Arbeidsnotater

STATISTISK SENTRALBYRÅ

Dronningensgt. 16, Oslo-Dep., Oslo 1. Tlf. 41 38 20, 41 36 60

WORKING PAPERS FROM THE CENTRAL BUREAU OF STATISTICS OF NORWAY

IO 72/11 5. juni 1972

INHOMOGENEOUS SEMI-MARKOV PROCESSES, SELECT ACTUARIAL TABLES, AND DURATION-DEPENDENCE IN DEMOGRAPHY,

by Jan M. Hoem,

CONTENTS

This is a preliminary version of the present paper. The author solicits comments and welcomes any suggestions which will improve upon the account. The final version of the paper is scheduled to appear shortly in the Proceedings of the Symposium on Population Dynamics at The Mathematics Research Center of the University of Wisconsin, Madison, Wisc., June 19-21, 1972.

The author is already aware of the following additional references, which will be incorporated before printing.

{1} Barrett, J.C. (1969): "A Monte Carlo simulation of human reproduction." Genus **25** : 1-22 (Subsection 2.F).

{2} Cohen, Joel (to appear): "When does a leaky compartment model appear to have no leaks?" Theor. Popul. Biol. (Subsection 6.D).

{3} Ferriss, Abbott L. (1970): "An indicator of marriage dissolution by marriage cohort." Social Forces **48** : 356-365 (Subsection 2.D).

{4} Higham, J.A. (1851): "On the value of selection as exercised by the policy-holder against the company." J. Inst. Act. **1** : 179-202 (Subsection 7.E).

{5} Holmberg, Ingvar (1970): "Fecundity, Fertility and Family Planning: Application of Demographic Micromodels." Gothenburg: Almqvist & Wiksell (Subsection 2.F).

{6} McFarland, David D. (1970): "Intragenerational social mobility as a Markov process: Including a time-stationary Markovian model that explains observed declines in mobility rates over time." Am. Sociol. Rev. **35** : 463-476 (Subsection 7.E).

{7} Plateris, Alexander A. (1967): "Divorce Statistics Analysis, United States, 1963." In Vital and Health Statistics, Series 21, No. 13. U.S. Gov't Printing Office (Subsection 2.D).

{8} Venkatacharya, K. (1969): "Certain implications of short marital durations in the analysis of live birth intervals" and "An examination of a certain bias due to truncation in the context of simulation models of human reproduction." Sankhyā B **31** : 53-68 and 397-412 (Subsection 6.G).

{9} Wysocki, R. and A.M. Kshirsagar (1970): "Distribution of transition frequencies of a Markov renewal process over an arbitrary interval of time." Metron **28** (1-4) : 147-155 (subsidiary reference list).

S U M M A R Y

This paper starts by giving a number of demographic and actuarial examples of time-inhomogeneous semi-Markovian models. The examples are presented in a uniform terminology, viz. that of forces of transition between states in a system (state space). The states correspond to demographic or actuarial statuses, and the central features of the substantive models are reflected in the pattern of the state space and in the specification of the forces of transition. A sample path usually corresponds to the history of an individual. Jumps between states correspond to demographic events.

Forces of transition have been highly useful in a number of fields of application, yet the standard literature on semi-Markov (and the related Markov renewal) processes has found little room for them. The place of these functions in the now classical, time-homogeneous theory is pointed out briefly, and they are used as a connecting link between this theory and that of the corresponding inhomogeneous processes. The rudiments of the latter theory is then spelled out. Its basic notions are introduced; it is shown how the device of an operational time can be used to transform an inhomogeneous semi-Markov process into a homogeneous one (necessary and sufficient conditions are given and a uniqueness theorem is proved); and certain important problems connected with retrospective investigations are studied. In a final section, the multiplicity of uses of words like "select", "selected", "selection", and "selectivity" are discussed, and some parallellities between this terminology and others are pointed out.

After the "ordinary" list of references, a bibliography of some further recent articles on homogeneous Markov renewal processes is given.

1. INTRODUCTION

1.A. Since Lévy, Smith, and Takács published their first papers on semi-Markov processes in 1954/55, and particularly since Pyke's two basic papers on the related Markov renewal processes appeared in 1961, these processes have attracted widespread attention, and each year has brought a new crop of articles on the subject. The development has largely parallelled that of Markov chains, and much energy has gone into extending results in the latter theory to the new field, which has become a mature branch of stochastic processes. There now exists both a general theory of Markov renewal processes and a number of studies of applications. [A review of Markov renewal theory has been published by Çinlar (1969). A mathematically less advanced introduction has been provided by Störmer (1970; see also de Smit, 1971). Neuts (1968) has put out a working bibliography going up to 1968, and a number of later articles are given in the subsidiary list of references appearing at the end of the present paper.]

In its present form, general Markov renewal theory seems to concern itself exclusively with processes which are homogeneous in time, and interest in similar time-inhomogeneous processes has been manifested primarily in certain fields of application. In one such area, viz. actuarial science, the concept of an inhomogeneous semi-Markov process has roots going back far beyond the inception of the modern theory. The central idea in this theory is the dependence of the process on duration in the current state, and a disability model with genuine duration-dependence appeared in the actuarial literature as early as in 1924 (Schoenbaum, 1924/25).

1.B. In the present paper, we shall describe a number of inhomogeneous semi-Markovian models useful in two fields of application, viz. demography and actuarial science. We shall also indicate how our examples appear as particular cases of a more general theory, whose basic characteristics we will outline. Our purpose in doing this will be threefold:

Firstly, a pointer to some of the possibilities in this direction will hopefully help attract greater interest from people working in the theory of stochastic processes.

Secondly, population mathematicians and actuaries may find some inspiration in seeing their particular models in the perspective of a more comprehensive theory.

Thirdly, one may hope for some cross-fertilization between the two applied fields. After the early start indicated above and some subsequent pre-War work, actuarial science has witnessed only a very modest development of semi-Markovian ideas. (The early history of the subject is sketched by Seal, 1970.) These notions caught hold in population studies at a much later date, but they must have found a more fertile soil in the new area, because a considerable number of papers have been based on them since the middle 1960-s. Though closely related in many ways, the two fields have largely developed independently of each other, however, with some duplication of effort as a result. It may be useful, then, to remind the two traditions of each other.

1.C. Our mathematics will be on the intermediate level. Our interest will be focused on the substantive models described and on what can be said about them, and the mathematics will appear only as a convenient language to say it in. We do not aim at complete mathematical generality, and shall frequently impose more rigorous assumptions than what is really necessary. We will use no measure theory, so statements about what holds almost everywhere are ruled out.

We shall give a few proofs, but not where the results can be argued by "direct reasoning" (i.e. by intuition) or are well known or almost immediate. All this means that we shall sweep some interesting mathematical problems under the rug.

1.D. The examples described in the next Section have some interest in themselves, and beside this they can be seen as an introduction motivating the general theory which follows. The presentation of the examples will be made in a uniform terminology, viz. that of forces of transition between states in a system. The central features of the substantive models will be reflected in the pattern of the state space and in the specification of the forces of transition. This sometimes means that our formulation differs somewhat in outlook from the one in the literature to which we refer. Part of our purpose is to emphasize the usefulness of force functions.

We shall stress those aspects of the examples which will be useful in the later account.

2. EXAMPLES FROM DEMOGRAPHY AND ACTUARIAL SCIENCE

2.A. An actuarial model of disability. A number of models have been suggested as a basis for disability or sickness insurance. (See Seal, 1970; Hoem, 1969a, 1969b; Hoem, Riis, and Sand, to appear; and their references.) In one which is of central importance, the insured lives are simply divided into two groups, called "active lives" (or "able lives") and "disabled lives", and transitions between the two groups are described by means of two forces of transition, viz. the force of disablement $\nu(x)$ for an x-year-old active life, and the force of recovery $\rho(x,u)$ for an x-year-old disabled life whose current disability has lasted since age x-u. The two groups have forces of mortality which we call, say, $\mu(x)$ and $\eta(x,u)$, respectively. The intuitive interpretation of $\rho(x,u)$ is that $\rho(x,u)\Delta x + o(\Delta x)$ represents the probability that the disabled life will recover within age $x + \Delta x$ and stay recovered until that age. Similarly for the other forces.

