DEMOGRAPHY Volume 6, Number 2 May 1969

A DISCRETE-TIME TWO-SEX AGE-SPECIFIC STOCHASTIC POPULATION PROGRAM INCORPORATING MARRIAGE

J. H. Pollard
School of Biological Studies, Macquarie University, North Ryde, N.S.W. 2113 Australia

Abstract—A discrete-time two-sex stochastic population model is developed. All entities (single males, single females, or couples) are grouped according to their ages, and during a unit time interval, each entity has a choice of several outcomes with fixed conditional probabilities. The model assumes that the number of marriages between men aged x and women aged y is equal to the minimum of the number of men aged x desiring marriage with a woman aged y and the number of women aged y desiring marriage with a man aged x. It follows that if a large excess of males of all ages is maintained in the population, the female component grows as a multi-type Galton-Watson process. Under such circumstances, the females have perfect freedom in their choice of marriage partner, and the use of a multi-type Galton-Watson process is very realistic. The same result is true for the male component of the population. If there are no males (or females), no marriages take place, so the model is realistic on this score also. A complex computer program is described, and a detailed numerical example given.

In 1966, a unisexual age-specific discrete-time stochastic model for projecting human populations was developed. This model evolved from an earlier discrete-time deterministic model due to H. Bernardelli (1941), E. G. Lewis (1942) and P. H. Leslie (1945), but it may be regarded as a special case of the multi-type Galton-Watson process (T. E. Harris, 1963). It was developed from the population mathematics viewpoint, but several generalizations were given (J. H. Pollard, 1966). Many of the techniques described are useful for analyzing the present two-sex model, and we therefore begin with a summary of earlier results and include a few extensions of these results.

The two-sex model is developed in discrete time, and entities (single males, single females, or couples) are grouped according to their ages. During a unit time interval, each entity of a particular type has fixed conditional probabilities of following various possible outcomes, and, except for marriage, the outcome followed determines the number of entities due to that entity at the end of the time interval. The number of marriages between single males aged x and single females aged y is equal to the minimum of the number of males aged x desiring marriage with a female aged y, and the number of females aged y desiring marriage with a male aged x.

The process is very similar to a multi-type Galton-Watson process with a small amount of interaction between certain of the entities. As a model for monogamous human populations, the process has certain desirable features: the model ensures that if a large excess of males of all ages is maintained in a population, the females have perfect choice in selecting

their marriage partners, and the female component of the population grows as a multi-type Galton-Watson process; a similar result applies to the males; also this model (in contrast with certain other two-sex models) allows no marriages to take place if no males (females) exist.

Many mathematical models exist for human populations, but none of them are suitable for detailed projection purposes without certain, rather subjective adjustments in the calculations; the two-sex model described in this paper avoids many of these difficulties. The demographer is frequently faced with the problem of investigating the effect on a population of a change in marriage rates, or of divorce rates, or due to changes in economic conditions, or due to changes in government immigration policy, etc. It is possible with this model to carry out objective numerical investigations of such problems on digital computers. However it does not seem possible to derive interesting asymptotic results, such as those obtained using the simpler mathematical models.

A computer program of some generality has been developed to use this model for projection purposes, and a numerical example is given. One important fact emerges from the numerical calculations: the probabilities themselves must be considered as random variables in any realistic population model.

1. Introduction

In constructing mathematical models for human populations ". . . it has usually been found convenient to ignore numerical differences between the two sexes, and to discuss only the growth of the female population, the male component being supposed to adjust its numbers accordingly" (D. G. Kendall, 1949). Under ideal circumstances, these unisexual population models should represent the population quite accurately. However, in practice, numerical differences and age structure differences between the two sexes are important, and must be borne in mind when analyzing a population. Furthermore, the various one-sex models, when applied to the two sexes separately, usually lead to incompatible results.

Various bisexual deterministic theories have been brought forward (e.g. P. H. Karmel, 1947; A. H. Pollard, 1948). A two-sex stochastic theory presents a very difficult problem, and so far only a few simplified models have been analyzed. D. G. Kendall [(1949), section 2, (ix)] mentions the problem of the two sexes and suggests a few different approaches:

(1) Births $\propto$ men $\times$ women (unstable population; explosion);
(2) Births $\propto$ $\sqrt{\text{men} \times \text{women}}$ (geometric mean);
(3) Births $\propto$ (men + women) (somewhat unrealistic); and
(4) Births $\propto$ min (men, women) (perhaps the most realistic)

Kendall's work inspired L. A. Goodman (1953) to extend his ideas further. However, in both these discussions an age-structure was ignored. This is clearly an oversimplification.

In this paper, we describe a discrete-time two-sex stochastic population model. All entities (single males, single females, or couples) are grouped according to their ages, and during a unit time interval, each entity has a choice of several outcomes with fixed conditional probabilities. *Except for the problem of marriage,* these considerations would lead us to a multi-type Galton-Watson process, and the results of an earlier paper (J. H. Pollard, 1966) would apply. Our model will assume that the number of marriages between men aged x and women aged y is equal to the *minimum* of the number of men aged x desiring marriage with women aged y and the number of women aged y desiring marriage with men aged x.

This model ensures that if a large excess of males of all ages is maintained in a population, the female component of the population will grow as a multi-type Galton-Watson process. Similarly, if a large excess of females of all ages is maintained in a population, the male component (ignoring illegitimate births) will grow as a multi-type Galton-Watson process. Under such circumstances, the females (or in the latter case, the males) have perfect freedom in their choice of marriage partner, and the use of a multi-type Galton-Watson process is very realistic.

If there are no males (or females), no marriages take place, so the model is realistic on this score also. It should be noted that deterministic means and stochastic means are *not* equal for this type of model, so a stochastic analysis must be used.

Many results published in an earlier paper (J. H. Pollard, 1966) are required in §6 to analyze the two-sex model. These are therefore summarized in §2, and some extensions are given in §3, §4 and §5. A numerical example using the population projection program is described in some detail in §6.9.

One important fact emerges from the numerical example: the calculated variances are much smaller than observed variances with actual population data, even when time trends in the probabilities are taken into account. The additional variability must be due to random fluctuations in the probabilities themselves. Many mathematical demographers do not realize the importance of this source of variability, although Z. M. Sykes (1967) noted the smallness of the variances.

2. A Summary of Some Earlier Results

In 1966, the author listed the moments of the numbers of the various types for a multi-type Galton–Watson process in a column vector dimension $(k + k^2 + \cdots + k^n)$, where k is the number of types, and n the highest order moment required. The moments were listed in this vector $\mathbf{m}(t)$ in increasing degree and dictionary order, and it was shown that $\mathbf{m}(t)$ obeyed a linear recurrence relation over time of the form:

$$\mathbf{m}(t+1) = \mathbf{TMBFm}(t). \qquad (1)$$

This linear recurrence relation was derived by examining the diagrammatic representation of such a process. Consider, for example, the simple two-type process described in Figure 1.

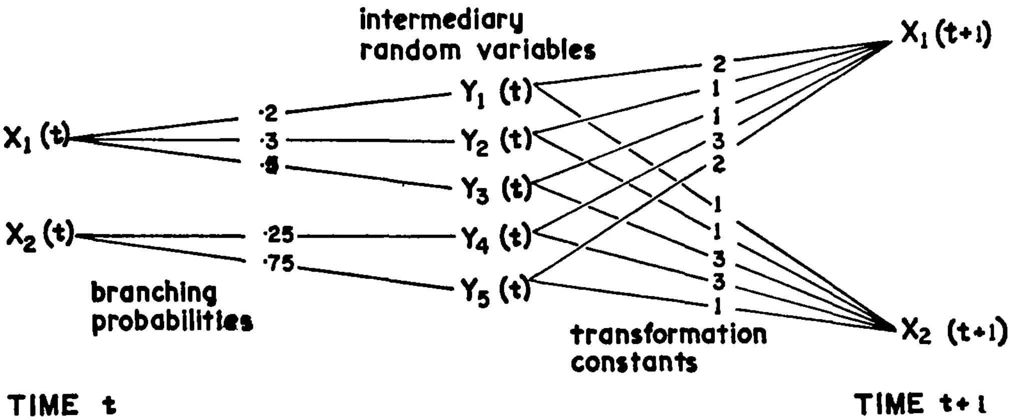

Fig. 1.—Diagrammatic Representation of a Simple Two-type Galton-Watson Process

During a unit time interval $(t, t + 1)$, each individual of type 1 has three alternatives with fixed multinomial probabilities 0.2, 0.3, and 0.5. If the individual follows the first alternative (with probability 0.2), there will be two individuals of type 1 and one individual of type 2 at time $t + 1$ corresponding to the single individual of type 1 at time t. Similarly, each individual of type 2 has two alternatives during the time interval with probabilities 0.25 and 0.75. If such an individual follows the first alternative (with probability 0.25) there will be three individuals of type 1 and three individuals of type 2 at time t. All the individuals in the process act independently.

The basic steps in the argument for deriving equation (1) are the following:

(1) The transformation from moments about the origin to falling factorial moments is linear. The moments about the origin are listed in the column vector $\mathbf{m}(t)$, so the factorial moments are listed appropriately in a column vector $\mathbf{F}\ \mathbf{m}(t)$.

(2) The factorial moments of order n of the intermediary random variables $\{Y_i\}$ are linear functions of the factorial moments of order n of the random variables $\{X_i(t)\}$ at time t. It follows that the factorial moments of the intermediary random variables $\{Y_i\}$ are listed in a vector $\mathbf{BFm}(t)$.

(3) The ordinary moments of the intermediary random variables $\{Y_i\}$ are linear functions of the factorial moments of the $\{Y_i\}$, so the ordinary moments of the intermediary random variables are listed appropriately in a vector $\mathbf{MBFm}(t)$.

(4) The vector random variable at time $t + 1$ is a linear transformation of the vector of intermediary random variables. So the moments at time t are listed appropriately in the vector $\mathbf{m}(t + 1)$ defined by equation (1).

The forms of matrices **T** and **B** are given in the above reference. For the two-type example depicted in Figure 1, for example, we define

$$\mathbf{P} = \begin{bmatrix} .2 & 0 \\ .3 & 0 \\ .5 & 0 \\ 0 & .25 \\ 0 & .75 \end{bmatrix}$$

and

$$\mathbf{Q} = \begin{bmatrix} 2 & 1 & 1 & 3 & 2 \\ 1 & 1 & 3 & 3 & 1 \end{bmatrix}. \tag{2}$$

The non-zero elements of **P** are the conditional multinomial probabilities for the individuals involved in the process. Matrix **Q** is made up of linear transformation constants.

It is necessary to define the Kronecker product of two matrices **W** and **Z**. Let $\mathbf{W} = (W_{ij})$ and $\mathbf{Z} = (Z_{ij})$ be matrices of dimension $\ell \times m$ and $r \times s$ respectively. Then the Kronecker product of **W** and **Z** is denoted by $\mathbf{W} \times \mathbf{Z}$ and is defined by

$$\mathbf{W} \times \mathbf{Z} = \begin{bmatrix} W_{11}\mathbf{Z} & W_{12}\mathbf{Z} & \cdots & W_{1m}\mathbf{Z} \\ W_{21}\mathbf{Z} & W_{22}\mathbf{Z} & \cdots & W_{2m}\mathbf{Z} \\ \cdot & \cdot & \cdot & \cdot \\ W_{\ell 1}\mathbf{Z} & W_{\ell 2}\mathbf{Z} & \cdots & W_{\ell m}\mathbf{Z} \end{bmatrix}$$

which is a matrix of dimension $\ell r \times ms$.

It is now possible to write down **T** and **B**:

$$\mathbf{T} = \begin{bmatrix} \mathbf{Q} & & & \\ & \mathbf{Q} \times \mathbf{Q} & & \\ & & \mathbf{Q} \times \mathbf{Q} \times \mathbf{Q} & \\ & & & \ddots \end{bmatrix}$$

and

$$\mathbf{B} = \begin{bmatrix} \mathbf{P} & & & \\ & \mathbf{P} \times \mathbf{P} & & \\ & & \mathbf{P} \times \mathbf{P} \times \mathbf{P} & \\ & & & \ddots \end{bmatrix}. \tag{3}$$

Complete details about the matrices **F** and **M**, however, were not given. It was merely stated that **F** had the form

$$\mathbf{F} = \begin{bmatrix} \mathbf{I} & & \\ \mathbf{F}_{21} & \mathbf{I} \times \mathbf{I} & \\ \mathbf{F}_{31} & \mathbf{F}_{32} & \mathbf{I} \times \mathbf{I} \times \mathbf{I} \\ \cdot & \cdot & \cdot \end{bmatrix}, \quad (4)$$

and that **M** had a similar form. Let us now consider **F** in some detail. Because of the redundant method of writing down the moments in the vector $\mathbf{m}(t)$, the form of the matrix is *not* unique, and indeed some of the $\mathbf{F}_{ij}$ submatrices have an infinite number of possible forms. We describe here the form generated by the computer program for TITAN, the computer of the Mathematical Laboratory at the University of Cambridge. It is perhaps the most elegant form.

The submatrix $\mathbf{F}_{ij}$ is of dimension $(k^i) \times (k^j)$, where k is the number of types in the branching process. The rows of this submatrix may be represented by numbers of the form:

$$\begin{matrix} 1 & 1 & 1 & \cdot & \cdot & \cdot & 1 & 1 & 1 \\ 1 & 1 & 1 & \cdot & \cdot & \cdot & 1 & 1 & 2 \\ 1 & 1 & 1 & \cdot & \cdot & \cdot & 1 & 1 & 3 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 1 & 1 & \cdot & \cdot & \cdot & 1 & 1 & k \\ 1 & 1 & 1 & \cdot & \cdot & \cdot & 1 & 2 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ k & k & k & \cdot & \cdot & \cdot & k & k & k \end{matrix} \quad (i \text{ digits}) \quad (5)$$

The columns of this submatrix may be represented by similar numbers, except that these will be j digits:

$$\begin{matrix} 1 & 1 & 1 & \cdot & \cdot & \cdot & 1 & 1 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ k & k & k & k & k & k & k & k & k \end{matrix} \quad (j \text{ digits}) \quad (6)$$

F has no non-zero submatrices above the diagonal. Further, we have fixed the form of the on-diagonal blocks; therefore $i > j$. We now wish to obtain the value of an element in the $\mathbf{F}_{ij}$ submatrix. It is possible to expand its row number in the form of (5) above, and then count the number of 1's, the number of 2's, $\cdots$, the number of k's. Let these numbers be $I_1, I_2, \cdots, I_k$ respectively.

Similarly, the column number of the element may be expressed in the form of (6) above, and we may then count the number of 1's, the number of 2's, $\cdots$, the number of k's. Let us call these numbers $J_1, J_2, \cdots, J_k$ respectively. Clearly,

$$i = \sum_{n=1}^{k} I_n, \quad \text{and} \quad j = \sum_{n=1}^{k} J_n. \quad (7)$$

The element of the submatrix $\mathbf{F}_{ij}$ may be shown to be

$$\frac{1}{j!} \prod_{n=1}^{k} s(I_n, J_n)(J_n)! \quad (8)$$

[$s(I_n, J_n)$ is a Stirling number of the first kind (J. Riordan, 1958, p. 32)]. To see that this is true, consider for example

$$\begin{aligned} &\mathcal{E}[U(U-1)(U-2) \cdots (U-m+1) \\ &\quad \cdot V(V-1)(V-2) \cdots (V-n+1)] \\ &= \mathcal{E}\left\{\left[\sum_{k=0}^{m} s(m, k)U^k\right]\left[\sum_{l=0}^{n} s(n, l)V^l\right]\right\} \\ &= \sum_{k=0}^{m} \sum_{l=0}^{n} s(m, k)s(n, l)\,\mathcal{E}[U^k V^l] \end{aligned}$$

The expectation $\mathcal{E}[U^k V^l]$ occurs $(k+l)!/k!l!$ times in the section of the moment vector corresponding to moments of order $(k+l)$. So the $(k+l)!/k!l!$ elements in submatrix $\mathbf{F}_{m+n,k+l}$ corresponding to the expectations $\mathcal{E}[U^k V^l]$ are $s(m, k)k!s(n, l)l!/(k+l)!$ The generalization of this result is expression (8).

