly from the assumption that $\lim_{t\to\infty} M(t|i,a_0) = K(i,a_0)$. So we cannot have $\lambda < c$. This completes the proof of the theorem.

<u>Theorem C.2</u>. If $\lambda = c$, then

$$U(t|i,a) \sim K_1 Ay^*(i,a,c)[(1-p)g^*(c) + (1-g^*(c))]t, \ t \to \infty .$$

Proof: It follows from equations (A.7) - (A.10), Appendix A that

$$\text{(C.3)}\quad U(t|i,a) = \int_0^t \lambda e^{-\lambda u}du \int_u^\infty y(i,a,\theta)d\theta + (1-p)\int_0^t e^{-\lambda u}y(i,a,u)du$$
$$+ A\int_0^t U(t-u|1,0)e^{-\lambda u}y(i,a,u)du \quad .$$

Taking Laplace transform on both sides of equation (C.3) we obtain

$$\text{(C.4)}\quad U^*(s|i,a) = \frac{\lambda}{s(s+\lambda)} + \left[\frac{1-p}{s} + AU^*(s|1,0) - \frac{\lambda}{s(s+\lambda)}\right]y^*(i,a,s+\lambda) .$$

Setting $i = 1$ and $a = 0$ in equation (C.4) and solving for $U^*(s|1,0)$ we obtain

$$\text{(C.5)}\quad U^*(s|1,0) = \left(\frac{1-p}{s}\, g^*(s+\lambda) + \frac{\lambda}{s(s+\lambda)}\,[1-g^*(s+\lambda)]\right)\frac{1}{1-Ag^*(s+\lambda)}, \quad s > 0 .$$

By letting $s \to 0$ in (C.5) we find that

$$s^2U^*(s|1,0) \to K[(1-p)g^*(c) + (1-g^*(c))]$$

which implies

$$\text{(C.6)}\quad s^2U^*(s|i,a) \to K_1Ay^*(i,a,c)[(1-p)g^*(c) + (1-g^*(c))] \quad \text{as} \quad s \to 0$$

and

$$\text{(C.7)}\quad \frac{U^*(\tau s|i,a\ )}{U^*(\tau|i,a\ )} \to \frac{1}{s} \quad \text{as} \quad \tau \to 0 \ ,$$

where $K_1 = [A\int_0^\infty te^{-ct}g(t)dt]^{-1}$.

It follows from standard Tauberian theorem (e.g. Feller [6], pp. 418-420) that

$$\text{(C.8)}\quad \int_0^t U(y|i,a)dy \sim \frac{1}{2}\, K_1Ay^*(i,a,c)[(1-p)g^*(c)+(1-g^*(c))]t^2, \quad t \to \infty .$$

Since $U(t|i,a)$ is non-decreasing in t, the Lemma on p. 422, Feller [6] can be applied to complete the proof of the theorem.

We also state the following two obvious theorems.

Theorem C.3. If $\lambda > c$, then

(i) $U(t|i,a) \to \dfrac{1+y^*(i,a,\lambda)-Ag^*(\lambda)}{1-A\,g^*(\lambda)}$ as $t \to \infty$

(ii) $Z^s(t|i,a) \to W^s(i,a)$ as $t \to \infty$, where W^s is a random variable.

(iii) $E\,W^s(i,a) = \lim_{t\to\infty} U(t|i,a)$

(iv) $P(W^s(i,a) < \infty) = 1$.

Proof: Proceed exactly the same arguments used in Section 6, Chapter II.

Theorem C.4. If $p > 0$, $c > \lambda$ and $f \in L^q$ for some $q > 1$, then

(i) $U(t|i,a) \sim C(i,a)e^{(c-\lambda)t}$, $t \to \infty$ where

$$C(i,a) = K\left(\frac{\lambda}{c-\lambda}\,\frac{1-g^*(c)}{c} + \frac{1-p}{c-\lambda}\,g^*(c)\right)A\,y^*(i,a,c)$$

(ii) $Z^s(t|i,a)/c(i,a)e^{(c-\lambda)t}$ converges to a random variable $W^s(i,a)$ with probability one as $t \to \infty$

(iii) $E\,W^s(i,a) = 1$

(iv) $\mathrm{Var}\,W^s(i,a) > 0$.