In stochastic process terminology, this is an inhomogeneous semi-Markovian model with the three states "active", "disabled", and "dead". The latter is, of course, absorbing. The forces of transition are sometimes called intensities or infinitesimal transition probabilities. They are functions of the age of the individual, which makes the process inhomogeneous in time. Two of the forces (ρ and η) also depend on duration in the current state, which makes the model semi-Markovian.

2.B. A three-state fertility model. In a recent paper, Chiang (1971) suggested using essentially the same model to describe the fertility histories of individual human females. In his set-up, the two transient states are called "fecundable" (corresponding to the "active lives") and "pregnant or infecundable" (corresponding to the "disabled lives"). A woman moves from one state to another as she conceives, becomes fecundable again, and possibly dies (as fecundable or infecundable). The function $\nu(\cdot)$ could be called the force of conception, and $\rho(x,u)\Delta x + o(\Delta x)$ represents the probability of a transition back into the fecundable state of an infecundable female of age x whose latest conception was at age x-u. In Chiang's version of the model, such a transition cannot take place at a duration less than some lower bound a, and a female may not stay infecundable for more than a maximum period of time b. In terms of the forces of transition, the first of these conditions means that $\rho(y+u,u) = 0$ for $0 \leqq u \leqq a$, for each fixed age y at conception. The second condition implies that

$$(2.1) \qquad \rho(y+u,u) \to \infty \quad \text{as } u \uparrow b.$$

The latter observation has some interest in principle, so let us give it a little consideration.

Assume that a woman conceives at age y, and let T be the time she will subsequently spend in the infecundable state before she dies or becomes fecundable again. Assuming that the distribution F of T is absolutely continuous and that $\rho(y+u,u)$ and $\eta(y+u,u)$ are, say, continuous in u, we will have

$$F(t) = P\{T \leq t\} = 1 - \exp\{-\int_0^t [\rho(y+u,u) + \eta(y+u,u)]du\}.$$

Chiang assumes that

$$P\{a \leq T \leq b\} = 1, \tag{2.2}$$

so $F(b) = 1$, and thus

$$\int_0^b [\rho(y+u,u) + \eta(y+u,u)]\,du = \infty .$$

Since $\eta(x,u)$ will be bounded by a constant for all ages and durations of interest, this implies (2.1).

The central assumption leading to this result is the absolute continuity of F. One could easily achieve (2.2) in other ways, e.g. by placing some positive probability mass in b. Such a feature does not seem to have a reasonable interpretation in medical terms, however, and, anyway, it is not included in Chiang's model.

Since $F(b) = 1$, permanent sterility following a pregnancy is ruled out here.

2.C. The qualifying period in disability insurance. When in Subsection 2.A the force of recovery $\rho(x,u)$ was taken to depend upon duration u as disabled as well as upon age x attained, the intention was that this would reflect the real-life phenomenon that equally old people who are ill, will have a propensity to recover which depends on the duration of the illness. Similarly for the fertility model of Subsection 2.B: the propensity to become fecundable again really depends upon time since conception. In other cases, duration-dependence can get introduced through the observational plan, even though the original model is taken to be genuinely Markovian. This can happen when less than all relevant information about the sample paths is collected. We shall give an example based upon the simple three-state disability model.

Following Du Pasquier's original formulation of the disability model (Du Pasquier, 1912/13), let us ignore the dependence of ρ and η on u, and let us take them to be functions of x only, so that we are faced with an inhomogeneous Markov chain (with a continuous time parameter). In practice, disability income benefits will not be paid from disablement in most cases, but only after disability has lasted for some time κ, called the qualifying period or waiting period. In principle, the insurer will receive no information about cases of disability lasting less than κ. Indeed, part of the justification for having a qualifying period is precisely that the insurer avoids having to keep track of short-time disability. This means, however, that he does not observe the actual, underlying disability process, but a secondary one, derived from the primary process by means of the "filter" created by the qualifying period. It is easily seen that the secondary process is not a (three-state) Markov chain (Hoem, 1969a, Theorem 3.1) although the primary process is. The secondary process is semi-Markovian; no additional information about the behaviour of an observed sample path before an arbitrary moment t will change our probability statements about its behaviour after time t, once the state and duration at time t is known.

The observed process will have the original $\rho(x)$ as a force of recovery at age x and the original $\eta(x)$ as a force of mortality for an x-year-old disabled life. Its forces of transition out of the "active" state will not equal $\nu(x)$ and $\mu(x)$, however. Since it will be impossible to observe a disablement during the qualifying period, its force of disablement, say $\overline{\nu}(x,u)$, will equal 0 for durations $u < \kappa$. At duration $u = \kappa$, this function will make a jump to some positive value. An expression for this quantity is given elsewhere (Hoem, 1969a, page 111), but it need not concern us here. What is important to us is that $\overline{\nu}(x,\cdot)$ is discontinuous at κ.

Evidently, both $\overline{\nu}(x,u)$ and the corresponding force of mortality, say $\overline{\mu}(x,u)$, will be functions of all four original forces, ν, μ, ρ, and η.

2.D. Marriage models. It is well known that the inclination of existing marriages to dissolve varies with the duration of the marriage. This is even reflected in a popular expression like "the seven year itch". Similarly the inclination to remarry depends on time since marriage dissolution. A class of marriage models incorporating this feature has been reviewed elsewhere (Hoem, 1970b). (Additional references are Rowntree and Carrier (1958, pp. 214-218), U.S. Bureau of the Census (1970), Land (1971).) In these models, one operates with a number of marital states called "unmarried", "in first marriage", and so on, and again one has forces of transition which depend on age attained and duration, separately. Most of these models account for one sex only, but a model for marriage dissolution can easily take both sexes into consideration simultaneously. (It is easy to describe the termination of a unit, viz. the marriage, but no one has been able so far to give a satisfactory mathematical model for how two arbitrary units in a population, a potential bride and a potential bridegroom, can join forces and make a couple.) If one subsequently translates the two-sex model into a single-sex one, to match up with a one-sex model for remarriage, say, one may get duration-dependence even when there was none in the original two-sex model. Just as in Subsection 2.C, the reason is that some relevant information is not collected. In the present case information on one of the spouses is left out.

Let us describe how this effect may arise. To simplify exposition, let us disregard remarriage and emigration. (It is easy to include these features, but only at some notational inconvenience.) Consider a couple whose marriage has lasted for u years, and where the bride and bridegroom were y and x years old at marriage, respectively. Dissolution of their marriage can take place by divorce or through the death of one of the spouses. We describe the dissolution process by postulating a standard multiple decrement model with three causes of decrement, viz. (i) death of wife, (ii) death of husband, and (iii) divorce. The time parameter is marital duration, and the three forces of attrition are (i) the force $\eta(y+u)$ of mortality for (married) females, (ii) the corresponding force $\mu(x+u)$ for (married) males, and (iii) a force of divorce $\sigma(x+u, y+u, u)$. While the forces of mortality are taken as functions of age attained only, we shall let the force of divorce possibly depend separately on marital duration also.

To derive a one-sex model for females, say, we will take the age X of the bridegroom of a bride who marries at a known age y, to be a random variable, and we postulate a distribution $G_y(z) = P\{X-y \leqq z\}$ for the age difference $Z = X-y$. We can then describe attrition from the marital state by a standard multiple decrement model with death, divorce, and widowhood as decrements and marital duration as the time variable. The forces of attrition are

(a) the force of mortality: $\eta(y+u)$,

(b) the force of widowhood: $\omega(y+u,u) = \int_{\Omega} \mu(z+y+u)\, d\,G_y(z)$,

and

(c) the force of divorce: $\delta(y+u,u) = \int_{\Omega} \sigma(z+y+u, y+u, u)\, d\,G_y(z)$,

where Ω is the range of the age difference Z. Since σ depends on u, separately, it is not surprising that δ should also turn out to be duration-dependent. Since male mortality is a function of age attained only, it may be less immediately obvious that ω should have the same property as δ has. The reason for the duration-dependence of ω is, of course, the empirical fact that the distribution $G_y(\cdot)$ of the age difference genuinely changes with the age y of the bride.