This expression may be regarded as a general form for an element of any submatrix $\mathbf{F}_{ij}$ of **F**. If $j > i$, at least one J_n

will be greater than the corresponding I_n and the Stirling number $s(I_n, J_n)$ will be zero. The element is therefore zero. To obtain the elements of the diagonal blocks, we must consider $i = j$; formula (8) does not, however, yield the useful submatrices $\mathbf{I} \times \mathbf{I}, \mathbf{I} \times \mathbf{I} \times \mathbf{I}, \cdots$.

The results for the submatrix $\mathbf{M}_{ij}$ of the matrix $\mathbf{M}$ are strictly analogous. The dimensions of $\mathbf{M}$ are $(K^i) \times (K^j)$, where K is the number of conditional branching probabilities. Replacing k by K in the above argument, we obtain

$$i = \sum_{n=1}^{K} I_n, \quad \text{and} \quad j = \sum_{n=1}^{K} J_n. \tag{9}$$

The element of the submatrix $\mathbf{M}_{ij}$ may be shown to be

$$\frac{1}{j!} \prod_{n=1}^{K} S(I_n, J_n)(J_n)! \tag{10}$$

[$S(I_n, J_n)$ is a Stirling number of the second kind (J. Riordan, 1958, p. 32)]. The comments made about formula (8) also apply to formula (10).

Frequently, expectations and quadratic moments are the only moments of interest. Indeed, these are the only moments required in §6 to analyze the two-sex model. The matrix $\mathbf{F}$ then has a very simple form:

$$\mathbf{F} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{F}_{21} & \mathbf{I} \times \mathbf{I} \end{bmatrix}. \tag{11}$$

The rows of $\mathbf{F}_{21}$ may be expressed as number pairs $(1, 1), (1, 2), \cdots, (1, k), \cdots, (k, k)$ as in (5) and the columns of $\mathbf{F}_{21}$ may be denoted by single numbers $1, 2, \cdots, k$. Then all the elements of $\mathbf{F}_{21}$ are zero, *except* the element in the (j, j) row and the j column $(j = 1, 2, \cdots, k)$. This element is minus one.

For expectations and quadratic moments, $\mathbf{M}$ too has a simple form:

$$\mathbf{M} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{M}_{21} & \mathbf{I} \times \mathbf{I} \end{bmatrix}. \tag{12}$$

All the elements of $\mathbf{M}_{21}$ are zero, *except* the element in the (j, j) row and the j column $(j = 1, 2, \cdots, K)$. This element is one.

These results for first and second order moments are *very* useful computationally. We may list expectations and second order moments in the vector $\mathbf{m}(t)$. It is *not* necessary to store the matrix $\mathbf{F}$, since premultiplication by $\mathbf{F}$ is equivalent to subtracting each expectation from the corresponding second order (squared) moment. It is not necessary to store $\mathbf{B}$, only $\mathbf{P}$, and $\mathbf{P}$ may be stored in a very compact form. Premultiplication of $\mathbf{Fm}(t)$ by $\mathbf{B}$ is straightforward. It is not necessary to store $\mathbf{M}$, since premultiplication by $\mathbf{M}$ is equivalent to the addition of each expectation (of an intermediary random variable) to the corresponding second order (squared) moment. $\mathbf{Q}$ needs to be stored (often in a compact form) but *not* $\mathbf{T}$, and premultiplication by $\mathbf{T}$ is easily achieved. Programming the moment analyses for such processes is straightforward, and numerical results have been obtained in this manner in several different contexts (e.g. D. J. Bartholomew, 1968, pp. 51–55; J. H. Pollard, 1968a).

It has been shown (J. H. Pollard, 1966) that if only first and second order moments are being considered, the moment recurrence relation (1) applies to expectations and quadratic moments about the origin. This result simplifies computation still further.

3. Multi-Type Galton-Watson Processes With Random Branching Probabilities

One possible generalization of the usual multi-type Galton-Watson process is obtained by assuming that the conditional branching probabilities are themselves random variables. The probabilities, as random variables, are assumed independent of the other random variables which represent numbers of individuals.

It is not difficult to conceive of situations in which this type of model is applicable. Consider, for example, the population model analyzed by the author in 1966. It is a well-known fact that mortality rates depend upon weather conditions: a severe winter will cause mortality rates (especially at the older ages, and at the very young ages) to rise; conversely, a mild winter will mean that the mortality rates experienced are lighter than usual. Thus there may be occasions when it is reasonable to consider the mortality probabilities as random variables. [One could also consider the linear transformation "constants" in matrix **Q** as random variables; however, from the point of view of constructing population models, there does not seem to be a case for doing so.]

Branching process calculations performed with fixed conditional probabilities and large populations usually lead to variances considerably smaller than those encountered in practical situations. This fact has been noted by Z. M. Sykes (1967). The additional variability is usually due to fluctuations in the probabilities themselves.

A numerical example is instructive. Consider 1,000,000 persons subject to a mortality rate q_x, where q_x has expected value .002 and standard deviation .0001. The variance in the number of deaths due to the finite size of the population is 1,000,000 × .002 × .998, equal to 1,996, whereas the variance in the number of deaths due to fluctuations in the mortality rate q_x is approximately $(1{,}000{,}000)^2 \times (.0001)^2$, equal to 10,000. Thus the total variability arises from two main sources: (i) statistical fluctuations due to the finite population size; and (ii) fluctuations in the conditional probabilities themselves. With large populations, the second source of variation is often the greater, but it is usually neglected by mathematical demographers.

When stochastic fluctuations in the probabilities are taken into account, the linear recurrence relation (1) is changed only slightly, and takes the form

$$\mathbf{m}(t+1) = \mathbf{TM}\,\mathcal{E}(\mathbf{B})\mathbf{Fm}(t) \qquad (13)$$

This result was proved by J. H. Pollard (1968b). For this type of model,

$$\mathcal{E}(\mathbf{P} \times \mathbf{P}) \neq \mathcal{E}(\mathbf{P}) \times \mathcal{E}(\mathbf{P}),$$

and consequently, the linear recurrence relation (13) applies only to moments about the origin and *not* to central quadratic moments.

4. Some Stochastic Processes Permitting Analyses Similar to that of the Galton-Watson Process

It has been shown that all multi-type Galton-Watson processes may be represented by diagrams like that in Figure 1. The intermediary random variables $\{Y_j(t)\}$ conditional on the random variables $\{X_j(t)\}$ are multinomial random variables, and the random variables $\{X_j(t+1)\}$ are linear multiples of the intermediary random variables.

It is possible to construct other stochastic processes using different conditional distributions. Some of these will have linear moment recurrence relations over time similar to equation (1).

Example 1. Consider Figure 1, and let

$Y_1(t) \mid X_1(t)$ be a Poisson random variable with mean $.2\, X_1(t)$;
$Y_2(t) \mid X_1(t)$ be a Poisson random variable with mean $.3\, X_1(t)$;
$Y_3(t) \mid X_1(t)$ be a Poisson random variable with mean $.5\, X_1(t)$;
$Y_4(t) \mid X_2(t)$ be a Poisson random variable with mean $.25\, X_2(t)$;
$Y_5(t) \mid X_2(t)$ be a Poisson random variable with mean $.75\, X_2(t)$.

These conditional Poisson distributions are mutually independent. The random variables $\{X_j(t+1)\}$ are obtained from the $\{Y_j(t)\}$ by linear transformations, and the transformation constants are non-negative. If the transformation constants are integers, it is soon apparent that the process is a special multi-type Galton-Watson process with an infinite number of conditional branching probabilities. An examination of the moments of the conditional random variables reveals that a linear moment recurrence relation exists for this type of process, and it has the form:

$$\mathbf{m}(t+1) = \mathbf{TMBm}(t) \quad (14)$$

The matrices **T**, **M**, and **B** are the same as those defined in §2, and all the results of §2 and §3 may be applied to this process.

Example 2. Consider the gamma density

$$f_\alpha(y) = e^{-y}y^{\alpha-1}/\Gamma(\alpha)$$

$Y_1(t) \mid X_1(t)$ has the gamma density with $\alpha = .2\ X_1(t)$;

$Y_2(t) \mid X_1(t)$ has the gamma density with $\alpha = .3\ X_1(t)$;

$Y_3(t) \mid X_1(t)$ has the gamma density with $\alpha = .5\ X_1(t)$;

$Y_4(t) \mid X_2(t)$ has the gamma density with $\alpha = .25\ X_2(t)$;

$Y_5(t) \mid X_2(t)$ has the gamma density with $\alpha = .75\ X_2(t)$.

These conditional gamma distributions are mutually independent. The random variables $\{X_j(t+1)\}$ are obtained from the $\{Y_j(t)\}$ by linear transformations, and the transformation constants are non-negative. An examination of the moments of the intermediary random variables reveals that a linear recurrence relation exists for the moments in this type of process, and it is of the form:

$$\mathbf{m}(t+1) = \mathbf{TEFEBm}(t) \quad (15)$$

The matrices **T**, **F** and **B** have their usual forms, and **E** (which is used for conversion between rising and falling factorial moments) is defined by

$$\mathbf{E} = \begin{bmatrix} (-\mathbf{I}) & 0 & 0 \\ 0 & (-\mathbf{I}) \times (-\mathbf{I}) & 0 \\ 0 & 0 & (-\mathbf{I}) \times (-\mathbf{I}) \times (-\mathbf{I}) \\ . & . & . & . \end{bmatrix}.$$

For this model, the type random variables $\{X_j(t)\}$ may take any non-negative values, not necessarily integral. All the results of §2 and §3 may be applied to the process.

Example 3. This example is obtained by considering the negative multinomial distribution (W. Feller, 1957). The distribution is obtained by considering the numbers of the various types of failure in a multinomial situation before obtaining exactly r successes. Let the probability of success at each trial be p, and the probability of a failure of type j at each trial be p_j, so that $p + \Sigma_{j=1}{}^n\, p_j = 1$.

The probability of k_1 failures of type 1, k_2 failures of type 2, . . . , k_n failures of type n, before exactly r successes is equal to

$$P_r(k_1, k_2, \cdots, k_n) = \frac{(r + \sum k_i - 1)!}{k_1! \cdots k_n!\,(r-1)!}\, p_1^{k_1} \cdots p_n^{k_n} p^r. \quad (16)$$

If r is set equal to $X_1(t)$, we may construct a trivariate distribution for $Y_1(t)$, $Y_2(t)$ and $Y_3(t)$ in Figure 1 by allowing these random variables to assume values k_1, k_2 and k_3 respectively, according to the above distribution. The random variables $Y_4(t)$ and $Y_5(t)$ are similarly-defined conditional random variables: r is set equal to $X_2(t)$, and probabilities p', p_1' . . . , p_n' replace the probabilities p, p_1, . . . , p_n in formula (16).

An examination of the moments of the $\{Y_j(t)\}$ reveals that a linear moment recurrence relation exists for the process, and it has the form:

$$\mathbf{m}(t+1) = \mathbf{TMBEFEm}(t). \quad (17)$$

The matrices **T**, **M**, **E** and **F** have their usual forms, and **B** is modified slightly: .2, .3, .5, .25, and .75 must be replaced by p_1/p, p_2/p, p_3/p, p_1'/p' and p_2'/p' respectively. Once again, all the results of §2 and §3 may be applied.

Many other models, permitting the same type of analysis, are possible. It should be noted that the product matrix in equations (14), (15) and (17) is always of the form:

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} & \mathbf{0} & \cdot \\ \mathbf{C}_{21} & \mathbf{A}\times\mathbf{A} & \mathbf{0} & \cdot \\ \mathbf{C}_{31} & \mathbf{C}_{32} & \mathbf{A}\times\mathbf{A}\times\mathbf{A} & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}.$$

5. Immigration

In two earlier papers (J. H. Pollard, 1966, 1967) techniques for dealing with immigration have been discussed. In both cases, the number and age-structure of immigrants are assumed independent of the overall population. The basic moment recurrence relation (1) is then modified to

$$\mathbf{m}(t+1) = \mathbf{TMBFm}(t) + \mathbf{r}(t+1), \quad (18)$$

where $r(t+1)$ is the immigration vector of moments. These methods are easily adapted and incorporated in the two-sex model of §6. Although not discussed in detail in §6, immigration may be readily incorporated in the two-sex analysis.

6. The Two-Sex Population Model

We consider at discrete points of time $t = 0, 1, 2, \ldots$ a population composed of three types of entity: single men, single women, and couples. The single men and the single women are grouped into age groups corresponding to the unit intervals of time. The couples are grouped according to the pair of ages (on the same discrete age-scale). [Thus, for example, an artificially simple population might be composed of the following entities: men aged 0, men aged 1, men aged 2; women aged 0, women aged 1, women aged 2; and four types of couple with age pairs (1, 1), (1, 2), (2, 1) and (2, 2).]

Consider first a single man aged x. During a unit time interval, he has various possible alternatives:

(1) die;
(2) merely survive to be aged $x+1$, and not marry;
(3) *wish* to marry a woman aged y_1, and survive;
(4) *wish* to marry a woman aged y_2, and survive;
(5) etc. (for the other marriage possibilities).

The outcome he follows is determined by fixed conditional multinomial probabilities.

A single woman aged y has similar possibilities, but in addition the possibility of an illegitimate birth [There is no theoretical difficulty in including multiple births. However, because one confinement in about eighty results in a multiple birth, and we assume a reasonably small time unit, we shall ignore them (J. H. Pollard, 1966).]:

(1) die;
(2) have an illegitimate son and survive;
(3) have an illegitimate daughter and survive;
(4) merely survive to be aged $y+1$;
(5) *wish* to marry man aged x_1 and survive;
(6) *wish* to marry man aged x_2 and survive;
(7) etc. (for the other marriage possibilities).

A married couple, husband aged x,

wife aged y, has the following possibilities during a unit time interval:

(1) merely survive as a couple;
(2) husband die and wife survive to be a single woman aged $y + 1$;
(3) wife die and husband survive to be a single man aged $x + 1$;
(4) divorce and both survive;
(5) son born and couple survives;
(6) daughter born and couple survives.

[For reasons given above, multiple births have been ignored.]

For single men, single women, and couples, certain possibilities involving probabilities of smaller order have been ignored [e.g. for a couple, the possibility of the husband dying and a son being born during the same time interval]. There is no theoretical difficulty in including these possibilities, and indeed, they should be included if the probabilities are appreciable.

All the outcomes listed above (except the "*wish* to marry" outcomes) immediately determine the numbers of the various entities at time $t + 1$ in a multitype Galton-Watson fashion. The only difficulty is caused by marriage: the model assumes that the number of marriages between men aged x and women aged y is equal to the *minimum* of the number of men aged x desiring marriage with a woman aged y and the number of women aged y desiring marriage with a man aged x.

It is clear that the entities could be further subdivided according to social class, race, duration of marriage, number of previous children, whether unmarried, widowed or divorced, etc. No theoretical difficulties arise, but computational and data difficulties will crop up. The computational difficulties may soon be a thing of the past with the large computers of the (near) future. As long as a single male (female) in category x may be assumed to have a fixed conditional probability of marrying a single female (male) from category y when there is a large excess of females (males) in all categories, this type of model is applicable. The probabilities of desiring marriage must be independent of the numbers of entities in the population.