2.E. Parity and birth intervals. If the incorporation of duration-dependence is important or desirable in many situations, it is quite essential in models intended for a closer study of human reproduction. Nature has set a definite lower bound for the interval between two live births to the same woman, so her propensity to have another child depends in a definite way on duration since the previous one. Marriage and other social arrangements similarly interfere and produce duration-dependence in fertility behaviour.

To indicate how human reproduction processes can be described by a semi-Markovian model, let us disregard widowhood and divorce, extra-marital births, and emigration, for simplicity. While an unmarried female will be said to be in state -1, a married woman will be designated a state equal to her parity. Any female thus starts in state -1 at birth, moves to state 0 at marriage, moves on to state 1 at her first live birth, and so on. There will be a force $\nu(x)$ of nuptiality, which is a function of age attained. There is a force of first births, say $\phi_0(x,u)$, depending on age x and on marital duration u. To births of order $n \geqq 2$, there corresponds a force ϕ_{n-1} which depends on age x, time w since the last previous birth ("the open interval"), and possibly also on marital duration u. (One will usually have to confine oneself to at most two out of the three arguments x, u, and w.) There is also a state of death with corresponding forces of mortality.

A recent paper by Sheps and Menken (1972) is based on essentially this type of model.

2.F. Human reproduction. Models of the sort sketched in the previous Subsection are essentially simple extensions of the life table model. They are geared to the type of data which can be collected in good official register systems, and do not go into the details of the human reproduction process, with conceptions, pregnancy outcomes, post-partum infecundable periods, the effect of contraception and abortion, and the like. The latter have been studied in the considerable literature which now exists on probabilistic models for human reproductive histories. Although there had been some previous work, notably due to Henry, it would be fair to say that the breakthrough along this line came with the 1964 papers by Perrin and Sheps (Perrin and Sheps, 1964; Sheps and Perrin, 1964). Later developments have been reviewed by Sheps and her associates (Sheps, Menken, and Radick, 1969; Sheps, 1971). An additional reference is Tolba (1966). Further papers are being published at a steady rate.

Most of this work has been based on semi-Markov models which are homogeneous in time, but some papers incorporating age-dependence (i.e., time-inhomogeneity) have also appeared (Sheps et al, 1969, Section 7.3; Potter, 1971; Venkatacharya, 1971). What makes these models semi-Markovian rather than, say, straight Markovian, is the feature that the probabilities of the various outcomes at the termination of a pregnancy (whether foetal loss, stillbirth, or live birth) depend on its duration. Usually, these models are not phrased in terms of forces of transition, but if such a framework were to be used, one would specify a force of foetal loss, say $\lambda_1(x,u)$, a force of stillbirth, say $\lambda_2(x,u)$, and a force of live birth, say $\lambda_3(x,u)$. All these functions would depend upon pregnancy duration u, and, possibly, upon age x attained. Their sum, $\lambda(x,u) = \Sigma\ \lambda_i(x,u)$, would be a force of pregnancy termination. The ratio $\pi_1(x,u) = \lambda_1(x,u)/\lambda(x,u)$ would represent the probability of a foetal loss in a pregnancy terminating at age x when conception was at age x-u. The ratios λ_2/λ and λ_3/λ have similar interpretations.

2.G. Internal migration. To improve upon the realism of ordinary Markov chain models for social mobility, McGinnis (1968, p. 716) has proposed "the following simple axiom about motion through time in social space:

Axiom of cumulative inertia: The probability of remaining in any state of nature increases as a strict monotone function of duration of prior residence in that state."

In two perceptive recent papers, Ginsberg (1971a,b) has shown how this idea can be used as a basis for semi-Markovian models for social mobility in general, and for internal migration in particular. If we disregard external migration for the moment, the states of the migration model represent the regions into which a country is divided, say, (plus a state of death), and one could work with forces $\mu_{ij}(x,u)$ of migration from region i to region j for x-year-olds whose last previous move was made at age x-u. One would get cumulative inertia by specifying the $\mu_{ij}(x,\cdot)$ to be (strictly) decreasing, but this is not a requirement in the model.

Ginsberg (1971b, pp. 8-9) does not rule out i to i moves; i.e., he keeps open the possibility of analysing moves from one location to another in the same region. Following such a move, the value of the duration variable would be set back to 0. Thus duration is the length of stay at the location, not the total sojourn without a break in the region.

The results of an analysis based on this model would probably have substantive (say, sociological) interest only if the number of regions were not too small. Perhaps a score or two is needed in a country like Norway, whose data Ginsberg plans to use. This would mean that the analysis of a process which is inhomogeneous in time due to age effects, will surely be difficult because of the massive amounts of data required to estimate parameters and carry out tests. It is always easier to stay within a time-homogeneous set-up. To overcome this very real difficulty, Ginsberg (1971, pp. 257-259) suggests using the device of an operational time to translate his original model into a homogeneous process. He feels that this has a good chance of working "if the interactions between age and location [region] are not too great". In a personal communication, he tells me that for the value of the operational time function at age x (for any x) he considers using the expected number of moves made by an individual within that age.

We shall look more closely at the concept of an operational time in Section 5 below.

3. ASPECTS OF THE THEORY OF HOMOGENEOUS SEMI-MARKOV PROCESSES

3.A. The purpose of the present and the next Section is to provide some basic mathematical machinery for the study of semi-Markovian models of a form which is useful in demographic and actuarial applications. To provide a link with the now classical homogeneous model, we shall discuss briefly some aspects of the latter, largely using notation introduced by Pyke and Schaufele (1964). The stringent mathematics have been given in that paper and others (see Çinlar, 1969), and we shall stay on an intuitive level.

3.B. There is, then, a finite or countable collection $\mathcal{J}$ of states i, representing, e.g., the various demographic statuses in the above examples, usually including the status "dead". With this interpretation, transitions between states would correspond to changes in demographic status (including, possibly, death). A sample path corresponds to the history of an individual (or sometimes an individual couple). These histories are taken as independent.

The state of a sample path at time t is $Z(t)$, time usually representing age attained. The transition into $Z(t)$ occurred at time $t-U(t)$, so that $U(t)$ is the duration of the current stay in $Z(t)$. The next transition occurs at time $t+V(t)$. It results in a jump from state $Z(t)$ to state $Z^+(t) = Z(t+V(t))$. The duration variable $U(\cdot)$ is then set back to zero, i.e., $U(t+V(t)) = 0$.

The stochastic process $\{(Z(t), V(t)); t \geqq 0\}$ is a homogeneous Markov process over the state space $\mathcal{J} \times [0,\infty\rangle$, and its transition function is

$$P_{ij}(t-s,u,v) = P\{Z(t) = j, U(t) \leqq v \mid Z(s) = i, U(s) = u\}.$$

Much of its behaviour is reflected in the function

$$Q_{ij}(u,v) = P\{ Z^{+}(t) = j,\ V(t) \leqq v \mid Z(t) = i,\ U(t) = u \},$$

which essentially describes what the sample path may be expected to do next when we know that it is now in state i with a currend duration of u. The distribution of the future sojourn V(t) in Z(t) is given as

$$H_{i}(u,v) = \sum_{j} Q_{ij}(u,v) = P\{ V(t) \leqq v \mid Z(t) = i,\ U(t) = u \}.$$

Two further quantities require our attention, viz.

$$p_{ij} = Q_{ij}(0,\infty) = P\{ Z^{+}(t) = j \mid Z(t) = i,\ U(t) = 0 \},$$

and, for the (i,j) where $p_{ij} > 0$ so that a direct transition $i \to j$ is possible,

$$F_{ij}(v) = Q_{ij}(0,v) / p_{ij} = P\{ V(t) \leqq v \mid Z(t) = i,\ U(t) = 0,\ Z^{+}(t) = j \}.$$

3.C. Frequently, it is convenient to base mathematical arguments on the distribution $F_{ij}(\cdot)$ of the total sojourn in state i, given that j is the next state to be visited, and many results are phrased in terms of this function. Some descriptions of its role makes one imagine a probabilistic mechanism where, upon entry into state i, the next state j to be visited is first determined according to the probability distribution $\{p_{ij};\ j \in \mathcal{J}\}$, and then the length of stay is determined subsequently according to $F_{ij}(\cdot)$.

The notion of such a sequencing in time is useful in many connections, but one must not get trapped, of course, by the wording of this interpretation into believing that it is an assumption in the mathematical theory. It may be equally fruitful to think of what goes on as happening in the reverse sequence, i.e., to take the length of stay v as determined first (according to $H_{i}(0,\cdot)$) and the next state as being fixed only subsequently.

A third presentation can be given in terms of a continuously operating mechanism for deciding at any duration u whether a transition will occur in the next instant and, if so, to what new state the jump will be made. The latter interpretation is the basis of the force of transition concept. In the present context, the force of transition from state i to state $j \neq i$ at duration u is defined as the limit

$$\mu_{ij}(u) = \lim_{h \downarrow 0} P_{ij}(h,u,\infty)/h,$$

which is assumed to exist for all i, $j \neq i$, $u \geq 0$. Evidently, $\mu_{ij}(u)\Delta u + o(\Delta u)$ is the probability of a transition to state j within time Δu, given a current stay in state i of duration u.

3.D. The $\mu_{ij}(\cdot)$ correspond to the q_{ij} of the Q-matrix of a Markov chain with a continuous time parameter (Chiang, 1960, p. 130). While the q_{ij} play a prominent part in the latter area, the standard literature on the general theory of semi-Markov processes has found little room for the $\mu_{ij}(\cdot)$. Yet forces of transition have been highly useful in applications in fields like demography, biostatistics, and actuarial science, and they could have played an even more prominent part there, as we have indicated in Section 2 above.

These functions will provide us with a connecting link between the models which are homogeneous in time and the corresponding inhomogeneous processes, to which we now turn.

4. RUDIMENTS OF A THEORY OF INHOMOGENEOUS SEMI-MARKOV PROCESSES: I. BASIC NOTIONS

4.A. Extending the notation of the previous Section, we now let

$$P_{ij}(s,t,u,v) = P\{Z(t) = j,\ U(t) \leq v \mid Z(s) = i,\ U(s) = u\},$$

with $P_{ij}(s,s,u,v) \equiv \delta_{ij}$ (a Kronecker delta);

$$Q_{ij}(t,u,v) = P\{Z^+(t) = j,\ V(t) \leq v \mid Z(t) = i,\ U(t) = u\};$$

$$H_i(t,u,v) = \sum_j Q_{ij}(t,u,v) = P\{V(t) \leq v \mid Z(t) = i,\ U(t) = u\};$$

and, for $i \neq j$,

$$\mu_{ij}(s,u) = \lim_{t \downarrow s} P_{ij}(s,t,u,\infty)/(t-s) = \frac{\partial}{\partial t} P_{ij}(s,t,u,\infty)\Big|_{t=s}; \tag{4.1}$$

with obvious verbal interpretations. We assume that the $\mu_{ij}(s,u)$ exist for all $i \neq j$, $s \geq 0$, $u \geq 0$, and take each $\mu_{ij}(\cdot,\cdot)$ to be continuous. (Extension to discontinuous forces, as needed in Subsection 2.C, is made <u>ad hoc</u>.)

We also assume that $\sum_j P_{ij}(s,t,u,\infty) \equiv 1$.

4.B. The μ_{ij} are of central interest, particularly when each $Q_{ij}(t,u,\cdot)$ and $H_i(t,u,\cdot)$ is absolutely continuous, as we shall assume. Then positive probability mass at any given duration is ruled out, something which constitutes a definite restriction of generality, but not beyond what seems of prime interest for the applications which we have in mind.

In the general time-homogeneous model, the $Q_{ij}(u,\cdot)$ are any (measurable non-negative right-continuous) functions such that $H_i(u,\cdot)$ is a distribution function, possibly defective, concentrated on $[0,\infty\rangle$. This flexibility in the choice of the Q_{ij} is essential for the great generality of Markov renewal theory. One gets interesting and useful sub-theories by introducing some special restriction, and, indeed, important theories, like that of Markov chains, can be seen as arising in this way. Here, we try to pursue this line of thought in another direction through restricting ourselves to absolutely continuous Q_{ij} and H_i. In particular, $H_i(t,u,0) \equiv 0$, so that all sojourns must take some time.

4.C. There is another restriction which we shall make, but which is one of formal appearance only. To simplify notation, we shall do away with direct transitions $i \to i$, i.e., we shall not permit $Z^+(t)$ to equal $Z(t)$ with positive probability. In its general formulation, the time-homogeneous model allows for such transitions, and our comments on Ginsberg's migration model (Subsection 2.G) show that this property may be useful in demography. Nevertheless, we shall let $Q_{ii} \equiv 0$, but this is no real restriction of generality. If direct transitions $i \to i$ constitute an important element in an application, one may easily "save" this feature and still have $Q_{ii} \equiv 0$ by using two copies of the state space $\mathcal{J}$, say $\mathcal{J}'$ and $\mathcal{J}''$. A direct transition from a state $i \in \mathcal{J}$ to itself would then be represented as a jump between the distinct states $i' \in \mathcal{J}'$ and $i'' \in \mathcal{J}''$.

4.D. We also introduce the (total) force of decrement from state i,

$$\mu_i(s,u) = \lim_{t \downarrow s} \{1 - P_{ii}(s,t,u,\infty)\}/(t-s) = -\frac{\partial}{\partial t} P_{ii}(s,t,u,\infty)\Big|_{t=s},$$

and assume that

$$(4.2) \qquad \mu_i(s,u) = \sum_{j \neq i} \mu_{ij}(s,u).$$

We will take each $\mu_i(\cdot,\cdot)$ to be continuous.

4.E. Let $N(s,t)$ be the number of transitions observed in any period $\langle s,t]$. The possibility that $N(s,t) = \infty$ seems to have no interesting interpretation in the applications we have in mind, and we shall assume $N(s,t) < \infty$ with probability 1. (Nobody experiences an infinity of demographic events.) We shall assume also that

$$\lim_{\Delta s \downarrow 0} P\{N(s,s+\Delta s) > 1 \mid Z(s) = i, \; U(s) = u\} / \Delta s = 0,$$

for all i, s, u. Intuitively, this means that the transitions (events) occur in an orderly fashion, one at a time.

We <u>will</u> get $P\{N(s,t) < \infty \mid Z(s) = i, \; U(s) = u\} = 1$ if $\mu_j(s', u')$ $\leq$ some constant c for all j, all $s' \in \langle s,t \rangle$, and all $u' \leq u+t-s$, because N(s,t) will then be stochastically smaller than some Poisson distributed variable with parameter $c \cdot (t-s)$. In practice, the μ_j are frequently uniformly bounded in this way, so we have a simple criterion for the finiteness of N(s,t). Chiang's example (Subsection 2.B) shows, however, that unbounded μ_j-s do have some interest. ($N < \infty$ is secured there, of course, by the finiteness of $\mathcal{J}$.)

4.F. Let

$$P_{ij}^{(n)}(s,t,u,v) = P\{ Z(t) = j, \; U(t) \leq v, \; N(s,t) = n \mid Z(s) = i, \; U(s) = u\}.$$

Then

$$P_{ij}^{(0)}(s,t,u,v) = \delta_{ij} \, \varepsilon(v+s-u-t) \, \overline{H}_i(s,u,t-s), \tag{4.3}$$

where $\varepsilon(x) = 0$ or 1 as $x < 0$ or $x \geq 0$, while $\overline{H}_i(s,u,t-s) = 1 - H_i(s,u,t-s)$ is the probability that the sample path will stay put in i at least until time t. A decomposition with respect to the possible values of $\{Z^+(s), V(s)\}$ gives, for $n \geq 1$,

$$P_{ij}^{(n)}(s,t,u,v) = \sum_{k \neq i} \int_s^t \overline{H}_i(s,u,w-s) \mu_{ik}(w,u+w-s) P_{kj}^{(n-1)}(w,t,0,v) \, dw. \tag{4.4}$$

Adding over n, we get

$$\begin{aligned} P_{ij}(s,t,u,v) &= \sum_{k \neq i} \int_s^t \overline{H}_i(s,u,w-s) \mu_{ik}(w,u+w-s) P_{kj}(w,t,0,v) \, dw \\ &\quad + \delta_{ij} \, \varepsilon(v+s-u-t) \, \overline{H}_i(s,u,t-s). \end{aligned} \tag{4.5}$$

4.G. Another set of formulas follow from a similar decomposition according to the values of U(t) and the state preceding Z(t). Before we list these relations, it is useful to introduce a <u>force of increment</u> (Hoem, 1969b, § 7.F), defined as follows:

$$\lambda_{ij}(s,t,u) = \lim_{\Delta t \downarrow 0} P\{Z(t) \neq j,\ Z(t+\Delta t)=j \mid Z(s)=i,\ U(s)=u\} / \Delta t \tag{4.6}$$

$$= \int_0^\infty \sum_{k \neq j} P_{ik}(s,t,u,dw)\,\mu_{kj}(t,w).$$

The quantity $\lambda_{ij}(s,t,u)\Delta t + o(\Delta t)$ is the probability that a sample path with $Z(s) = i$, $U(s) = u$, will not be in state k at time t, but will make a jump to this state before time $t+\Delta t$. (For example, it may be the probability that an s-year-old married female with marital duration u will be married at a later age t too, but will then get a divorce before age $t+\Delta t$.)