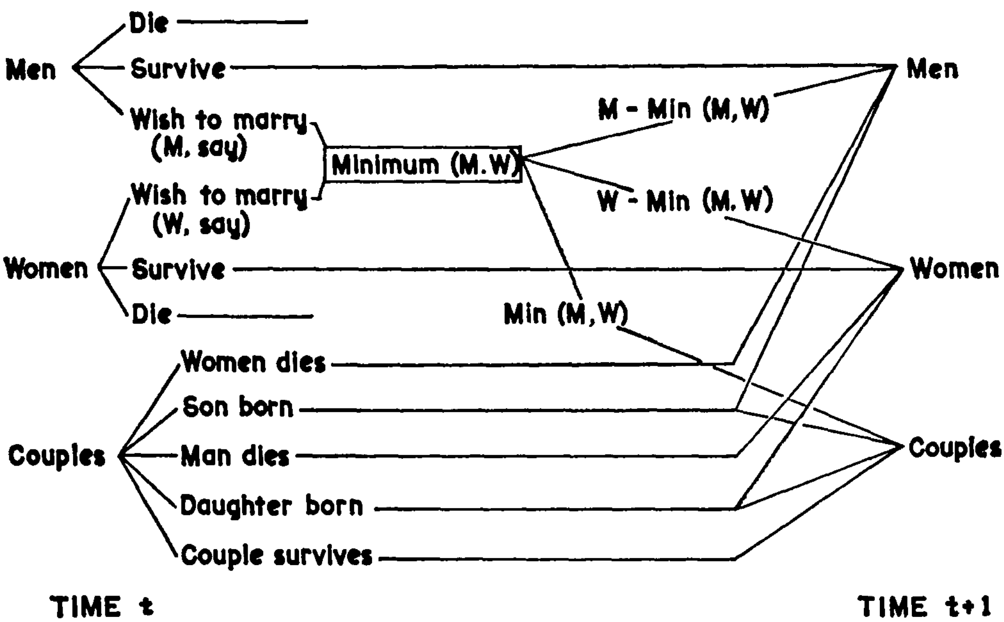

FIG. 2.—The Two-sex Model with No Age Structure

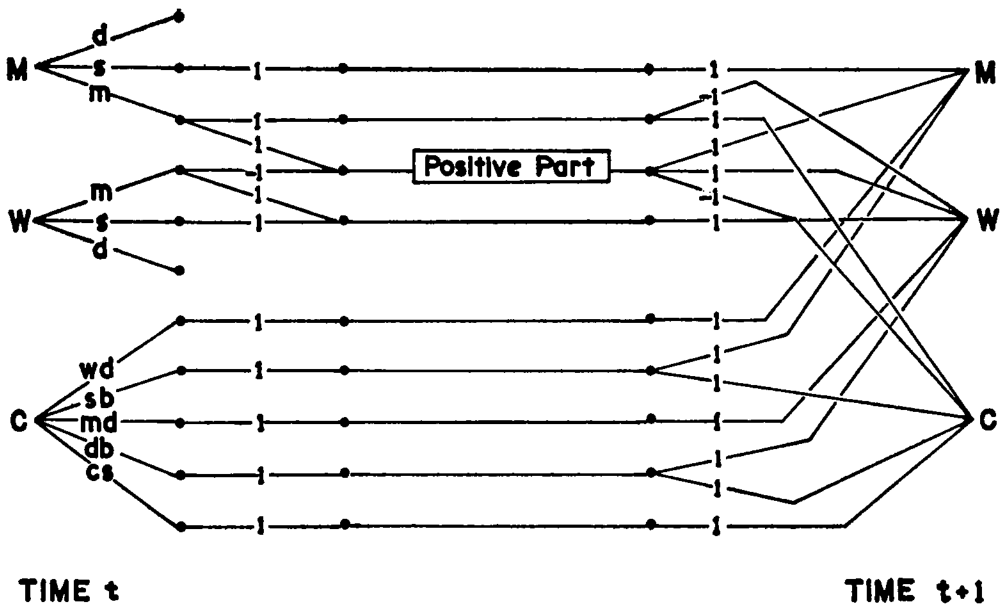

FIG. 3.—An Alternative Representation of the Two-sex Model with No Age Structure

6.1 The Principal Difficulty

To simplify the discussion in this section and in some of the following sections, age-structure, divorce and illegitimate births will be ignored. It is then possible to represent the model diagrammatically as in Figure 2. Representing this type of population with age-structure diagrammatically is almost impossible, but not necessary because it is possible to discuss the more complicated cases using the simplified diagram in Figure 2. An alternative representation is given in Figure 3 and it is soon apparent that the two processes are identical. The representation given in Figure 3 is the more useful form, and the one used throughout §6.

It is clear that the techniques discussed in §2 are useful for analyzing stages 1, 2 and 4 of the process in Figure 3. The only stage requiring a different treatment is stage 3 when the moments of the positive part of a random variable need to be computed. [It should be noted in passing that for a multi-type Galton-Watson process, the linear transformation constants are non-negative integers; the techniques we use are applicable for any real linear transformations, but only make sense in the present context if they are integers (positive or negative).]

In the case of the one-sex stochastic model, a linear recurrence relation was derived for expectations and central quadratic moments. Ideally, we should like to derive a recurrence relation for the expectations and quadratic moments in the two-sex model, or alternatively produce a numerical recurrence method for these moments. There is one major difficulty however: to obtain a recurrence method for the two-sex model, it is necessary to know something about the distributions of *some* of the random variables at stage 3 in Figure 3; such knowledge was not necessary for the multi-type Galton-Watson recurrence relation. The moment recurrence method for the two-sex model can be written symbolically as

$$\mathbf{m}(t+1) = \mathbf{T}_2 * \mathbf{T}_1\mathbf{MBFm}(t).$$

All the symbols have their usual meaning (§2), and * represents the moment process which occurs as the positive part of a random variable is taken.

Consider two random variables X and Y with expectations μ_1 and μ_2 respectively, variances σ_1^2 and σ_2^2 respectively, and covariance $\rho\sigma_1\sigma_2$. We define $X+$ equal to the positive part of X. (i.e., $X+$ is equal to X if X is positive, and equal to zero if X is negative or zero.) The problem is then the following: *knowing these moments, how accurately can we compute the first and second order moments of $X+$ and Y?*

The following points should be noted:

(1) It seems that the expectation and variance of X^+ will vary very little for a wide range of possible distributions of X, all having the same first two moments. This is to be expected, because moments are averages. [This point is discussed in some detail in § 6.2.]

(2) When $\mu_1 > 3\sigma_1$ (say), the expected value and variance of X^+ are approximately μ_1 and σ^2_1 respectively.

(3) When $\mu_1 < -3\sigma_1$ (say), the expected value and variance of X^+ are both approximately zero.

(4) For the time interval (0, 1) in the two-sex model, the positive part taken is that of the difference between two binomial random variables. The difference is approximately normal for large populations.

(5) For most populations we consider, the difference random variables which have their positive parts taken are usually small compared with the other random variables involved. When this is not so, results (2) and (3) above usually apply.

6.2 The Effect of the Distribution of X on the Moments of X+

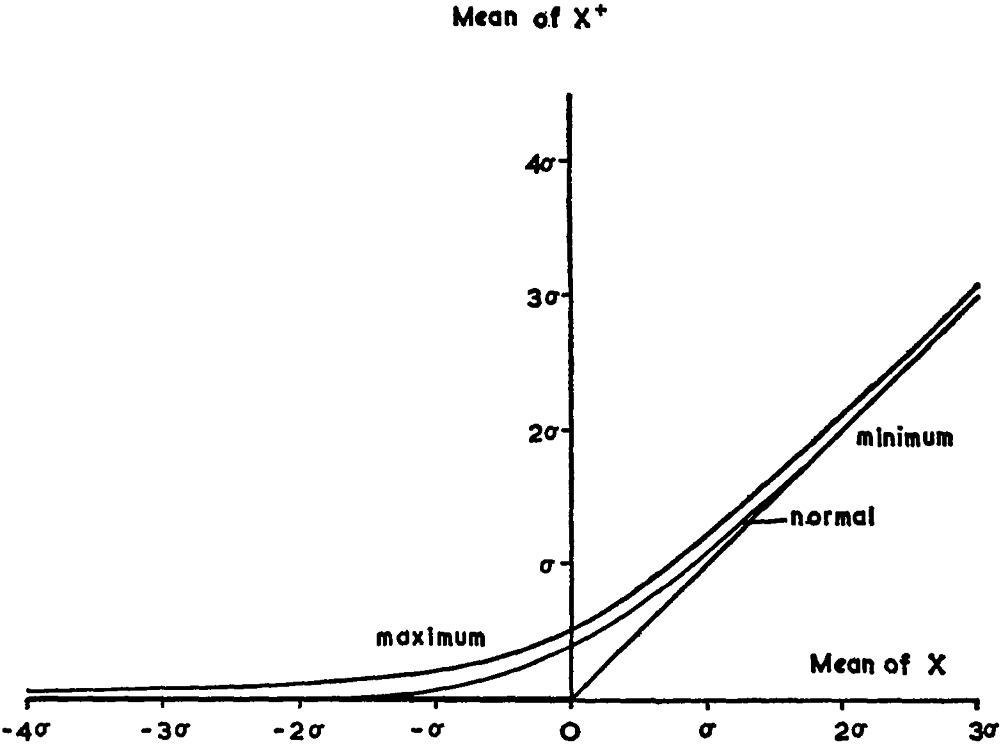

FIG. 4.—Results of the Linear Programming Calculations. (The maximum and minimum values for the mean of X^+ are plotted against the mean of X. Also given is the curve when X is normal)

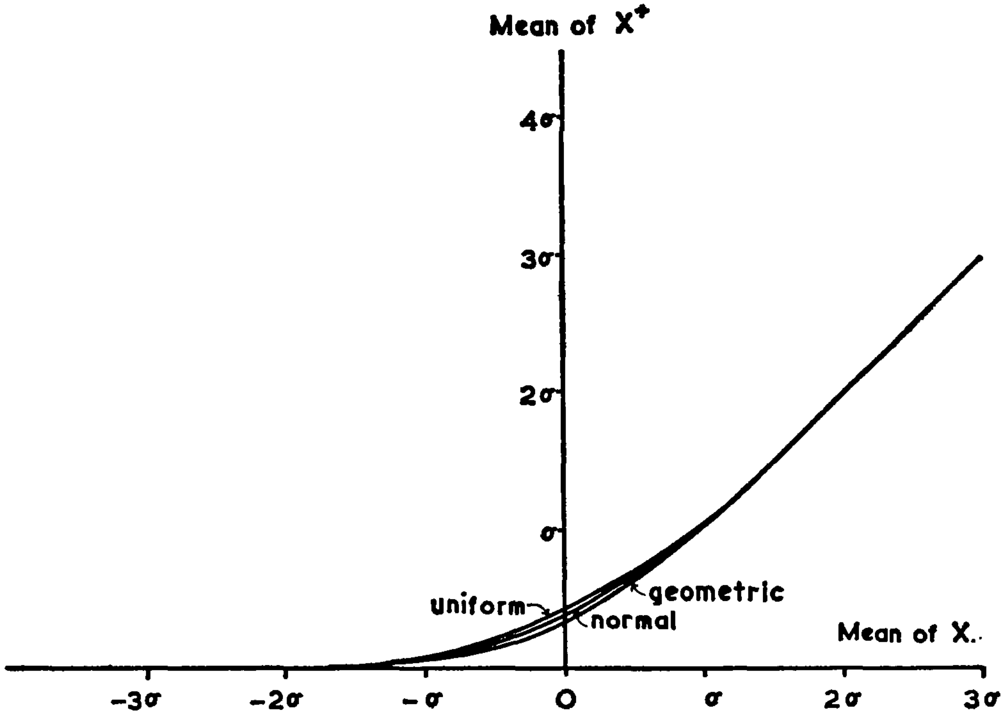

FIG. 5.—The Mean of X^+ as a Function of the Mean of X, when X has the Uniform Distribution, the Normal Distribution, and the Geometric Distribution

Let us look at the expectation of $X+$ and investigate the limits between which it must lie for all possible distributions of X. In a discrete formulation such as this, the problem reduces to a linear programming problem: we wish to maximize and minimize

$$\sum_{j=0}^{\infty} jp_j,$$

subject to

$$\sum_{j=-\infty}^{\infty} p_j = 1,$$

$$\sum_{j=-\infty}^{\infty} jp_j = \mu_1,$$

$$\sum_{j=-\infty}^{\infty} j^2 p_j = \mu_1^2 + \sigma_1^2,$$

and

$$p_j \geq 0, \qquad (\text{all } j).$$

A suitable program was written for TITAN, and with $\sigma_1 = 50$, this linear programming problem was solved for various values of μ_1. The results of this investigation are presented graphically in Figure 4. The results when X is normal are also given in the diagram.

The rather unusual distributions (with only three non-zero p_j) which give rise to the maxima and minima were available from the computer output. These unusual distributions (especially near $\mu_1 = 0$) suggest that the bounds given in Figure 4 are wider than necessary.

The variance may be examined in a similar manner, but it leads to a nonlinear programming problem. No calculations were performed, firstly because of the greater amount of computer time required, and secondly because this method would give wide bounds like those obtained in the expectation calculations.

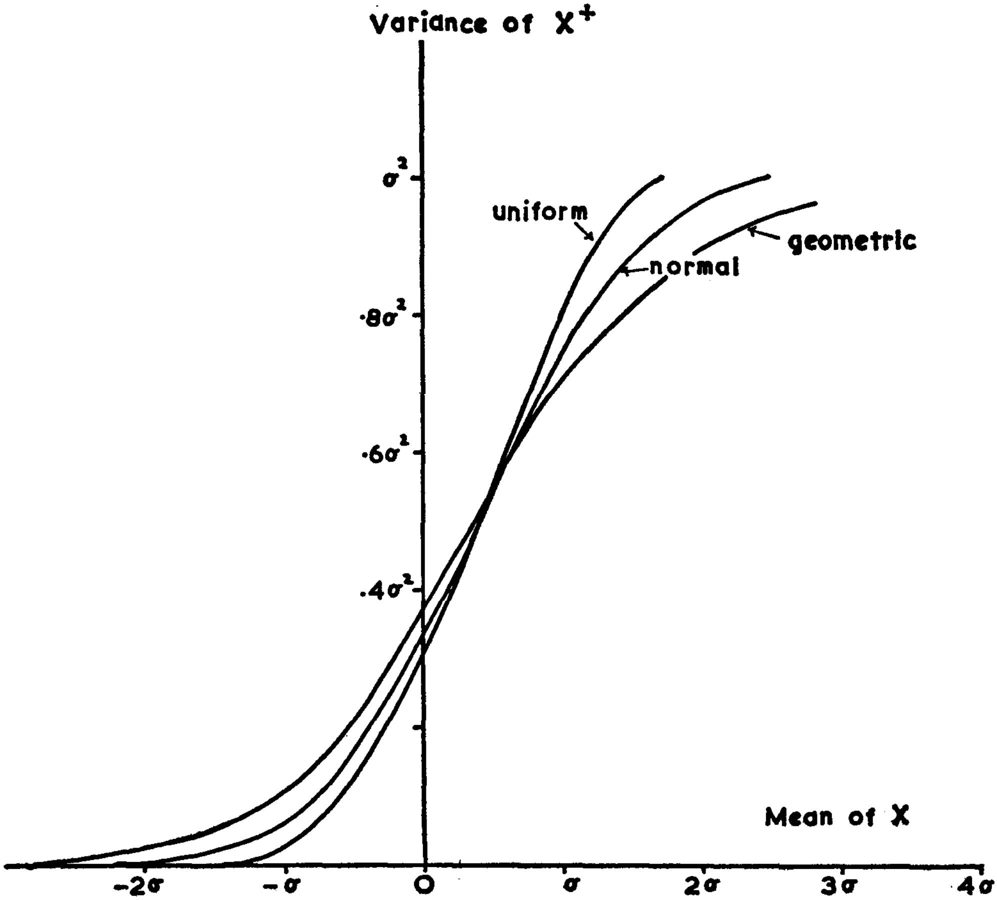

FIG. 6.—The Variance of X^+ as a Function of the Mean of X When X has the Uniform, Normal and Geometric Distributions

It is of interest at this stage to examine the expectation and variance of X^+ when X has a certain known distribution. Two cases were therefore examined: (i) X having the discrete uniform distribution

$$p_j = \frac{1}{n - m + 1}, \quad j = m, m + 1, \cdots, n; \text{ and}$$

(ii) X having the discrete double geometric distribution

$$p_j = K(\lambda) \exp\{-|j - \mu|/\lambda\}, \quad -\infty < j < \infty.$$

Both λ and $(n - m)$ were large. The results obtained are given in Figures 5 and 6, together with the appropriate normal curves.

Neither of these two distributions resembles the normal distribution, and yet the curves obtained in both cases lie close to the curves for the normal case. These calculations support the remark number (1) of §6.1; it is to be expected that the expectation and variance of X^+ vary very little for a wide range of possible distributions of X all having the same first two moments.

6.3 The Approximate Computation Procedure

In §6.2, two discrete random variables

X and Y were defined with expectations μ_1 and μ_2 respectively, variances σ_1^2 and σ_2^2 respectively and covariance $\rho\sigma_1\sigma_2$. If X is defined over all integers and Y over all non-negative integers, and we wish to take the positive part $X+$ of X, approximations to the first and second order moments of $X+$ may be obtained as follows:

$$\mathcal{E}(X^+) = \sum_{i=0}^{\infty} ip_i$$
$$\doteqdot \frac{1}{\sqrt{2\pi}\,\sigma_1} \cdot \int_0^{\infty} x \exp\left\{-\frac{1}{2}\left[\frac{x-\mu_1}{\sigma_1}\right]^2\right\} dx$$
$$= \mu_1 F\left(\frac{\mu_1}{\sigma_1}\right) + \sigma_1 f\left(\frac{\mu_1}{\sigma_1}\right) \qquad (19)$$

where

$$f(u) = \frac{1}{\sqrt{2\pi}} \exp\{-\tfrac{1}{2}u^2\},$$

and

$$F(u) = \int_{-\infty}^{u} f(x)\,dx.$$

An approximation to $\mathcal{E}(X^+)^2$ is found in a similar manner:

$$\mathcal{E}(X^+)^2 \doteqdot (\sigma_1^2 + \mu_1^2)F\left(\frac{\mu_1}{\sigma_1}\right) + \sigma_1\mu_1 f\left(\frac{\mu_1}{\sigma_1}\right). \qquad (20)$$

An approximation to the produce moment of $X+$ and Y is obtained by considering a bivariate normal integral; the covariance of $X+$ and Y then has a very simple form:

$$\text{Cov}\,(X^+, Y) \doteqdot \rho\sigma_1\sigma_2 F\left(\frac{\mu_1}{\sigma_1}\right). \qquad (21)$$

In Figure 3, there is only one random variable which must have its positive part taken. However, for a population with an age structure, there are many such variables. It is therefore necessary to consider the case in which both X and Y are distributed over all the integers, and both X and Y have their positive parts taken ($X+$ and $Y+$ respectively). Once again, an approximation to the product moment is obtained using the bivariate normal integral, but this integral is troublesome to evaluate. A computer can readily perform the calculation, but a large number of such integrals are required for a reasonably realistic population model, and the time required would be prohibitive.

A simple method is available, however, and it makes use of the Mehler expansion of a bivariate normal density (M. G. Kendall, 1948, 355–356; H. O. Lancaster, 1958):

$$\frac{1}{2\pi\sqrt{1-\rho^2}} \cdot \exp\left\{\frac{-1}{2(1-\rho^2)}[x^2 - 2\rho xy + y^2]\right\}$$
$$= \frac{1}{\sqrt{2\pi}} \exp\{-\tfrac{1}{2}x^2\} \cdot \frac{1}{\sqrt{2\pi}} \exp\{-\tfrac{1}{2}y^2\} Q(x, y),$$

where $Q(x, y) = 1 + \rho xy + 1/2!\,\rho^2 (1 - x^2)(1 - y^2) + 1/3!\,\rho^3 (x^3 - 3x)(y^3 - 3y) + \ldots$. After expanding the bivariate density in this manner, integrating, and subtracting the product of the expectations, we obtain:

$$\text{Cov}\,(X^+, Y^+) \doteqdot \rho\sigma_1\sigma_2 F(\mu_1/\sigma_1)F(\mu_2/\sigma_2) + [\tfrac{1}{2}\rho^2\sigma_1\sigma_2 + \tfrac{1}{6}\rho^3\mu_1\mu_2] f(\mu_1/\sigma_1) f(\mu_2/\sigma_2) + \cdots. \qquad (22)$$

The correlation coefficient ρ will always be strictly less than one for our problems (and usually much less!) so we have no convergence problems. Note that it is only necessary to determine $F(\mu/\sigma)$ and $f(\mu/\sigma)$ once for each random variable whose positive part is required. Rather than compute these two functions, values can be obtained more quickly from tables of the normal ordinate and integral stored in the computer. The moments of the positive parts are then readily evaluated.

This approximate procedure depends heavily on two assumptions (discussed in some detail in §6.4, §6.5 and §6.6): (i) the random variables whose positive parts are required have distributions in pairs close to bivariate normal; and (ii) the moments of the positive parts, being averages, do not depend too heavily on the actual distributions.

The moments obtained using these methods are of course approximate, and the question arises: how good is the approximation? We shall show that the approximation is *extremely* good, and that the errors involved are negligible.

6.4 Some Monte Carlo Experiments

One possible method of examining the accuracy of the suggested recurrence method is to compare results using it with the results of Monte Carlo experiments. We describe here four such experiments.

Experiment 1. Consider a population consisting of three types of entity: men, women and couples. During a unit time interval, there are three possible outcomes for men:

(1) man has desire to marry with probability .3;
(2) man merely survives with probability .6; and
(3) man dies with probability .1.

There are also three possibilities for a woman:

(1) woman has desire to marry with probability .3;
(2) woman merely survives with probability .65; and
(3) woman dies with probability .05.

For couples, four outcomes are possible:

(1) couple survives and has one son with probability .105;
(2) couple survives and has one daughter with probability .1;
(3) couple merely survives with probability .6; and
(4) couple ceases to exist with probability .195.

A Monte Carlo experiment was performed with this type of population. At time $t = 0$, there were 1,000 men, 1,000 women and 1,000 couples, and the experiment was performed with 40 observations on the first 100 time units. Over that long time period, the Monte Carlo means did not differ significantly from the (approximate) theoretical means. Furthermore, the variances were not sig-

TABLE 1.—Results from the First Monte Carlo Experiment

THEORETICAL			MONTE CARLO		
expectations at time t = 10			means at time t = 10		
192.188	321.523	560.473	193.675	321.300	557.575
covariance matrix, t = 10			observed covariance matrix, t = 10		
226.534	- 37.366	167.573	233.919	140.948	258.737
- 37.366	1064.137	-155.012	140.948	783.260	107.428
167.573	- 155.012	725.981	258.737	107.428	804.544
expectations at time t = 50			means at time t = 50		
3.97966	37.8253	11.4906	4.02500	34.32500	11.8250
covariance matrix, t = 50			observed covariance matrix, t = 50		
5.90590	- 3.36712	6.67877	3.97438	- 4.83313	6.95438
-3.36712	127.79124	- 7.47330	-4.83313	134.66937	-13.11813
6.67877	- 7.47330	24.34680	6.95438	- 13.11813	23.39438

nificantly large or small. Some of the results output are given in Table 1. A comparison of the covariances in Table 1 may be puzzling. The sample covariances have large sampling variances. The covariances were not themselves tested directly. However, the theoretical covariances at time t are used to compute the theoretical variances at later points of time. The fact that these variances are compatible with the Monte Carlo results is an indirect test of the covariances.

The population under consideration is rapidly approaching extinction.

Experiment 2. In the above experiment, the differences between the number of men desiring marriage and the number of women desiring marriage became large and negative as t increased. Consequently, we should expect comment (3) of §6.1 to apply, and the approximate method of computation to give good results. It therefore seems desirable to examine a case in which the population size remains more or less constant, and in which the difference between the numbers of each sex desiring marriage is always close to zero. Another experiment was therefore performed using the same model as Experiment 1. Initially there were 500 men, 500 women and 1,000 couples. The ten probabilities were: .18, .79, .03; .18, .80, .02; .105, .1, .705, .09; enumerated in the same order as in the first experiment.

The same theoretical calculations were made, and a Monte Carlo experiment with 31 observations for $t = 0$ to 100 performed. Once again, the Monte Carlo results did not differ significantly from the (approximate) theoretical calculations. Table 2 contains some of the results output.

In this second experiment, the deterministic means remain at 500 for men, 500 for women and 1,000 for couples. The theoretical stochastic means, however, differ from these, and the Monte Carlo results seem to bear this out.

Experiment 3. Experiments 1 and 2 each contained one random variable whose positive part was taken. An experiment was therefore performed with two types of men, two types of women and one type of couple, and in this experiment two random variables had their positive parts taken.

The Monte Carlo experiment was performed with 46 observations on the first 50 time units, and once again, these re-

TABLE 2.—Results from the Second Monte Carlo Experiment

THEORETICAL			MONTE CARLO		
expectations at time t = 10			means at time t = 10		
516.020	517.670	978.700	517.484	521.290	980.613
covariance matrix, t = 10			observed covariance matrix, t = 10		
780.652	− 131.936	176.979	951.411	− 287.624	253.736
− 131.936	783.376	152.868	− 287.624	620.142	− 97.081
176.979	152.868	796.206	253.736	− 97.081	896.495
expectations at time t = 99			means at time t = 99		
491.230	503.744	924.991	482.452	509.419	929.161
covariance matrix, t = 99			observed covariance matrix, t = 99		
2671.237	6206.332	2986.126	2356.635	1189.101	2535.185
6206.332	3003.016	2860.655	1189.101	3276.760	2989.029
2986.126	2860.655	6363.501	2535.185	2989.029	5222.200

TABLE 3.—Results from the Third Monte Carlo Experiment

		THEORETICAL		
		expectations at time t = 50		
15.081	29.842	15.078	29.837	49.899
		covariance matrix at time t = 50		
35.580	5.273	16.984	17.125	65.883
5.273	112.577	17.057	15.849	62.143
16.984	17.057	35.571	5.299	66.132
17.125	15.849	5.299	112.135	61.975
65.883	62.143	66.132	61.975	259.851
		MONTE CARLO		
		means at time t = 50		
14.261	29.652	14.957	28.717	48.652
		observed covariance matrix, t = 50		
34.454	5.417	15.337	4.748	49.982
5.417	166.923	9.637	9.597	40.009
15.337	9.637	29.389	9.509	43.159
4.748	9.597	9.509	98.724	47.315
49.982	40.009	43.159	47.315	177.749

sults did not differ significantly from the (approximate) theoretical results. In Table 3, the theoretical expectations and covariance matrix for $t = 50$ are given, together with the observed means and observed covariance matrix for $t = 50$.

Experiment 4. The above three Monte Carlo experiments suggest that the approximate recurrence method is extremely good. However, as a further test, one other Monte Carlo experiment was performed. This experiment used the same model and data as Experiment 2, and produced frequency polygons for the numbers of men, women and couples at time $t = 10$. The polygons, based on a sample size of 299, are reproduced in Figure 7. The normal density curves included in Figure 7 have parameters obtained from the left hand side of Table 2. The χ_{20}^2 values of goodness-of-fit are 10.464 for males, 16.633 for females, and 15.515 for couples (They are not independent of course). Each of these values is much less than the expected value of χ_{20}^2 (and almost significantly small!). The fit is apparently very good.

It should be noted that the deterministic means are: males—500; females—500; and couples—1,000. That is, the stochastic means are considerably different from the deterministic means.

6.5 Some Numerical Calculations

From the observations made in §6.1, and also from Figures 4, 5 and 6, it appears that the largest errors made in calculating the first two moments of the positive part of a random variable occur when the expected value of the random variable lies close to zero. Furthermore, because we are interested in large populations, many of the conditional binomial probabilities may be represented accurately by probabilities of the form:

$$p_j = K \exp\{-(j - np)^2/(2npq)\} \qquad (23)$$

It is of interest to consider the discrete trivariate distribution

$$P(X = i,\ Y = j,\ Z = k) = C \exp\{-(\mathbf{x}'\mathbf{V}^{-1}\mathbf{x})/2\} \qquad (24)$$

where

$$\mathbf{x} = \begin{bmatrix} i - \mu_1 \\ j - \mu_2 \\ k - \mu_3 \end{bmatrix},$$

and μ_1, μ_2 and μ_3 are suitable means, and $\mathbf{V}$ is a suitable covariance matrix of full rank.

Numerically, it is possible to obtain the trivariate distribution of X, Y and Z^+, where $Z^+ = \max(Z, 0)$, and it is then easy to compute the joint distribution of

$$U = X$$

and

$$W = Y + Z^+.$$

If one considers the associated (continuous) trivariate normal distribution, it is soon apparent that U and W have a reasonably well-behaved bivariate distribution when the matrix $\mathbf{V}$ is of full rank. Let us assume that the means μ_1 and μ_2 are large and positive. Let us further assume that these two random variables must be non-negative. It is then possible to define two random variables U^* and W^*, conditional on U and W, as follows:

$$P(U^* = j \mid U = n) = \binom{n}{j} p_1{}^j (1 - p_1)^{n-j}, \quad (25)$$

$$P(W^* = j \mid W = n) = \binom{n}{j} p_2{}^j (1 - p_2)^{n-j}. \quad (26)$$

Using equation (23), it is easy to obtain an accurate approximation to the joint distribution of U^* and W^*. We shall be interested in the form of this joint distribution.

A glance at Figure 3 shows that we are in effect examining part of the process from stage 2 in time interval $(t, t + 1)$ until stage 2 in the time interval $(t + 1, t + 2)$. These numerical computations were carried out with several different parameters, and the joint distribution of U^* and W^* examined. U^* and W^* were virtually indistinguishable from bivariate normal variables, and when, for example, the conditional distribution of $W^*|U^*$ was plotted on normal probability paper, a straightedge was necessary to distinguish the graph from a straight line. The random variables U and W, on the other hand, had a bivariate distribution which would resemble a bivariate normal density, but for a moderately pronounced skewness.

These calculations suggest that if the random variables at stage 2 in Figure 3 have distributions which pairwise resemble bivariate normal distributions, the random variables at stage 4 have distributions which pairwise resemble skewed bivariate normal densities. The conditional multinomial processes at stage 1 in the following time interval then have the effect of rectifying the skewness present, and the random variables at stage 2 again have bivariate distributions similar to bivariate normal densities.