We get

$$P_{ij}^{(n)}(s,t,u,v) = \sum_{k \neq j} \int_{\tau=t-v}^{t} \int_{\omega=0}^{\infty} P_{ik}^{(n-1)}(s,\tau,u,dw)\,\mu_{kj}(\tau,w)\,\overline{H}_j(\tau,0,t-\tau)\,d\tau$$

for $n \geqq 1$, $v \leqq t-s$; and, using the force of increment function,

$$P_{ij}(s,t,u,v) = \tag{4.7}$$

$$\int_{\max(s,t-v)}^{t} \lambda_{ij}(s,\tau,u)\overline{H}_j(\tau,0,t-\tau)\,d\tau + \delta_{ij}\,\varepsilon(v+s-u-t)\,\overline{H}_i(s,u,t-s).$$

4.H. Some further formulas are

$$\overline{H}_i(t,u,v) = \exp\{-\int_0^v \mu_i(t+y,u+y)\,dy\} \quad \text{for } v \geqq 0, \tag{4.8}$$

$$Q_{ij}(t,u,v) = \int_0^v \overline{H}_i(t,u,w)\,\mu_{ij}(t+w,u+w)\,dw,$$

$$\mu_{ij}(s,u) = \frac{\partial}{\partial v} Q_{ij}(s,u,0),$$

and

$$\lambda_{ij}(s,s,u) = \mu_{ij}(s,u). \tag{4.9}$$

The Chapman-Kolmogorov equations are

$$P_{ij}(s,t,u,v) = \int_0^\infty \sum_k P_{ik}(s,t',u,dw)\,P_{kj}(t',t,w,v), \tag{4.10}$$

for $0 \leqq s \leqq t' \leqq t$.

4.I. Both demographers and actuaries will take interest in the mean value of particular functionals of the stochastic process. Thus, demographers will want to know such things as the mean number of births to a woman, the mean age at marriage, the mean duration of a marriage, and so on. Such expected values can be built up from two kinds of elements: the mean total sojourn in given states, and the mean number of transitions of various kinds. We give three examples. Given that $Z(s) = i$, $U(s) = u$, the mean total subsequent sojourn in state j is

$$\int_s^\infty P_{ij}(s,t,u,\infty)\, dt,$$

the mean number of subsequent jumps $j \to k$ is

$$\int_{t=s}^{\infty} \int_{v=0}^{\infty} P_{ij}(s,t,u,dv)\, \mu_{jk}(t,v),$$

and the mean number of subsequent arrivals in state j equals

$$\int_s^\infty \lambda_{ij}(s,t,u)\, dt.$$

Actuaries will be more interested in mean cash values (called "actuarial values") of various streams of money. Two examples follow. The actuarial value at time s of an income benefit to an insured life for which $Z(s) = i$, $U(s) = u$, in the amount of $B_j(t,v)$ at time $t \geqq s$ if $Z(t) = j$, $U(t) = v$, equals

$$\int_{t=s}^{\infty} \int_{v=0}^{\infty} e^{-\delta(t-s)} \sum_j B_j(t,v)\, P_{ij}(s,t,u,dv),$$

where δ is the force of interest, assumed constant. Similarly, the actuarial value at time s of a benefit paid upon arrival in state j, in the amount of $C_j(t)$ if this occurs at time t, equals

$$\int_s^\infty e^{-\delta(t-s)} C_j(t)\, \lambda_{ij}(s,t,u)\, dt.$$

One could also calculate other characteristics of these functionals, such as their standard deviations (Hoem, 1969b, Section 7), but users display little interest in them.

4.J. In demographic and actuarial applications, one will typically be concerned with a restricted age interval, not extending beyond the maximum lifetime of an individual. In fertility studies, for instance, one's interest will seldom go further than age 50 for females. Chiang's fertility model (Subsection 2.B) shows that there may also be some upper bound to possible duration. It is easy to incorporate such bounds by specifying an interval $[0, \zeta\rangle$ to which s and t must belong, and a similar interval for u and v. To save some writing, we suppress this feature.

5. RUDIMENTS II. OPERATIONAL TIME

5.A. A change of time scale is used for many purposes in the theory of Markov processes. As we mentioned in Subsection 2.G, Ginsberg proposes to use it to transform a time-inhomogeneous semi-Markov process into a homogeneous one. The first one who did something similar, seems to have been Filip Lundberg (1903), who showed that a Poisson process with a time-dependent intensity $\lambda(t)$ is changed into a Poisson process with a constant intensity of 1 when one switches to the time scale

$$(5.1) \qquad \Lambda(t) = \int_0^t \lambda(s)ds.$$

Ove Lundberg (1940, p. 57) later used the same device to "homogenize" a pure birth process with time-dependent birth intensities of the form

$$(5.2) \qquad \lambda_n(t) = c_n \lambda(t),$$

and Bühlmann (1970, pp. 50-51) has recently shown that an inhomogeneous pure birth process can be transformed into a homogeneous one by a change of time scale only if (5.2) holds for some function $\lambda(\cdot)$ and some set of constants $\{c_n\}$. We shall now first extend these ideas to semi-Markovian processes. To make full use of our results, we shall subsequently specialize to general Markov chains with a continuous time parameter.

5.B. For any non-decreasing, right-continuous real function $\Lambda(\cdot)$, defined over $[0, \infty\rangle$ and with $\Lambda(0) = 0$, let

$$\Lambda^{-1}(s) = \inf\{t: \Lambda(t) \geq s\},$$

as usual. The change of time-scale which we shall study, consists in replacing real time t by a new time $\Lambda(t)$, in the sense that we replace any real function $f(t)$ by $\tilde{f}(t) = f(\Lambda^{-1}(t))$. Thus $Z(t)$ and $U(t)$ are replaced by $\tilde{Z}(t) = Z(\Lambda^{-1}(t))$ and $\tilde{U}(t) = U(\Lambda^{-1}(t))$, respectively, and for $N(s,t)$ we substitute $\tilde{N}(s,t) = N(\Lambda^{-1}(s), \Lambda^{-1}(t))$. The transformed process will have transition probabilities of the form

$$(5.3) \quad \tilde{P}_{ij}^{(n)}(s,t,u,v) = P_{ij}^{(n)}\{\Lambda^{-1}(s), \Lambda^{-1}(t), \Lambda^{-1}(s) - \Lambda^{-1}(s-u), \Lambda^{-1}(t) - \Lambda^{-1}(t-v)\},$$

$$(5.4) \quad \tilde{Q}_{ij}(t,u,v) = Q_{ij}\{\Lambda^{-1}(t), \Lambda^{-1}(t) - \Lambda^{-1}(t-u), \Lambda^{-1}(t+v) - \Lambda^{-1}(t)\},$$

$$(5.5) \quad \tilde{P}_{ij}(s,t,u,\infty) = P_{ij}\{\Lambda^{-1}(s), \Lambda^{-1}(t), \Lambda^{-1}(s) - \Lambda^{-1}(s-u), \infty\},$$

and so on. The process $(\tilde{Z}, \tilde{U}) = \{(\tilde{Z}(t), \tilde{U}(t));\ t \geq 0\}$ is, of course, time-homogeneous if (and only if) $\tilde{P}_{ij}(s,t,u,v)$ depends on s and t only via their difference $t-s$.