6.6 Some Analytical Results

It was observed in §6.5 that conditional multinomial processes seem to rectify skewness in a bivariate distribution which otherwise resembles a bivariate normal density. In this section, therefore, analytical results associated with conditional multinomial processes are discussed. The following elementary theorems should first be noted:

Theorem 1. Let U be a random variable having the Binomial distribution $B(n, p)$. Let U^* be a random variable conditional on U and having the conditional binomial distribution $B(U, p_1)$. Then U^* has the binomial distribution $B(n, pp_1)$, If n is not too small, a normal approximation is accurate.

Theorem 2. Let U be a random variable having the Poisson distribution with mean λ. $\{W_j\}$ $(j = 1, 2, \ldots, k)$ are conditional multinomial random var-

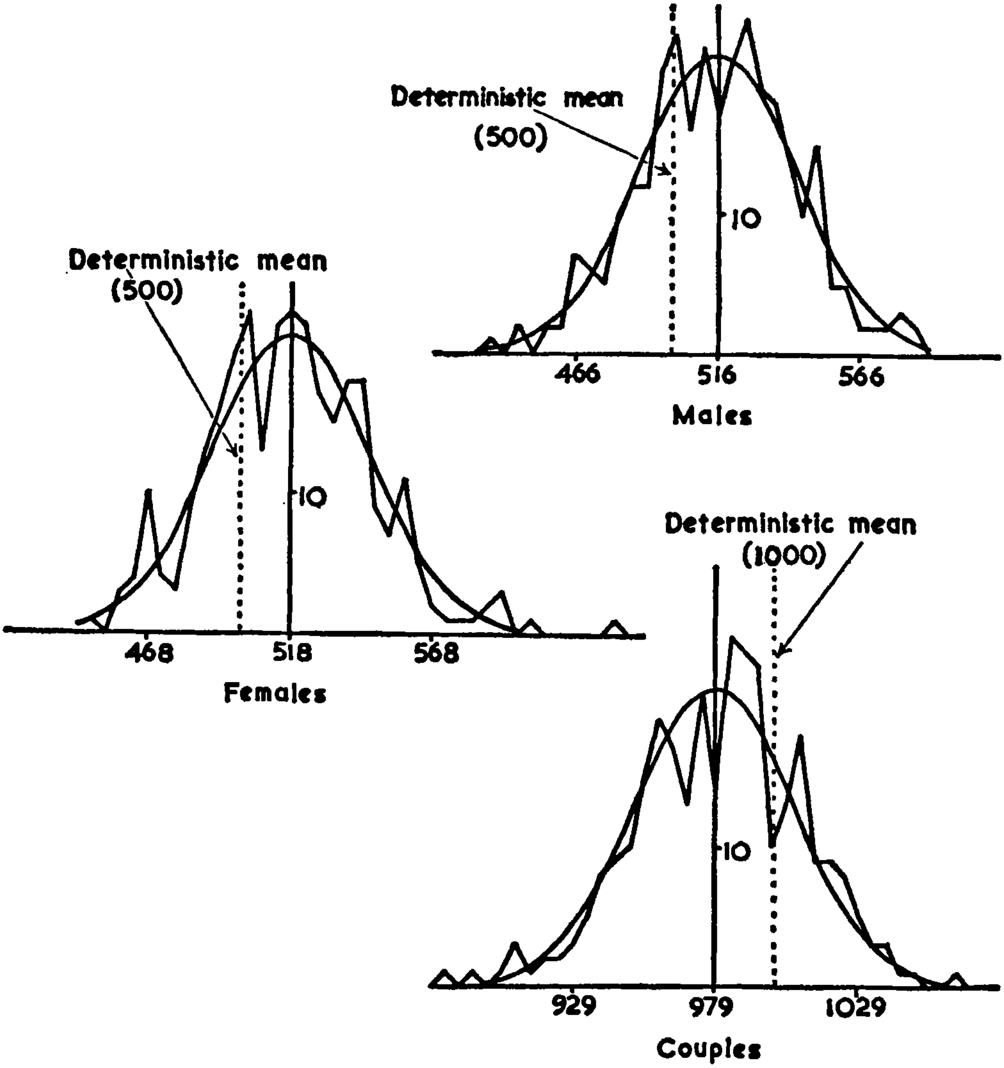

FIG. 7.—Results of the Fourth Monte Carlo Experiment

iables conditional on U, and having the conditional distribution Mult. $(U; p_1, p_2, \ldots, p_k)$. Then the $\{W_j\}$ are mutually independent Poisson variates with means $\{\lambda p_j\}$.

Both these results assume that U has a known well-behaved distribution. The following does not, and is closer to the situation we need to investigate.

Theorem 3. Let U be a random variable taking positive integral values. Let it have fixed finite variance σ^2 and a mean μ. $U^*|U$ is a conditional binomial random variable $B(U, p)$. Then if $\mu \to \infty$ and $p \to 0$ such that $\mu p \to \lambda$, the limiting distribution of U^* is Poisson with mean λ.

If λ is not too small, a normal approximation is accurate. Theorem 3 may be generalized for two dimensions:

Theorem 4. Let U and W be two correlated random variables taking positive integral values and having fixed finite variances σ_1^2 and σ_2^2 respectively. Let their respective means be μ_1 and μ_2. $U^*|U$ is a conditional binomial random variable $B(U, p_1)$ and $W^*|W$ is a conditional binomial random variable $B(W, p_2)$. The two conditional distributions are independent. Then if $\mu_1 \to \infty$ and $p_1 \to 0$ such that $\mu_1 p_1 \to \lambda_1$, and $\mu_2 \to \infty$ and $p_2 \to 0$ such that $\mu_2 p_2 \to \lambda_2$, U^* and W^* have in the limit independent Poisson distributions with parameters λ_1 and λ_2 respectively. In this case, a normal approximation will be accurate, provided λ_1 and λ_2 are not too small.

The conditions for the above results are very similar to the conditions encountered with the two-sex model. However, none of them is completely appropriate to the two-sex situation. Consider a random variable U^* conditional on U, and having the conditional binomial distribution $B(U, p)$. U has the distribution $\{p_j\}$ $(j = 0, 1, 2, \ldots)$ with mean μ and variance σ^2. Let us examine the case in which p is small (less than .1, say), μ is large (greater than 1,000, say) and σ^2 is smaller than μ.

The probability that U^* is equal to j (P_j, say) is given by

$$P(U^* = j) = P_j = \sum_n \binom{n}{j} p^j q^{n-j} p_n. \quad (27)$$

Let us now assume μ to be an integer; this assumption simplifies the algebra, but does not invalidate the final result. The right-hand side of equation (27) may be expanded in the form:

$$\binom{\mu}{j} p^j q^{\mu-j} \Bigg\{ \Bigg[p_\mu + \frac{(\mu+1)q}{(\mu+1-j)} p_{\mu+1} + \frac{(\mu+1)(\mu+2)q^2}{(\mu+1-j)(\mu+2-j)} p_{\mu+2} + \cdots \Bigg] + \Bigg[\frac{(\mu-j)}{\mu q} p_{\mu-1} + \frac{(\mu-j)(\mu-1-j)}{\mu(\mu-1)q^2} p_{\mu-2} + \cdots \Bigg] \Bigg\}. \quad (28)$$

Writing $(\mu p + d)$ for j, we have:

$$\frac{(\mu+r)q}{\mu+r-j} = 1 + \frac{(d-rp)}{q\mu} + \frac{(d-rp)(d-r)}{q^2\mu^2} + 0\left\{\frac{(d-rp)(d-r)^2}{q^3\mu^3}\right\}, \quad (29)$$

where d is not too large and $|r| < 3\sigma$. Taking logarithms, we have:

$$\log\left\{\frac{(\mu+r)q}{\mu+r-j}\right\} = \frac{(d-rp)}{q\mu} + \frac{(d-rp)(d-r)}{q^2\mu^2} - \frac{(d-rp)^2}{2q^2\mu^2} + 0\left\{\frac{(d-rp)(d-r)^2}{q^3\mu^3} - \frac{(d-rp)^2(d-r)}{q^3\mu^3}\right\}. \quad (30)$$

Summing for $r = 1$ to k, and neglecting terms in the sum which are very small, we obtain:

$$\sum_{r=1}^{k} \log\left\{\frac{(\mu+r)q}{\mu+r-j}\right\} = \left\{\frac{(2d-p)}{2q\mu}\right\}k - \left\{\frac{p}{2q\mu}\right\}k^2. \quad (31)$$

The same relation is true for the left-hand tail of the distribution. Hence an approximation for P_j is given by

$$P_j = \binom{\mu}{j} p^j q^{\mu-j} \left\{ \sum_k p_k \cdot \exp\left[\left\{\frac{2d-p}{2q\mu}\right\}k - \left\{\frac{p}{2q\mu}\right\}k^2\right]\right\}, \quad (32)$$

where the summation is from $k = -$ integer part (3σ) to $k =$ integer part of (3σ), assuming the $\{p_j\}$ distribution to be reasonably well-behaved.

If the $\{p_j\}$ distribution is well-behaved and has a distribution not unlike the shape of the normal density curve, we may consider a continuous density curve approximately the $\{p_j\}$ distribution, and expand this continuous density curve in a Gram-Charlier series (Cramér, 1961):

$$f(x) = \frac{1}{\sigma}\left\{\phi\left(\frac{x}{\sigma}\right) + \frac{C_3}{3!}\phi^{(3)}\left(\frac{x}{\sigma}\right) + \frac{C_4}{4!}\phi^{(4)}\left(\frac{x}{\sigma}\right) + \cdots\right\} \quad (33)$$

where $C_3 = -\mu_3/\sigma^3$ and $C_4 = \mu_4/\sigma^4 - 3$.

$$\phi^{(n)}(x) = \frac{d^n}{dx^n}\phi(x) = \frac{d^n}{dx^n}\left\{\frac{1}{\sqrt{2\pi}}\exp(-x^2/2)\right\}. \quad (34)$$

Then the sum in equation (32) may be approximated by

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left[\left\{\frac{(2d-p)}{2q\mu}\right\}k - \left\{\frac{p}{2q\mu}\right\}k^2\right]f(x)\,dx. \quad (35)$$

Furthermore,

$$\binom{\mu}{j}p^j q^{\mu-j} \doteqdot \frac{1}{\sqrt{2\pi}\sqrt{\mu pq}}\exp(-d^2/(2\mu pq)). \quad (36)$$

Evaluating the integral (35), and combining the result with (36), we obtain

$$P_j = \frac{1}{\sqrt{2\pi}\sqrt{pq\mu + p^2\sigma^2}} \cdot \exp\left\{-\frac{d^2}{2(\mu pq + p^2\sigma^2)}\right\}g(d), \quad (37)$$

for moderate d, where $g(d)$ is given by

$$g(d) = \left[1 + \frac{C_4}{8}\left\{1 - \frac{2pq\mu}{pq\mu + p^2\sigma^2} + \frac{(pq\mu)^2}{(pq\mu + p^2\sigma^2)^2}\right\}\right] + \frac{d}{q\mu}\left[\frac{C_3}{2}\left\{\frac{pq\mu\sigma}{pq\mu + p^2\sigma^2}\right\} \cdot \left\{1 - \frac{(pq\mu)^2}{(pq\mu + p^2\sigma^2)^2}\right\}\right]. \quad (38)$$

Thus $g(d)$ is a constant with an error term of $0(C_3\sigma d/\mu)$. If the coefficients of skewness and excess of the $\{p_j\}$ distribution are small, $g(d)$ is close to unity. Thus, in this very special case, we have shown that the distribution of U^* is close to a normal density curve, indeed the normal curve with mean μp and variance $(\mu pq + p^2\sigma^2)$. It is simple to prove that μp is the exact mean of U^* and that $(\mu pq + p^2\sigma^2)$ is the exact variance of U^*.

The conditions under which formula (37) is true should be emphasized:

(1) μ is large ($>$ 1,000, say);
(2) $\sigma^2 < \mu$;
(3) the $\{p_j\}$ distribution may be accurately approximated by a Gram–Charlier series;
(4) p is small ($<$ .1, say); and
(5) $|d|$ is moderate in size ($< 3\sqrt{\mu pq + p^2\sigma^2}$, say).

It is clear that bivariate formulae exist corresponding to equations (28), (31) and (32). However, the simplification of these formulae is considerably more difficult than the simplification of the formula for P_j.

Much work remains to be done for the analysis of the general situation when σ^2 may be much larger than μ. The algebra involved in the preliminary analysis of the above very special case is very tedious, and it seems likely that the analysis of the more general situation will be even more tiring.

One further comment should be made concerning the distributions of the random variables in this bisexual model: the linear transformations at stages 2 and 4 in Figure 3 should, due to the Central Limit Theorem, encourage normality; the random variables concerned are not independent, but many of them are only slightly correlated.

6.7 Some Generalizations of the Model

It was mentioned in §3 that mortality probabilities, fertility probabilities, etc.

may themselves be considered as random variables. For the multi-type Galton-Watson process, the basic moment recurrence relation is altered only slightly in this situation. It is soon apparent that the two-sex computation procedure requires a similar minor modification to allow for this extra complication. We show later in §6.9 that the probabilities must be considered as random variables in any realistic population model.

Immigration is mentioned in §5, and it is discussed in greater detail elsewhere (J. H. Pollard, 1966, 1967). The methods outlined are easily incorporated in the analysis of the two-sex model.

Time trends in the probabilities (or distributions of the probabilities) may be readily incorporated in the model. The adjustment necessary for the computation procedure is straightforward.

6.8 The Basis of the General Computer Program

The computer program discussed in this section does *not* include allowances for

(i) probabilities which are themselves random variables;
(ii) immigration; and
(iii) time trends for probabilities (or for distributions of probabilities).

However the basic computer program can be modified to take all these factors into account, since the theoretical alterations and programming alterations are fairly trivial. There are a few practical difficulties, however (mainly associated with data), but we shall ignore them for the moment.

The difficulties encountered with the basic program are not caused by theoretical complications, but rather by the storage limitations of even moderately large computers. The theoretical calculations are straightforward. Five magnetic tape decks are required, and it is in organizing the data within the machine that skillful programming is required. The program (used in the numerical example of §6.9) was written in FORTRAN II for the IBM 7094 computer system at the University of Chicago Computation Center.

Data are input to the computer by punched card, and the cards are accepted in the following order:

(*1*) *Structural Constants Card.* This card gives five integer numbers to the machine:

(i) the number of types of entity, T $(= M + F + C)$;
(ii) the size of the time step and age step;
(iii) the number of male age groups, M;
(iv) the number of female age groups, F; and
(v) the number of groups of couples, C.

(*2*) *Male Probabilities Cards.* For each age group there may be one, two or three cards; the first number on each card is an integer giving the youngest age of the age group, and the last number is either 1, 2 or 3. The male probabilities cards are accepted by the computer in any order whatsoever. For each card, a fractional number is read after the age group integer, and if the last number on the card is 1, this fraction is the probability of the single male merely surviving. If the last card number is 2 or 3, the fraction is ignored. Five number pairs lie between the fraction and the final number on the card; each pair consists of an integer (female age group) and a fraction (probability), and the fifteen possible pairs describe the age preferences for brides of single men in that age group; the pairs may be in any order on the cards. The data in Appendix Table 1 have this format.

(*3*) *Female Probabilities Cards.* These have the same format and obey the same rules as the Male Probabilities Cards, except that the last number on each card must be either -1, -2, or -3. For females, the first fraction on the card is *not* ignored when the final integer is -2 or -3; the fractions here represent the probabilities of an illegitimate son or daughter respectively. The data in Appendix Table 2 have this format.

(*4*) *Couple Probabilities Cards.* There is one card for each couple group, and each card contains eight numbers:

(i) the age group of the husband (an integer);
(ii) the age group of the wife(an integer);
(iii) the probability that the couple merely survives;
(iv) the probability that the husband dies, the wife survives and no child is born;
(v) the probability that the wife dies, the husband survives and no child is born;
(vi) the probability that the couple survives and a son is born;
(vii) the probability that the couple survives and a daughter is born; and
(viii) the probability of divorce.

The data in Appendix Table 3 have this format.

(*5*) *Initial Single Male Population Cards.* There is one card for each male age group, and each card contains three numbers:

(i) the age group involved (an integer);
(ii) the initial number in the population of that age group (an integer); and
(iii) the integer "1" to indicate "male".

(*6*) *Initial Single Female Population Cards.* The same format and rules apply as for single males. An integer "-1" indicates "female."

(*7*) *Initial Married Population Cards.* There is one card for each couple group, and each card contains four numbers:

(i) the age group of the husband;
(ii) the age group of the wife;
(iii) the initial number of couples in that category; and
(iv) the integer "0" to indicate "couple."

(*8*) *Projection Output Cards.* Each card contains one integer, and the integers must form a strictly monotonic increasing sequence. As soon as this rule is violated, the program is terminated, and this is the method for stopping. The integers indicate the points of time at which the projected population is to be output by printer.

The random variables representing the numbers of the various entities are ordered as a vector in the machine as follows: single males (in ascending age groups), then single females (in ascending age groups) and finally the couples (ordered according to the order of input of the couple probabilities cards). All the probabilities are listed in one enormous vector, and another list of numbers indicates how many probabilities are to be associated with each type of entity. Marriage-desire probabilities of males aged x for females aged y and of females aged y for males aged x must be paired off. Then the linear transformation constants for stage 4 must be determined; the matrix involved is enormous, but as most of the elements are zero, and the others are either plus or minus one, the information required may be stored in a very compact form.

All these preliminaries take up half the written program. The recurrence procedure loop then follows.

6.9 A Numerical Example

The population projection program of §6.8 was used to project an hypothetical human population using a time unit of two years. The single male population was divided into thirty age groups 0-, 2-, 4-, . . . , 58-, and the single female population into twenty-five age groups 0-, 2-, 4-, . . . , 48-; 160 types of couple were considered.

The data for the calculations were based on the Australian population in 1960. It was the original intention of the author to project the Australian population from 1960, but for two reasons, this goal was abandoned:

(i) certain important data were not readily available to the author; and
(ii) the preparation of the data involved a

TABLE 4.—Single Male Population

Age	Initial single male population	Single male population at time t = 1 unit: Expected	Variance	Single male population at time t = 2 units: Expected	Variance	Single male population at time t = 3 units: Expected	Variance
0 . .	228998	216300	168347	201258	164491	194244	163174
2 . .	227952	228309	687	215649	167985	200652	164106
4 . .	220945	227596	355	227953	1041	215313	167797
6 . .	213380	220698	247	227341	609	227697	1293
8 . .	212513	213158	222	220468	476	227105	844
10 . .	208887	212332	180	212977	402	220281	662
12 . .	203486	208701	186	212143	369	212787	591
14 . .	199595	203252	234	208461	425	211899	612
16 . .	171708	199256	339	202906	578	208107	777
18 . .	149634	171217	490	198686	905	202326	1153
20 . .	130472	145297	4293	165804	5846	192374	7137
22 . .	96624	116304	13621	127021	21393	143426	26872
24 . .	66991	74490	20224	88073	37639	91286	51274
26 . .	46999	53055	13583	57189	32908	67023	50628
28 . .	37802	35076	12234	38779	26270	39582	41561
30 . .	34558	29949	9043	26561	20931	27966	31287
32 . .	29905	29543	6677	25210	13326	22330	22915
34 . .	27232	27112	5069	26905	10487	22813	16287
36 . .	23872	24993	4713	24973	9243	24696	13835
38 . .	22202	22630	3818	23797	8105	23847	12347
40 . .	20203	21479	3492	22001	6774	23141	10745
42 . .	17247	19876	3147	21199	6280	21777	9170
44 . .	17112	17155	2794	19833	5752	21184	8612
46 . .	17842	17102	2502	17264	5043	19983	7927
48 . .	17537	17925	2371	17222	4563	17488	6915
50 . .	17711	17606	2073	18012	4278	17338	6208
52 . .	16236	17650	1726	17560	3621	17970	5690
54 . .	15621	16013	1292	17451	2990	17362	4741
56 . .	15443	15270	1011	15642	2207	17104	3903
58 . .	14887	14975	866	14789	1767	15133	2867

considerable amount of clerical work, and the author did not have any computing assistance.

Divorce and ex-nuptial births were included in the calculations, and the 1393 probabilities for the population are listed in Appendix Tables 1, 2 and 3. The initial population structure, and the projected populations for $t = 1$, 2 and 3 units (i.e., 2, 4 and 6 years) are given in Tables 4, 5 and 6. For each time unit, the projection calculations took 62 minutes; this is quite a short time when it is realized that there are $(1393)^2 = 1.94$ million covariances to be calculated, output to magnetic tape, and later read from magnetic tape several times. Much of the computer time was taken up by magnetic tape operations; with a time-sharing machine, the computing time required should be much less.

The numerical example illustrates the power of this projection technique. In practice, a time unit of one year is recommended, and the size of the problem then increases by a factor of almost 16. With one exception, the probability of two or more vital events in one year may be safely neglected; the exception is the probability of marriage and a birth in the one year, and *the computer program should be modified to deal with this situation.* A time unit of two years was used for the numerical example for reason (ii) above, and also to save computer time at the research stage.

Although expectations and variances only are given in Tables 4, 5 and 6, the

TABLE 5.—Single Female Population

Age	Initial single female population	Single female population at time t = 1 unit: Expected	Variance	Single female population at time t = 2 units: Expected	Variance	Single female population at time t = 3 units: Expected	Variance
0 . .	218002	207828	163559	193374	159431	186643	157955
2 . .	218495	217405	596	207259	163232	192845	159087
4 . .	209040	218228	266	217139	859	207006	163086
6 . .	204602	208844	196	218023	471	216935	1061
8 . .	203087	204457	145	208695	344	217868	625
10 . .	199321	202971	116	204340	261	208576	463
12 . .	194014	199213	108	202862	225	204230	371
14 . .	188769	193898	116	199094	227	202740	346
16 . .	160080	188610	158	193735	279	198927	394
18 . .	122108	156176	3809	184010	4638	189010	4875
20 . .	77638	105379	14656	134727	21240	158814	25389
22 . .	42123	53845	17410	73260	30103	93822	39707
24 . .	24803	22345	14835	28999	25255	41379	41351
26 . .	16905	16482	7344	14864	13908	18937	20024
28 . .	14207	12525	5157	12257	8697	11218	11731
30 . .	14167	11731	4300	10360	6600	10267	8555
32 . .	13093	12521	4226	10376	6041	9191	6953
34 . .	13689	12150	3636	11815	6529	9809	6998
36 . .	13785	13260	3575	11886	5970	11744	8238
38 . .	14841	13747	3349	13283	6049	12043	7596
40 . .	15405	15151	3396	14074	5971	13667	7982
42 . .	15146	16037	3298	15860	6353	14794	8199
44 . .	16640	15974	3070	17018	6272	16927	9070
46 . .	18984	17724	3122	17148	5882	18366	9034
48 . .	20153	20491	3486	19108	6077	18623	8576

random variables are of course not independent. The covariances are available on magnetic tape.

The smallness of the variances in all age groups should be noted. Consider single males in age group 0- at time $t = 1$ for example: the variance is 168,-347, so that the standard deviation is 410. No demographer would predict 216,300 single males in age group 0- with a standard deviation of 410! We conclude that birth probabilities, marriage probabilities, divorce probabilities, etc. must be considered as random variables themselves. [Their distributions may of course change with time.]

This important fact, mentioned in §3 appears to have been noticed by only one other author (Z. M. Sykes, 1967), although it is obvious from the simple numerical example in §3.

Some of the expectations in Table 6 undergo substantial changes from one time period to the next and these need explanation. The population being investigated is an hypothetical one; the marriage rates etc. are also hypothetical, and not necessarily the ones experienced in the past to give the population its present hypothetical form. Thus, the substantial changes in the expectations in Table 6 from one time period to the next are caused by a sudden change in marriage rates etc. at time $t = 0$.

The results obtained using this model will be different from those obtained using the simpler one-sex Leslie approach. It is of interest to compare some of the numerical results for the two different models. The obvious expectation for comparison purposes is the expected number of females aged 0 at times $t = 1$, 2, 3. The one-sex age-specific female birth rates may be calculated using the illegitimate female birth rates of Appendix Table 2, the legitimate female birth rates from Appendix Table 3, the initial single female population of Table 5 and the initial married population of Table 6. This method of calculating the one-sex, age-specific female birth rates means that the expected number of females aged 0 at time $t = 1$ for the one-sex model is the same as the expected

TABLE 6.—Married Population

Ages M	Ages F	Initial married population	Married population at time t = 1 unit: Expected	Variance	Married population at time t = 2 units: Expected	Variance	Married population at time t = 3 units: Expected	Variance
20	18	5998	1079	1072	1271	1263	1306	1297
20	20	11184	1861	1833	2380	2345	2804	2763
20	22	2226	634	629	861	855	1101	1093
20	24	2389	217	215	277	276	376	364
20	26	30	3	3	2	2	3	3
20	28	28	22	22	26	26	30	30
22	18	5595	1311	1300	1545	1532	1587	1573
22	20	22240	11165	5010	7723	7350	9100	8660
22	22	2489	15990	4638	8467	7881	10826	10092
22	24	3043	4274	1978	3268	3132	4444	4283
22	26	36	2475	125	310	308	398	396
22	28	34	238	178	225	162	208	207
22	30	28	60	31	57	50	63	53
24	18	4769	728	725	858	854	881	877
24	20	13952	10475	4744	7585	7226	8937	8513
24	22	15660	30255	7439	22164	14143	21827	18844
24	24	16032	8911	5473	24107	11015	19607	16469
24	26	1571	4025	992	5145	2773	4418	4161
24	28	1876	540	483	2944	624	751	741
24	30	217	195	131	380	316	363	299
24	32	128	67	40	92	64	86	78
26	18	3493	359	358	422	421	434	433
26	20	5649	7417	2651	4150	4042	4890	4762
26	22	37286	19180	5064	17632	10976	16778	15015
26	24	38735	19293	3676	34189	10753	26927	17308
26	26	11463	17133	1378	10219	6625	25526	12257
26	28	15008	1999	457	4478	1491	5677	3313
26	30	3035	1997	163	693	635	3098	828
26	32	199	283	71	269	205	463	381
26	34	226	127	2	67	40	92	64
28	18	1039	171	171	202	202	207	207
28	20	2338	4731	1279	1974	1949	2326	2297
28	22	20444	8411	2758	11169	6197	8993	8486
28	24	19708	40533	3647	23602	8842	23457	14423
28	26	22107	41161	2342	21656	5443	37158	12479
28	28	24432	12324	1066	17904	2322	10975	7005
28	30	4394	15364	715	2456	908	4893	1899
28	32	2176	3290	323	2215	416	898	823
28	34	1332	300	105	378	170	347	283
28	36	465	250	29	149	27	89	63
30	18	370	78	78	92	92	95	95
30	20	733	1683	659	988	728	1121	1016
30	22	10944	3751	1430	6634	3175	4402	3464
30	24	11187	22419	2277	11117	5274	14849	8979
30	26	21924	21241	1858	41653	5218	25388	10296
30	28	24107	22859	1227	41654	3525	22290	6085
30	30	21102	24798	982	12770	1715	18262	2942
30	32	4186	4780	361	15539	1205	2757	1209
30	34	3143	2370	220	3458	519	2363	601
30	36	2338	1404	107	376	185	451	246

TABLE 6.—Married Population (Continued)

Ages M	Ages F	Initial married population	Married population at time t = 1 unit: Expected	Variance	Married population at time t = 2 units: Expected	Variance	Married population at time t = 3 units: Expected	Variance
30	38	422	481	28	268	53	166	47
32	20	276	491	127	186	185	187	186
32	22	4103	1416	684	2265	1227	1509	1229
32	24	7011	12136	1397	5306	2691	7978	4249
32	26	20584	12113	1159	23098	3316	12224	6300
32	28	23978	22458	1073	21763	2733	41824	6116
32	30	29118	24470	1021	23156	2032	41665	4415
32	32	14906	21378	847	24923	1701	13010	2179
32	34	9292	4466	293	5015	667	15582	1546
32	36	4631	3283	176	2498	362	3555	674
32	38	3161	2380	106	1457	186	435	239
32	40	680	436	27	493	53	282	74
34	22	1629	364	94	576	219	260	258
34	24	2139	4560	542	1898	1153	2664	1613
34	26	13600	7565	714	12561	2041	5990	3341
34	28	11499	20904	848	12532	1798	23312	3954
34	30	26171	24204	897	22637	1793	21942	3333
34	32	27308	29294	977	24611	1745	23257	2624
34	34	19676	15134	636	21494	1489	24916	2273
34	36	14779	9412	379	4650	535	5158	920
34	38	4531	4703	192	3373	334	2580	491
34	40	2494	3181	96	2408	187	1493	243
34	42	627	687	28	447	50	502	75
36	24	723	2053	456	923	592	1175	810
36	26	4259	2484	395	4837	904	2309	1556
36	28	6137	13826	585	7862	1166	12751	2474
36	30	24313	11744	571	20960	1450	12704	2227
36	32	26632	26269	836	24248	1556	22658	2336
36	34	26321	27365	818	29302	1733	24613	2321
36	36	24015	19767	627	15250	1141	21507	2027
36	38	16832	14816	463	9519	734	4802	790
36	40	5113	4613	216	4772	390	3457	502
36	42	3112	2515	99	3191	200	2425	273
36	44	689	632	25	692	54	455	71
38	26	647	929	223	2219	660	1163	835
38	28	3713	4465	321	2709	667	5003	1165
38	30	13933	6330	374	13866	1007	7989	1471
38	32	23801	24293	697	11854	988	20898	1924
38	34	25204	26598	749	26224	1505	24174	2098
38	36	24100	26296	663	27295	1499	29193	2384
38	38	25903	23957	580	19773	1124	15295	1567
38	40	15507	16796	414	14802	851	9580	1032
38	42	4513	5161	206	4667	405	4814	567
38	44	2063	3110	102	2526	190	3189	299
38	46	739	690	26	633	48	693	78
40	28	1801	805	176	1079	391	2333	816
40	30	8769	3838	242	4556	538	2825	839
40	32	16909	13950	452	6426	646	13820	1345
40	34	21776	23688	631	24160	1281	11875	1324
40	36	22408	25114	684	26450	1397	26075	2097

number for the two-sex model (207,828). The one-sex female survivorship probabilities are obtained by summing the relevant entries of Appendix Table 2. The expected members aged 0 at times $t = 1, 2, 3$ for the one-sex model are 207,828, 210,137 and 216,766 respectively. The relevant figures for the two-sex model are 207,828, 193,374 and 186, 643 respectively. These figures differ considerably. The main reason for the difference is the change in marriage rates mentioned in the previous paragraph. This factor cannot be dealt with by the one-sex model. For a population experiencing near-constant marriage rates, the figures obtained using the different models would be very much the same as each other. Another factor contributing to the difference is the fact that the present two-sex model does not allow a marriage and a birth to occur in the one time unit. This restriction was imposed on the model to simplify the initial computer program. There is no theoretical difficulty in eliminating it, and indeed this must be done in any practical situation.

6.10 A Criticism of the Stochastic Model

The stochastic model (as opposed to the method of analysis) may be criticised because it assumes that an individual man aged x makes up his mind that he desires to marry a woman aged y during a unit time interval; if there are insufficient women aged y desiring to marry a man aged x, he does not marry and does not even try to marry, as a second preference a woman aged (say) $y - 1$ during that unit time interval.