5.C. If $\Lambda(\cdot)$ is continuous and strictly increasing, formulas (5.3) to (5.5) have the converses

$$(5.6) \quad P_{ij}^{(n)}(s,t,u,v) = \tilde{P}_{ij}^{(n)}\{\Lambda(s), \Lambda(t), \Lambda(s) - \Lambda(s-u), \Lambda(t) - \Lambda(t-v)\},$$

$$(5.7) \quad Q_{ij}(t,u,v) = \tilde{Q}_{ij}\{\Lambda(t), \Lambda(t) - \Lambda(t-u), \Lambda(t+v) - \Lambda(t)\},$$

and

$$(5,8) \quad P_{ij}(s,t,u,\infty) = \tilde{P}_{ij}\{\Lambda(s), \Lambda(t), \Lambda(s) - \Lambda(s-u), \infty\}.$$

If the forces $\tilde{\mu}_{ij}$ of transition corresponding to the $\tilde{P}_{ij}$ exist, and if (5.1) holds with a right-continuous $\lambda(\cdot)$, it follows from (5.8) that

$$\mu_{ij}(s,u) = \tilde{\mu}_{ij}\{\Lambda(s), \Lambda(s) - \Lambda(s-u)\}\lambda(s).$$

The process $(\tilde{Z}, \tilde{U})$ will be a time-homogeneous semi-Markov process if and only if each $\tilde{\mu}_{ij}(s,u)$ is independent of s. In this case,

$$(5.9) \quad \mu_{ij}(s,u) = \tilde{\mu}_{ij}\{\Lambda(s) - \Lambda(s-u)\}\lambda(s).$$

5.D. Let us agree to call any real function $\Lambda(\cdot)$, defined over $[0, \infty\rangle$, an operational time provided it satisfies (5.1) with a positive, right-continuous $\lambda(\cdot)$, and provided the corresponding process $(\tilde{Z}, \tilde{U})$ is a time-homogeneous semi-Markov process. The relation (5.1) of course makes $\Lambda(\cdot)$ absolutely continuous and strictly increasing, and it makes $\Lambda(0) = 0$. We prove the following theorem.

Theorem 5.10. There exists an operational time $\Lambda(\cdot)$ if and only if one can write the forces of transition in the form (5.9), where $\lambda(\cdot)$ is a right-continuous, positive real function, $\Lambda(\cdot)$ satisfies (5.1), and the $\tilde{\mu}_{ij}(\cdot)$ are continuous.

Proof. (i) Assume that (5.9) holds as specified, and let

$$\tilde{\mu}_i(x) = \sum_{j \neq i} \tilde{\mu}_{ij}(x).$$

Then

$$\mu_i(s,u) = \tilde{\mu}_i\{\Lambda(s) - \Lambda(s-u)\}\,\lambda(s) \tag{5.11}$$

by (5.9) and (4.2). The substitution of (5.11) in (4.8) and the introduction of $w = \Lambda(t+y) - \Lambda(t-u)$ as a new variable of integration gives

$$\bar{H}_i(t,u,v) = \exp\{-\int_{\Lambda(t)-\Lambda(t-u)}^{\Lambda(t+v)-\Lambda(t-u)} \tilde{\mu}_i(w)\,dw\},$$

and, consequently,

$$\begin{aligned}\bar{H}_i\{\Lambda^{-1}(s),\ \Lambda^{-1}(s) - \Lambda^{-1}(s-u),\ \Lambda^{-1}(t) - \Lambda^{-1}(s)\}\\ = \exp\{-\int_0^{t-s} \tilde{\mu}_i(u+y)\,dy\}.\end{aligned} \tag{5.12}$$

Thus, by (5.3) and (4.3),

$$\tilde{P}_{ij}^{(0)}(s,t,u,v) = \delta_{ij}\,\varepsilon[(s-u) - (t-v)]\,\exp\{-\int_0^{t-s} \tilde{\mu}_i(u+y)\,dy\},$$

since $\Lambda^{-1}(s-u) \geqq \Lambda^{-1}(t-v)$ if and only if $s-u \geqq t-v$. This function therefore depends on s and t only via $t-s$. Now make the induction assumption that this is the case for $\tilde{P}_{ij}^{(n)}(s,t,u,v)$ for all i, j, s, t, u, and v for all $n < m$, and write $\tilde{P}_{ij}^{(n)}(s,t,u,v) = \tilde{P}_{ij}^{(n)}(t-s,u,v)$ for such n. Relations (4.4) and (5.3) give

$$\tilde{P}_{ij}^{(m)}(s,t,u,v) = \sum_{k \neq i} \int_{\Lambda^{-1}(s)}^{\Lambda^{-1}(t)} \bar{H}_i\{\Lambda^{-1}(s),\ \Lambda^{-1}(s)-\Lambda^{-1}(s-u),\ w-\Lambda^{-1}(s)\}\,\lambda(w)\ \cdot$$

$$\cdot\ \tilde{\mu}_{ik}\{\Lambda(w)-s+u\}\, P_{kj}^{(m-1)}\{w,\ \Lambda^{-1}(t),\ 0,\ \Lambda^{-1}(t)-\Lambda^{-1}(t-v)\}\,dw\,.$$

The substitution of $z = \Lambda(w)-s+u$ gives, by (5.12) and (5.3),

$$\tilde{P}_{ij}^{(m)}(s,t,u,v) = \sum_{k \neq i} \int_{u}^{u+t-s} \exp\{-\int_u^z \tilde{\mu}_i(y)dy\}\,\tilde{\mu}_{ik}(z)\,\tilde{P}_{kj}^{(m-1)}(u+t-s-z,\ 0,\ v)\,dz\,,$$

so that this function too depends on s and t only via $t-s$. By induction, all $\tilde{P}_{ij}^{(n)}(s,t,u,v)$ then have this property, and so does their sum $\tilde{P}_{ij}(s,t,u,v)$, as was to be proved.

(ii) Now assume conversely that an operational time $\Lambda(\cdot)$ exists. Its inverse $\Lambda^{-1}(\cdot)$ is right-differentiable with

$$\frac{d^+}{dt}\Lambda^{-1}(t) = 1/\lambda\{\Lambda^{-1}(t)\},$$

and, for $i \neq j$,

$$\frac{\tilde{P}_{ij}(s,t,u,\infty)}{t-s} = \frac{P_{ij}\{\Lambda^{-1}(s), \Lambda^{-1}(t), \Lambda^{-1}(s)-\Lambda^{-1}(s-u), \infty\}}{\Lambda^{-1}(t)-\Lambda^{-1}(s)} \cdot \frac{\Lambda^{-1}(t)-\Lambda^{-1}(s)}{t-s}\,.$$

As $t \downarrow s$, the right hand side here converges to

$$\mu_{ij}\{\Lambda^{-1}(s),\ \Lambda^{-1}(s)-\Lambda^{-1}(s-u)\}\,\frac{d^+}{ds}\Lambda^{-1}(s)\ ,$$

which means that the forces of transition corresponding to the $\tilde{P}_{ij}$ exist. It then follows from the argument below (5.8) that (5.9) holds, as was to be proved. □

5.E. If $\Lambda(\cdot)$ is an operational time, then evidently so is $\Lambda_a(\cdot) = a\Lambda(\cdot)$ for any constant $a > 0$. This transformation only corresponds to a change of time unit, however, and we will regard all members of the class $\{\Lambda_a(\cdot);\ a > 0\}$ as essentially the same operational time.

The question then arises whether there may exist two or more essentially different operational times. It turns out that we may prove the following uniqueness theorem.

Theorem 5.13. If $\Lambda(\cdot)$ is an operational time and there exists a pair (i,j), with $i \neq j$, for which $\tilde{\mu}_{ij}(0) > 0$, then $\Lambda(\cdot)$ is unique (up to a multiplicative positive constant).