This criticism may be valid at the personal level. However, the model is essentially a macroscopic one, and the criticism then is not so well founded. Consider a cohort of young men. When aged 17-22, say, a shortage of slightly younger women will cause many of the young

TABLE 6.—Married Population (Continued)

Ages M	Ages F	Initial married population	Married population at time t = 1 unit: Expected	Variance	Married population at time t = 2 units: Expected	Variance	Married population at time t = 3 units: Expected	Variance
40	38	22609	24017	633	26178	1255	27141	2100
40	40	24170	25732	655	23820	1131	19708	1560
40	42	13533	15427	407	16696	813	14733	1195
40	44	4568	4541	176	5183	383	4698	573
40	46	2511	2070	84	3095	199	2525	275
40	48	743	737	27	688	50	632	70
42	30	1007	1887	147	895	285	1162	493
42	32	7027	8775	298	3901	420	4592	701
42	34	14706	16806	461	13889	831	6462	864
42	36	17803	21632	594	23485	1194	23943	1812
42	38	21817	22287	571	24941	1301	26228	1985
42	40	22333	22458	624	23855	1221	25982	1817
42	42	21479	23933	623	25479	1279	23604	1666
42	44	14179	13423	379	15290	810	16538	1214
42	46	4558	4572	158	4544	334	5178	544
42	48	2511	2503	95	2068	164	3071	294
44	32	1513	1086	115	1934	254	950	359
44	34	7920	7017	242	8728	552	3924	563
44	36	12136	14593	424	16636	879	13769	1175
44	38	16780	17668	505	21419	1135	23219	1717
44	40	20883	21625	592	22097	1127	24695	1888
44	42	20721	22088	607	22240	1222	23620	1772
44	44	18135	21199	575	23626	1236	25153	1893
44	46	13758	13991	405	13267	746	15103	1201
44	48	4212	4535	169	4556	310	4529	482
46	34	1246	1558	103	1136	201	1955	339
46	36	8061	7865	256	6969	457	8643	783
46	38	11112	12029	369	14432	812	16421	1271
46	40	16007	16620	483	17478	979	21148	1648
46	42	18172	20637	587	21364	1154	21833	1670

TABLE 6.—Married Population (Continued)

Ages M F	Initial married population	Married population at time t = 1 unit		Married population at time t = 2 units		Married population at time t = 3 units	
		Expected	Variance	Expected	Variance	Expected	Variance
46 44	19803	20437	598	21775	1201	21948	1806
46 46	22197	17858	533	20853	1144	23247	1841
46 48	13979	13532	413	13760	801	13068	1112
48 36	1323	1284	91	1582	189	1167	271
48 38	8114	7972	258	7777	493	6892	656
48 40	10775	10990	360	11878	719	14224	1185
48 42	13480	15799	502	16395	963	17224	1445
48 44	18432	17917	537	20323	1163	21033	1707
48 46	19973	19472	624	20084	1195	21391	1799
48 48	21004	21760	675	17526	1064	20446	1715
50 38	2032	1345	86	1305	169	1590	263
50 40	7365	7993	274	7849	508	7656	723
50 42	10217	10608	372	10817	713	11674	1065
50 44	14113	13256	470	15522	1005	16101	1448
50 46	18084	18084	597	17581	1089	19924	1736
50 48	18876	19550	684	19067	1255	19656	1801
52 40	1070	2025	106	1354	164	1313	239
52 42	6371	7222	274	7831	547	7686	757
52 44	7009	10003	382	10388	745	10589	1069
52 46	13403	13797	534	12961	941	15165	1511
52 48	17602	17656	666	17653	1206	17164	1647
54 42	1002	1078	77	2001	205	1349	235
54 44	5722	6215	264	7040	549	7628	822
54 46	9001	6847	314	9734	764	10111	1117
54 48	12402	13040	551	13413	1067	12604	1412
56 44	2373	1000	73	1076	149	1963	301
56 46	5767	5548	268	6019	527	6816	824
56 48	9117	8717	426	6644	625	9413	1146
58 46	2132	2299	134	986	141	1061	215
58 48	5211	5554	303	5342	533	5790	787

men to wait longer before marriage, and then to choose a bride whose age differs from his own by a greater amount. This process will be reflected in the stochastic model.

The stochastic model has been constructed in order to study the behaviour of the whole population, and consequently, this criticism of the model does not cause us much concern.

7. CONCLUSION

The demographer is frequently faced with the problem of investigating the effect on a population of a change in marriage rates, or of divorce rates, or due to changes in economic conditions, or due to changes in government immigration policy, etc. The present two-sex model permits objective numerical investigations of some of these problems to be carried out on a digital computer. The demographer need only change some data constants at specified times, and then look to see what happens to first and second order moments. We describe a recurrence method for expectations and second order moments, which with a slight modification for marriage, is the usual multi-type Galton-Watson recurrence relation (J. H. Pollard, 1966).

Simpler two-sex models exist (e.g., L. A. Goodman, 1968; A. H. Pollard, 1948). These models are useful for proving certain mathematical results, but they cannot be used for detailed projection purposes. The present model is too complex to prove sophisticated mathe-

APPENDIX TABLE 1.—Male Probabilities

[The format of this table is explained in section 6.8.]

Age	Merely survive	Probabilities of desiring marriage					Indi-cator
0	.99699	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	1
2	.99844	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	1
4	.99888	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	1
6	.99896	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	1
8	.99915	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	1
10	.99911	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	1
12	.99885	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	1
14	.99830	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	1
16	.99714	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	1
18	.92384	16 .02908	18 .03105	20 .00990	22 .00219	24 .00031	1
18	.00000	26 .00015	0 .00000	0 .00000	0 .00000	0 .00000	2
20	.74094	16 .04001	18 .09814	20 .08599	22 .02356	24 .00616	1
20	.00000	26 .00161	28 .00025	0 .00000	0 .00000	0 .00000	2
22	.53447	16 .02954	18 .11985	20 .18598	22 .09505	24 .02320	1
22	.00000	26 .00603	28 .00186	30 .00093	0 .00000	0 .00000	2
24	.52011	16 .09655	18 .17459	20 .12123	22 .05613	24 .01869	1
24	.00000	26 .00663	28 .00213	30 .00102	32 .00001	0 .00000	2
26	.54989	16 .02792	18 .07460	20 .11854	22 .10480	24 .06178	1
26	.00000	26 .03112	28 .01419	30 .00778	32 .00412	34 .00229	2
28	.61182	16 .00429	18 .02378	20 .06353	22 .10095	24 .08925	1
28	.00000	26 .05261	28 .02650	30 .01208	32 .00663	34 .00351	2
28	.00000	36 .00195	0 .00000	0 .00000	0 .00000	0 .00000	3
30	.67297	18 .00360	20 .01999	22 .05341	24 .08487	26 .07504	1
30	.00000	28 .04424	30 .02228	32 .01016	34 .00557	36 .00295	2
30	.00000	38 .00164	0 .00000	0 .00000	0 .00000	0 .00000	3
32	.72347	20 .00304	22 .01685	24 .04504	26 .07156	28 .06327	1
32	.00000	30 .03730	32 .01879	34 .00856	36 .00470	38 .00248	2
32	.00000	40 .00138	0 .00000	0 .00000	0 .00000	0 .00000	3
34	.81564	22 .02253	24 .02414	26 .02448	28 .02368	30 .02207	1
34	.00000	32 .01908	34 .01563	36 .01149	38 .00805	40 .00552	2
34	.00000	42 .00368	0 .00000	0 .00000	0 .00000	0 .00000	3
36	.84215	24 .01913	26 .02050	28 .02079	30 .02011	32 .01874	1
36	.00000	34 .01620	36 .01327	38 .00976	40 .00683	42 .00468	2
36	.00000	44 .00312	0 .00000	0 .00000	0 .00000	0 .00000	3
38	.86316	26 .01638	28 .01755	30 .01780	32 .01722	34 .01605	1
38	.00000	36 .01387	38 .01137	40 .00836	42 .00585	44 .00401	2
38	.00000	46 .00267	0 .00000	0 .00000	0 .00000	0 .00000	3
40	.87973	28 .01416	30 .01517	32 .01539	34 .01488	36 .01387	1
40	.00000	38 .01199	40 .00982	42 .00722	44 .00506	46 .00347	2
40	.00000	48 .00231	0 .00000	0 .00000	0 .00000	0 .00000	3
42	.89152	30 .01276	32 .01367	34 .01387	36 .01341	38 .01250	1
42	.00000	40 .01081	42 .00885	44 .00651	46 .00456	48 .00313	2
44	.90461	32 .01122	34 .01202	36 .01219	38 .01179	40 .01099	1
44	.00000	42 .00950	44 .00778	46 .00572	48 .00401	0 .00000	2
46	.91354	34 .01023	36 .01096	38 .01111	40 .01075	42 .01002	1
46	.00000	44 .00866	46 .00710	48 .00522	0 .00000	0 .00000	2
48	.92166	36 .00938	38 .01005	40 .01019	42 .00986	44 .00919	1
48	.00000	46 .00794	48 .00651	0 .00000	0 .00000	0 .00000	2
50	.93069	38 .00838	40 .00898	42 .00911	44 .00881	46 .00821	1
50	.00000	48 .00710	0 .00000	0 .00000	0 .00000	0 .00000	2
52	.93798	40 .00753	42 .00807	44 .00818	46 .00791	48 .00737	1
54	.94398	42 .00667	44 .00714	46 .00725	48 .00701	0 .00000	1
56	.94730	44 .00595	46 .00638	48 .00647	0 .00000	0 .00000	1
58	.94778	46 .00538	48 .00576	0 .00000	0 .00000	0 .00000	1

APPENDIX TABLE 2.—Female Probabilities

[The format of this table is explained in section 6.8]

Age	Merely survive	Probabilities of desiring marriage					Indicator
0	.99726	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	-1
2	.99878	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	-1
4	.99906	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	-1
6	.99929	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	-1
8	.99943	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	-1
10	.99946	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	-1
12	.99940	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	-1
14	.99636	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	-1
14	.00143	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	-2
14	.00137	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	-3
16	.96111	18 .00674	20 .00819	22 .00455	24 .00224	26 .00107	-1
16	.00740	28 .00049	0 .00000	0 .00000	0 .00000	0 .00000	-2
16	.00710	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	-3
18	.83096	18 .01524	20 .04258	22 .04022	24 .02194	26 .01036	-1
18	.01505	28 .00535	30 .00267	0 .00000	0 .00000	0 .00000	-2
18	.01445	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	-3
20	.63723	18 .00817	20 .06280	22 .10505	24 .06864	26 .03621	-1
20	.02346	28 .01848	30 .01054	32 .00566	0 .00000	0 .00000	-2
20	.02254	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	-3
22	.39074	18 .00514	20 .04899	22 .15289	24 .13881	26 .08553	-1
22	.03315	28 .05184	30 .03124	32 .01813	34 .01049	0 .00000	-2
22	.03185	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	-3
24	.46321	18 .00011	20 .00427	22 .04063	24 .12680	26 .11512	-1
24	.04335	28 .07094	30 .04300	32 .02591	34 .01504	36 .00870	-2
24	.04165	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	-3
26	.55335	18 .00405	20 .01382	22 .02985	24 .04644	26 .05879	-1
26	.04983	28 .06045	30 .04884	32 .03527	34 .02433	36 .01574	-2
26	.04787	38 .00995	0 .00000	0 .00000	0 .00000	0 .00000	-3
28	.60809	20 .00336	22 .01147	24 .02477	26 .03853	28 .04877	-1
28	.05202	30 .05015	32 .04052	34 .02926	36 .02018	38 .01306	-2

matical results, but it is useful for projection purposes. We should expect many of the multi-type Galton-Watson results to apply to it, however.

The recurrence method for expectations and central quadratic moments involves one approximation, which we have shown (numerically) to be very accurate. An analytical study of the error involved presents an enormous problem, and the analytical results of §6.6 are hardly even a beginning.

It has been suggested that we could calculate the moments of X^+ more accurately if we knew the higher-order moments of X. This is undoubtedly true, but we need to apply the method recursively and the higher-order moments themselves are then highly suspect. Furthermore, the additional computation would be enormous, and there would be difficulties finding suitable distributions to effect the approximations.

It could be argued that it would be better to simulate the population and hence not need to use an approximate method of analysis. Clearly it is a question of computer time, and Monte Carlo methods are notorious for consuming time. Possibly the expectations could be obtained by suitable simulation, but the time required to get reliable estimates of the second-order moments would be prohibitive.

A computer program of some generality has been written and used to project an hypothetical population. The smallness of the projected variances is noted in §6.9. This fact leads us to the conclusion that fluctuations in population data are caused by two different sources of variation:

(i) statistical fluctuations due to the finite numbers in the population; and

(ii) random fluctuations in the actual probabilities.

Usually the second source of variation is the greater, although it is generally ignored by mathematical demographers. A simple numerical example in section 6.9 clearly illustrates the importance of the second source of variability. The distributions of the random probabilities must be investigated thoroughly before more accurate projections can be made. The methods of §3, which allow for this source of variation, are readily incorporated in the two-sex model.

Nothing has been said about the availability of suitable demographic data to be used in this type of analysis. Many of the probabilities required are at present available, and indeed the only considerable difficulty is that of obtaining the age-specific probabilities of desiring marriage. There is no obvious simple manner of calculating such probabilities, and some 'high-class cookery' method will probably be necessary. (See H. Tetley, 1950, Vol. 1, p. 263.) The methods of preparing life tables are often of this nature, so such a method for obtaining age specific probabilities of desiring marriage should not be too distasteful.