Proof. Assume that $\Gamma(t) = \int_0^t \gamma(s)ds$ is another operational time, and let us designate the forces of the corresponding homogeneous process by $\hat{\mu}_{ij}$. Let $\tilde{\mu}_{k\ell}(0) > 0$. Since

$$\mu_{k\ell}(s,u) = \tilde{\mu}_{k\ell}\{\Lambda(s) - \Lambda(s-u)\}\,\lambda(s)$$

$$= \hat{\mu}_{k\ell}\{\Gamma(s) - \Gamma(s-u)\}\,\gamma(s),$$

we get $\lambda(s) = \gamma(s)\,\hat{\mu}_{k\ell}(0)/\tilde{\mu}_{k\ell}(0)$ by letting $u = 0$. The result of the theorem is then immediate. □

5.F. The results of the previous parts of this Section will be specialized to Markov chains if we assume that the $\mu_{ij}(s,u)$ are independent of u. An operational time for a time-inhomogeneous Markov chain with a continuous time parameter is, of course, a continuous real function $\Lambda(\cdot)$, defined over $[0, \infty\rangle$, which satisfies (5.1) with a positive, right-continuous $\lambda(\cdot)$, and which transforms the chain into a homogeneous process. Disregarding homogeneous chains where all entries in the Q-matrix are zero, we get the following theorem.

Theorem 5.14. There exists an operational time $\Lambda(\cdot)$ for a Markov chain if and only if one can write its forces of transition in the form

$$\mu_{ij}(s) = q_{ij}\,\lambda(s),$$

where $\lambda(\cdot)$ is a right-continuous, positive real function satisfying (5.1). Such an operational time is unique (up to a multiplicative positive constant).

We have, of course, restricted ourselves to chains satisfying the assumptions which follow from Subsections 4.A, D, and E.

In his study of the operational time of a pure birth process with variable intensities, Bühlmann (1970, Section 2.2.3) proves the analogue of Theorem 5.14, except the part about uniqueness, with slightly less restrictive conditions in the definition of the operational time. A perusal of his proof reveals that uniqueness is easily proved in his situation too.

5.G. In the case of a Poisson process with a variable intensity $\lambda(t)$, $\Lambda(t)$ represents the expected number of occurrences of the event under observation (the expected number of renewals) during $[0, t]$. Thus a nice and simple interpretation of the operational time exists for this situation. It would be a great help if some similar interpretation could be found for the more general models considered in this paper. This would give a welcome guide in looking for an operational time in a concrete application, and also in checking whether a useful change of time scale is possible. Unfortunately, such an interpretation seems to be lacking.

6. RUDIMENTS III. PURGED AND PARTIAL PROCESSES. THE IMPORTANCE OF THE OBSERVATIONAL PLAN

6.A. People will frequently be interested in studying a hypothetical situation where one or more of the forces in operation are eliminated. For example, one will often want to know what number of births a female should expect if there were no mortality, or the probability that a newly-wed 25-year-old woman would become a widow if divorce were impossible (the alternative being that her husband becomes a widower). Such questions lead to the study of partial models, which arise from the original models by the substitution of 0 for the forces which are to be eliminated.

We have studied this approach as applied to Markov chains in a previous paper (Hoem, 1969c). The extension to semi-Markovian models is quite straightforward, and we shall only sketch some of the main ideas.

6.B. Assume, then, that the state space $\mathcal{J}$ can be partitioned into two disjoint parts, $\mathcal{K}$ and $\mathcal{R}$, where $\mathcal{R}$ is absorbing, i.e., $\mathcal{K}$ cannot be reached from $\mathcal{R}$. Suppose that one wants to eliminate the possibility that a sample path makes a jump into $\mathcal{R}$. One achieves this by substituting 0 for all $\mu_{ij}(\cdot,\cdot)$ where $i \in \mathcal{K}$, $j \in \mathcal{R}$. The effect is that of <u>r</u>emoving the states in $\mathcal{R}$ and <u>k</u>eeping those in $\mathcal{K}$.

Now assume that a set of transition probabilities $P_{ij}(s,t,u,v; \mathcal{K})$ of a semi-Markov process over the state space $\mathcal{K}$ can be constructed uniquely from the forces $\{\mu_{ij}(\cdot,\cdot);\ i \in \mathcal{K},\ j \in \mathcal{K}\}$. In actuarial and demographic applications, this is uniformly the case. We will then call these functions <u>the partial probabilities</u> corresponding to $\mathcal{K}$, and the semi-Markovian model for which they are transition functions, will be said to be <u>partial</u> relative to the original one.

6.C. The procedure which produces the partial model corresponding to $\mathcal{K}$, is not equivalent to conditioning upon non-absorption in $\mathcal{R}$, except in special cirsumstances. Thus, for example, the (partial) probability of becoming a widow when divorce is impossible, is something else generally than the conditional probability that a woman becomes a widow, given that she will have no divorce.

For the fundamental conditional probabilities, we pick some moment τ, which may represent age at menopause, say, or the terminal age ω of the life table, or age attained on a given census day, or some other moment or age of interest. We then condition upon the event $Z(\tau) \in \mathcal{K}$. The corresponding transition probabilities are

$$_\tau P_{ij}(s,t,u,v) = P\{Z(t)=j,\ U(t)\leqq v \mid Z(s)=i,\ U(s)=u,\ Z(\tau)\in \mathcal{K}\}$$

for $0 \leqq s \leqq t \leqq \tau$. (Using our marriage example once more and letting $\tau = \omega$, this may be the probability that a woman in her first marriage at age s, who has then been married for u years, and who will never get a divorce, will be married, or possibly remarried, at age t, with a marital duration then not exceeding v.) Evidently,

$$_\tau P_{ij}(s,t,u,v) = \frac{P\{Z(t)=j,\ U(t)\leqq v,\ Z(\tau)\in\mathcal{K} \mid Z(s)=i,\ U(s)=u\}}{P\{Z(\tau)\in\mathcal{K} \mid Z(s)=i,\ U(s)=u\}}.$$

The denominator here equals $P_{i\mathcal{K}}(s,\tau,u,\infty) = \Sigma_{j\in\mathcal{K}} P_{ij}(s,\tau,u,\infty)$. The numerator equals

$$\int_0^{\min(v,t-s)} P_{ij}(s,t,u,dw)\, P_{j\mathcal{K}}(t,\tau,w,\infty)$$

$$+\ \delta_{ij}\ \varepsilon(v+s-t-u)\ \overline{H}_i(s,u,t-s)\ P_{i\mathcal{K}}(t,\tau,u+t-s,\infty).$$

By (4.7) the integral equals

$$\int_0^{\min(v,t-s)} \lambda_{ij}(s,t-w,u)\ \overline{H}_j(t-w,0,w)\ P_{j\mathcal{K}}(t,\tau,w,\infty)\,dw$$

$$= \int_{\max(s,t-v)}^{t} \lambda_{ij}(s,x,u)\ \overline{H}_j(x,0,t-x)\ P_{j\mathcal{K}}(t,\tau,t-x,\infty)\,dx.$$

Thus,

$$_\tau P_{ij}(s,t,u,v) = \Big\{\int_{\max(s,t-v)}^{t} \lambda_{ij}(s,x,u)\ \overline{H}_j(x,0,t-x)\ P_{j\mathcal{K}}(t,\tau,t-x,\infty)\,dx$$

(6.1)

$$+\ \delta_{ij}\ \varepsilon(v+s-t-u)\ \overline{H}_i(s,u,t-s)\ P_{i\mathcal{K}}(t,\tau,u+t-s,\infty)\Big\} / P_{i\mathcal{K}}(s,\tau,u,\infty).$$

The ${}_{\tau}P_{ij}(s,t,u,v)$ will be a new set of transition probabilities for a semi-Markov process over the state space $\mathcal{K}$. The corresponding forces of transition are, for $i \neq j$,

$$ {}_{\tau}\mu_{ij}(s,u) = \lim_{t \downarrow s} {}_{\tau}P_{ij}(s,t,u,\infty)/(t-s) = \frac{\partial}{\partial t}\, {}_{\tau}P_{ij}(s,t,u,\infty)\Big|_{t=s}. $$