Appendix Table 2.—Female Probabilities (Continued)

Age	Merely survive	Probabilities of desiring marriage					Indicator
28	.04998	40 .00826	0 .00000	0 .00000	0 .00000	0 .00000	-3
30	.65824	22 .00280	24 .00955	26 .02063	28 .03209	30 .04062	-1
30	.05090	32 .04176	34 .03374	36 .02437	38 .01681	40 .01087	-2
30	.04890	42 .00688	0 .00000	0 .00000	0 .00000	0 .00000	-3
32	.70685	24 .00232	26 .00791	28 .01709	30 .02659	32 .03366	-1
32	.04692	34 .03461	36 .02796	38 .02020	40 .01393	42 .00901	-2
32	.04508	44. 00570	0 .00000	0 .00000	0 .00000	0 .00000	-3
34	.75261	26 .00193	28 .00659	30 .01423	32 .02213	34 .02802	-1
34	.04039	36 .02881	38 .02327	40 .01681	42 .01159	44 .00750	-2
34	.03881	46 .00474	0 .00000	0 .00000	0 .00000	0 .00000	-3
36	.79570	28 .00162	30 .00552	32 .01192	34 .01855	36 .02348	-1
36	.03188	38 .02414	40 .01950	42 .01409	44 .00971	46 .00628	-2
36	.03062	48 .00397	0 .00000	0 .00000	0 .00000	0 .00000	-3
38	.83297	30 .00138	32 .00469	34 .01013	36 .01576	38 .01995	-1
38	.02321	40 .02051	42 .01657	44 .01197	46 .00825	48 .00534	-2
38	.02229	50 .00338	0 .00000	0 .00000	0 .00000	0 .00000	-3
40	.86859	32 .00115	34 .00392	36 .00847	38 .01318	40 .01668	-1
40	.01454	42 .01716	44 .01386	46 .01001	48 .00690	50 .00447	-2
40	.01396	52 .00282	0 .00000	0 .00000	0 .00000	0 .00000	-3
42	.89718	34 .00098	36 .00333	38 .00720	40 .01120	42 .01418	-1
42	.00704	44 .01458	46 .01178	48 .00851	50 .00587	52 .00379	-2
42	.00676	54 .00240	0 .00000	0 .00000	0 .00000	0 .00000	-3
44	.91940	36 .00080	38 .00274	40 .00591	42 .00920	44 .01165	-1
44	.00281	46 .01197	48 .00967	50 .00699	52 .00482	54 .00312	-2
44	.00269	56 .00197	0 .00000	0 .00000	0 .00000	0 .00000	-3
46	.93192	38 .00069	40 .00236	42 .00511	44 .00794	46 .01005	-1
46	.00061	48 .01034	50 .00835	52 .00603	54 .00416	56 .00269	-2
46	.00059	58 .00170	0 .00000	0 .00000	0 .00000	0 .00000	-3
48	.94250	40 .00058	42 .00199	44 .00429	46 .00668	48 .00845	-1
48	.00005	50 .00869	52 .00702	54 .00507	56 .00350	58 .00226	-2
48	.00005	0 .00000	0 .00000	0 .00000	0 .00000	0 .00000	-3

APPENDIX TABLE 3.—Couple Probabilities

Ages M F	Merely survive	Husband dies	Wife dies	Son born	Daughter born	Divorce
20 18.	.18439	.00334	.00118	.42323	.39702	.00084
20 20.	.21093	.00334	.00122	.39927	.38361	.00163
20 22.	.24070	.00334	.00120	.38363	.36859	.00254
20 24.	.28626	.00334	.00127	.35975	.34564	.00374
20 26.	.33412	.00334	.00142	.33496	.32182	.00434
20 28.	.38354	.00334	.00158	.30916	.29704	.00534
22 18.	.21458	.00309	.00118	.39775	.38215	.00125
22 20.	.24112	.00309	.00122	.38379	.36874	.00204
22 22.	.27090	.00309	.00120	.36815	.35371	.00295
22 24.	.31645	.00309	.00127	.34427	.33077	.00415
22 26.	.36431	.00309	.00142	.31948	.30695	.00475
22 28.	.41374	.00309	.00158	.29368	.28216	.00575
22 30.	.45372	.00309	.00184	.27344	.26272	.00519
24 18.	.25292	.00291	.00118	.37773	.36292	.00234
24 20.	.27947	.00291	.00122	.36377	.34950	.00313
24 22.	.30924	.00291	.00120	.34813	.33448	.00404
24 24.	.35480	.00291	.00127	.32425	.31153	.00524
24 26.	.40265	.00291	.00142	.29946	.28772	.00584
24 28.	.45208	.00291	.00158	.27366	.26293	.00684
24 30.	.49207	.00291	.00184	.25342	.24348	.00628
24 32.	.52833	.00291	.00217	.23493	.22572	.00594
26 18.	.28912	.00297	.00118	.35856	.34450	.00367
26 20.	.31566	.00297	.00122	.34460	.33109	.00446
26 22.	.34544	.00297	.00120	.32896	.31606	.00537
26 24.	.39099	.00297	.00127	.30508	.29312	.00657
26 26.	.43885	.00297	.00142	.28029	.26930	.00717
26 28.	.48828	.00297	.00158	.25449	.24451	.00817
26 30.	.52827	.00297	.00184	.23425	.22506	.00761
26 32.	.56453	.00297	.00217	.21576	.20730	.00727
26 34.	.59799	.00297	.00257	.19854	.19075	.00718
28 18.	.33065	.00310	.00118	.33704	.32382	.00421
28 20.	.35719	.00310	.00122	.32308	.31041	.00500
28 22.	.38697	.00310	.00120	.30744	.29538	.00591
28 24.	.43252	.00310	.00127	.28356	.27244	.00711
28 26.	.48038	.00310	.00142	.25877	.24862	.00771
28 28.	.52981	.00310	.00158	.23297	.22383	.00871
28 30.	.56979	.00310	.00184	.21273	.20439	.00815
28 32.	.60606	.00310	.00217	.19424	.18662	.00781
28 34.	.63951	.00310	.00257	.17702	.17008	.00772
28 36.	.65735	.00310	.00302	.16794	.16135	.00724
30 18.	.37407	.00328	.00118	.31434	.30201	.00512
30 20.	.40061	.00328	.00122	.30038	.28860	.00591
30 22.	.43039	.00328	.00120	.28474	.27357	.00682
30 24.	.47594	.00328	.00127	.26086	.25063	.00802
30 26.	.52380	.00328	.00142	.23607	.22681	.00862
30 28.	.57284	.00328	.00158	.21027	.20241	.00962
30 30.	.61321	.00328	.00184	.19003	.18258	.00906
30 32.	.64948	.00328	.00217	.17154	.16481	.00872
30 34.	.68293	.00328	.00257	.15432	.14827	.00863
30 36.	.70077	.00328	.00302	.14524	.13954	.00815

ACKNOWLEDGMENTS

The early stages of this research were carried out in Cambridge; I should like to thank the Director of the Mathematical Laboratory at the University of Cambridge for allowing me time on TITAN, and the Shell Company of Australia for a Research Scholarship. The work was completed at the Population Research and Training Center at the

University of Chicago, and I should like to acknowledge a grant to the Population Center from the Ford Foundation.

REFERENCES

Bartholomew, D. J. 1968. Stochastic Models for Social Processes. New York: Wiley.

Bernardelli, H. 1941. Population waves. Journal of Burma Research Society 31: 1–18.

Feller, W. 1957. An Introduction to Probability Theory and Its Applications. Second edition, Vol. 1. New York: Wiley.

APPENDIX TABLE 3.—Couple Probabilities (Continued)

Ages M F	Merely survive	Husband dies	Wife dies	Son born	Daughter born	Divorce
30 38.	.72468	.00328	.00360	.13294	.12773	.00777
32 20.	.44042	.00356	.00122	.28002	.26904	.00574
32 22.	.47020	.00356	.00120	.26438	.25401	.00665
32 24.	.51575	.00356	.00127	.24050	.23107	.00785
32 26.	.56361	.00356	.00142	.21571	.20725	.00845
32 28.	.61304	.00356	.00158	.18991	.18246	.00945
32 30.	.65302	.00356	.00184	.16967	.16302	.00889
32 32.	.68929	.00356	.00217	.15118	.14525	.00855
32 34.	.72274	.00356	.00257	.13396	.12871	.00846
32 36.	.74058	.00356	.00302	.12488	.11998	.00798
32 38.	.76449	.00356	.00360	.11258	.10817	.00760
32 40.	.93179	.00356	.00429	.02692	.02588	.00756
34 22.	.51056	.00401	.00120	.24368	.23412	.00643
34 24.	.55611	.00401	.00127	.21980	.21118	.00763
34 26.	.60397	.00401	.00142	.19501	.18736	.00823
34 28.	.65340	.00401	.00158	.16921	.16257	.00923
34 30.	.69338	.00401	.00184	.14897	.14313	.00867
34 32.	.72965	.00401	.00217	.13048	.12536	.00833
34 34.	.76310	.00401	.00257	.11326	.10882	.00824
34 36.	.78094	.00401	.00302	.10418	.10009	.00776
34 38.	.80485	.00401	.00360	.09188	.08828	.00738
34 40.	.93156	.00401	.00429	.02692	.02588	.00734
34 42.	.95898	.00401	.00520	.01252	.01202	.00727
36 24.	.59257	.00472	.00127	.20094	.19306	.00744
36 26.	.64043	.00472	.00142	.17615	.16924	.00804
36 28.	.68986	.00472	.00158	.15035	.14445	.00904
36 30.	.72984	.00472	.00184	.13011	.12501	.00848
36 32.	.76611	.00472	.00217	.11162	.10724	.00814
36 34.	.79956	.00472	.00257	.09440	.09070	.00805
36 36.	.81740	.00472	.00302	.08532	.08197	.00757
36 38.	.84131	.00472	.00360	.07302	.07016	.00719
36 40.	.93104	.00472	.00429	.02692	.02588	.00715
36 42.	.95846	.00472	.00520	.01252	.01202	.00708
36 44.	.97402	.00472	.00626	.00404	.00390	.00706
38 26.	.66276	.00571	.00142	.16433	.15789	.00789
38 28.	.71219	.00571	.00158	.13853	.13310	.00889
38 30.	.75218	.00571	.00184	.11829	.11365	.00833
38 32.	.78844	.00571	.00217	.09980	.09589	.00799
38 34.	.82190	.00571	.00257	.08258	.07934	.00790
38 36.	.83973	.00571	.00302	.07350	.07062	.00742
38 38.	.86365	.00571	.00360	.06120	.05880	.00704
38 40.	.93020	.00571	.00429	.02692	.02588	.00700
38 42.	.95762	.00571	.00520	.01252	.01202	.00693
38 44.	.97318	.00571	.00626	.00404	.00390	.00691
38 46.	.97822	.00571	.00746	.00100	.00098	.00663
40 28.	.73402	.00693	.00158	.12688	.12190	.00869
40 30.	.77400	.00693	.00184	.10664	.10246	.00813
40 32.	.81027	.00693	.00217	.08815	.08469	.00779
40 34.	.84372	.00693	.00257	.07093	.06815	.00770
40 36.	.86156	.00693	.00302	.06185	.05942	.00722

APPENDIX TABLE 3.—Couple Probabilities (Continued)

Ages M F	Merely survive	Husband dies	Wife dies	Son born	Daughter born	Divorce
40 38.	.88547	.00693	.00360	.04955	.04761	.00684
40 40.	.92918	.00693	.00429	.02692	.02588	.00680
40 42.	.95660	.00693	.00520	.01252	.01202	.00673
40 44.	.97216	.00693	.00626	.00404	.00390	.00671
40 46.	.97720	.00693	.00746	.00100	.00098	.00643
40 48.	.97820	.00693	.00887	.00008	.00008	.00584
42 30.	.78624	.00841	.00184	.09974	.09583	.00794
42 32.	.82251	.00841	.00217	.08125	.07806	.00760
42 34.	.85596	.00841	.00257	.06403	.06152	.00751
42 36.	.87379	.00841	.00302	.05495	.05280	.00703
42 38.	.89771	.00841	.00360	.04265	.04098	.00665
42 40.	.92789	.00841	.00429	.02692	.02588	.00661
42 42.	.95531	.00841	.00520	.01252	.01202	.00654
42 44.	.97087	.00841	.00626	.00404	.00390	.00652
42 46.	.97591	.00841	.00746	.00100	.00098	.00624
42 48.	.97691	.00841	.00887	.00008	.00008	.00565
44 32.	.83343	.01017	.00217	.07496	.07202	.00725
44 34.	.86688	.01017	.00257	.05774	.05548	.00716
44 36.	.88472	.01017	.00302	.04866	.04675	.00668
44 38.	.90864	.01017	.00360	.03636	.03493	.00630
44 40.	.92648	.01017	.00429	.02692	.02588	.00626
44 42.	.95390	.01017	.00520	.01252	.01202	.00619
44 44.	.96946	.01017	.00626	.00404	.00390	.00617
44 46.	.97450	.01017	.00746	.00100	.00098	.00589
44 48.	.97550	.01017	.00887	.00008	.00008	.00530
46 34.	.87429	.01241	.00257	.05300	.05092	.00681
46 36.	.89212	.01241	.00302	.04392	.04220	.00633
46 38.	.91604	.01241	.00360	.03162	.03038	.00595
46 40.	.92459	.01241	.00429	.02692	.02588	.00591
46 42.	.95201	.01241	.00520	.01252	.01202	.00584

Goodman, L. A. 1953. Population growth of the sexes. Biometrics 9: 212–225.

———. 1968. Stochastic models for the population growth of the sexes. Manuscript, to be published.

Harris, T. E. 1963. The Theory of Branching Processes. Berlin: Springer-Verlag.

Karmel, P. H. 1947. The relation between male and female reproduction rates. Population Studies 1: 249–274.

Kendall, D. G. 1949. Stochastic processes and population growth. Journal of the Royal Statistical Society, B, 11: 230–264.

Kendall, M. G. 1948. The Advanced Theory of Statistics. Fourth edition, Vol. 1, London: Griffin.

Lancaster, H. O. 1958. The structure of bivariate distributions. Annals of Mathematical Statistics 29: 719–736.

Leslie, P. H. 1945. On the use of matrices in certain population mathematics. Biometrika 33: 183–212.

Lewis, E. G. 1942. On the generation and growth of a population. Sankhya 6: 93–96.

Pollard, A. H. 1948. The measurement of reproductivity. Journal of the Institute of Actuaries 74: 288–318.

Pollard, J. H. 1966. On the use of the direct matrix product in analyzing certain stochastic population models. Biometrika 53: 397–415.

———. 1967. Hierarchical population models with Poisson recruitment. Journal of Applied Probability 4: 209–213.

———. 1968a. The multi-type Galton-Watson process in a genetical context. Biometrics 24: 147–158.

———. 1968b. A note on multi-type Galton-Watson processes with random branching probabilities. Biometrika 55: 589–590.

Riordan, J. 1958. An Introduction to Combinatorial Analysis. New York: Wiley.

Sykes, Z. M. 1967. A stochastic version of the matrix model for population dynamics. Paper read at the International Biometrics Conference in Sydney.

Tetley, H. 1950. Actuarial Statistics. Second edition, Vol. I. Cambridge: Cambridge University Press.

APPENDIX TABLE 3.—Couple Probabilities (Continued)

Ages M F	Merely survive	Husband dies	Wife dies	Son born	Daughter born	Divorce
46 44.	.96757	.01241	.00626	.00404	.00390	.00582
46 46.	.97261	.01241	.00746	.00100	.00098	.00554
46 48.	.97361	.01241	.00887	.00008	.00008	.00495
48 36.	.89560	.01522	.00302	.04076	.03916	.00624
48 38.	.91952	.01522	.00360	.02846	.02734	.00586
48 40.	.92187	.01522	.00429	.02692	.02588	.00582
48 42.	.94929	.01522	.00520	.01252	.01202	.00575
48 44.	.96485	.01522	.00626	.00404	.00390	.00573
48 46.	.96989	.01522	.00746	.00100	.00098	.00545
48 48.	.97089	.01522	.00887	.00008	.00008	.00486
50 38.	.88303	.01872	.00360	.04530	.04354	.00581
50 40.	.91842	.01872	.00429	.02692	.02588	.00577
50 42.	.94584	.01872	.00520	.01252	.01202	.00570
50 44.	.96140	.01872	.00626	.00404	.00390	.00568
50 46.	.96644	.01872	.00746	.00100	.00098	.00540
50 48.	.96744	.01872	.00887	.00008	.00008	.00431
52 40.	.91449	.02296	.00429	.02692	.02588	.00546
52 42.	.94191	.02296	.00520	.01252	.01202	.00539
52 44.	.95747	.02296	.00626	.00404	.00390	.00537
52 46.	.96251	.02296	.00746	.00100	.00098	.00509
52 48.	.96351	.02296	.00887	.00008	.00008	.00450
54 42.	.93707	.02795	.00520	.01252	.01202	.00524
54 44.	.95263	.02795	.00626	.00404	.00390	.00522
54 46.	.95767	.02795	.00746	.00100	.00098	.00494
54 48.	.95867	.02795	.00887	.00008	.00008	.00435
56 44.	.94720	.03390	.00626	.00404	.00390	.00470
56 46.	.95224	.03390	.00746	.00100	.00098	.00442
56 48.	.95324	.03390	.00887	.00008	.00008	.00383
58 46.	.94552	.04108	.00746	.00100	.00098	.00396
58 48.	.94652	.04108	.00887	.00008	.00008	.00337