If $\lambda_{ij}(s,x,u)\frac{\partial}{\partial t}[\overline{H}_j(x,0,t-x)\, P_{j\mathcal{K}}(t,\tau,t-x,\infty)]$ is a continuous function of (x,t), as we shall simply assume, straightforward differentiation is permitted in (6.1), and we get

$$ {}_{\tau}\mu_{ij}(s,u) = \lambda_{ij}(s,s,u)\,\overline{H}_j(s,0,0)\,\frac{P_{j\mathcal{K}}(s,\tau,0,\infty)}{P_{i\mathcal{K}}(s,\tau,u,\infty)}. $$

Thus, by (4.8) and (4.9),

$$ (6.2) \qquad {}_{\tau}\mu_{ij}(s,u) = \mu_{ij}(s,u)\,\frac{P_{j\mathcal{K}}(s,\tau,0,\infty)}{P_{i\mathcal{K}}(s,\tau,u,\infty)}. $$

We see that the ${}_{\tau}\mu_{ij}(s,u)$ differ from the $\mu_{ij}(s,u)$ unless $P_{i\mathcal{K}}(s,\tau,u,\infty)$ is independent of $i \in \mathcal{K}$ and of u, i.e., unless the probability at time s of remaining in $\mathcal{K}$ until time τ is independent of position in $\mathcal{K}$ and of duration. Since the ${}_{\tau}\mu_{ij}(s,u)$ will typically uniquely determine the ${}_{\tau}P_{ij}(s,t,u,v)$, the latter will therefore differ from the partial probabilities $P_{ij}(s,t,u,v;\mathcal{K})$ (for $0 \leq s \leq t \leq \tau$), again unless $P_{i\mathcal{K}}(s,\tau,u,\infty)$ is independent of u and of $i \in \mathcal{K}$.

6.D. The exception here is sufficiently interesting to merit separate attention. It does happen in practice that $P_{i\mathcal{K}}(s,\tau,u,\infty)$ is independent of u and $i \in \mathcal{K}$ as stated, in which case the partial probabilities have an interpretation as conditional probabilities. For instance, $\mathcal{R}$ may be the death state, which makes $P_{i\mathcal{K}}(s,\tau,u,\infty)$ simply the probability of surviving to time τ (given state i and duration u at time s), and the probability of survival may often be taken as independent of demographic status and duration, at least approximately. (For an application to fertility, see Hoem, 1970a, Subsection 6.C.)

We shall prove a theorem containing conditions sufficient to make $P_{i\mathcal{K}}(s,\tau,u,\infty)$ independent of $i \in \mathcal{K}$ and of u. Let

$$ \mu_{i\mathcal{A}}(s,u) = \sum_{j \in \mathcal{A}-i} \mu_{ij}(s,u) $$

for any subset $\mathcal{A} \subseteq \mathcal{J}$. In some models, $\mu_{i\mathcal{R}}(s,u)$ is independent of $i \in \mathcal{K}$ and of u. (For mortality, this is often approximately correct, as mentioned above.) This means that transitions $\mathcal{K} \to \mathcal{R}$ occur at a rate which is independent of movements within $\mathcal{K}$. It should be possible, then, to calculate $P_{i\mathcal{K}}(s,\tau,u,\infty)$ as if one is faced with a model with the two "states" $\mathcal{K}$ and $\mathcal{R}$ only. Thus, if $\mu_{i\mathcal{R}}(s,u) = \gamma(s)$ for all $i \in \mathcal{K}$ and all u, where $\gamma(\cdot)$ is continuous, one would expect to get

$$(6.3) \qquad P_{i\mathcal{K}}(s,\tau,u,\infty) = \exp\{-\int_s^\tau \gamma(t)dt\}$$

for all $i \in \mathcal{K}$ and all u, under quite weak conditions. Our theorem is as follows.

<u>Theorem 6.4.</u> Assume that $\lim P_{i\mathcal{R}}(s,t,u,\infty)/(t-s) = \gamma(s)$ as $t \downarrow s$, uniformly in $i \in \mathcal{K}$ and $u \geq 0$. Then (6.3) holds under the conditions stated.

<u>Proof</u>: Let $0 \leq s < t < t + \Delta t$, and let $i \in \mathcal{K}$. Then

$$\left| \frac{1}{\Delta t}\{P_{i\mathcal{R}}(s,t+\Delta t,u,\infty) - P_{i\mathcal{R}}(s,t,u,\infty)\} - \gamma(t)\, P_{i\mathcal{K}}(s,t,u,\infty)\right|$$

$$= \left| \frac{1}{\Delta t} \sum_{k \in \mathcal{K}} \int_0^\infty P_{ik}(s,t,u,dv)\, P_{k\mathcal{R}}(t,t+\Delta t,v,\infty) - \gamma(t) \sum_{k \in \mathcal{K}} \int_0^\infty P_{ik}(s,t,u,dv)\right|$$

$$\leq \sum_{k \in \mathcal{K}} \int_0^\infty P_{ik}(s,t,u,dv) \left| \frac{1}{\Delta t} P_{k\mathcal{R}}(t,t+\Delta t,v,\infty) - \gamma(t)\right|$$

$$< \varepsilon\, P_{i\mathcal{K}}(s,t,u,\infty) \quad \text{for } 0 < \Delta t < \eta(\varepsilon,t),$$

by the uniform convergence. Since $P_{i\mathcal{K}}(s,t,u,\infty) = 1 - P_{i\mathcal{R}}(s,t,u,\infty)$, we get

$$\frac{\partial}{\partial t} P_{i\mathcal{K}}(s,t,u,\infty) = -\gamma(t)\, P_{i\mathcal{K}}(s,t,u,\infty),$$

from which the theorem follows. □

This generalizes and corrects Theorem 2 in Hoem (1969c, p. 152).

If (6.3) holds, (6.1) and (4.7) give

$$P_{ij}(s,t,u,v) = {}_\tau P_{ij}(s,t,u,v) \exp\{-\int_s^t \gamma(y)\,dy\}$$

for all i and $j \in \mathcal{K}$. In this case, therefore, a particularly simple relationship exists between the P_{ij} and the ${}_\tau P_{ij}$.

6.E. Conditioning on an event $Z(\tau) \in \mathcal{K}$ is something which typically happens in retrospective studies, often inadvertently. Let us consider an example from biostatistics this time. Suppose that one interviews the patients in a group of health institutions in a particular week to get their medical histories. Evidently, the histories of previous patients, not currently admitted, will not be represented among the data. Now focus on present and previous patients (now possibly dead) aged τ when the interviews are made. Pick an individual among them who at some previous age s was a patient of a given category in a particular institution of the group, and who had been so for a duration u. It is easy to conceive of situations where the probability that such an individual will be a patient in one of the institutions at age τ also (and not dead, recovered, or hospitalized somewhere else) may depend on his illness, on its duration at age s, and possibly also on what institution was then involved. If this is the case, and is important, then occurrence/exposure rates calculated on the basis of the retrospective data will be estimates of the functions ${}_{\tau}\mu_{ij}$ rather than estimates of the μ_{ij}, as probably intended.

Fertility rates calculated on the basis of retrospective questions in population censuses may be analogously affected. So may abortion rates and other measures calculated from retrospective fertility histories, and similarly in many other situations.

6.F. In retrospective studies, some information is left out simply because there is no one available to give it. It happens that the same kind of information is deleted from prospective studies also even though it does get collected. For example, when individuals are lost to follow-up, even data collected before they were lost sight of, is sometimes removed. Similarly, data concerning individuals who are dead or who have left the country, is kept separate from data on the current population in some population registers, and it is then enticing to concentrate on the current file and leave the other files alone.

When collected data are removed like this, one may perhaps say that the rest of the data, on which analysis is subsequently based, are purged of the data deleted. We have therefore suggested (Hoem, 1969c) that functions like the ${}_{\tau}P_{ij}$ and the ${}_{\tau}\mu_{ij}$ be called purged probabilities and purged forces, respectively.

6.G. The purging effect is an artefact of the observational plan. If data are first collected on a prospective basis and then purged, this effect arises because a full use is not made of all information available.

In retrospective studies, the "purging" effect comes in addition to other problems like recall errors. We are therefore reminded once again of the importance of the observational plan and of the fact that the results of substantive analysis are going to be influenced by it. Among demographers, Sheps and her associates have argued this case forcefully in a number of recent papers (Sheps et al., 1970; Menken and Sheps, 1970; Sheps and Menken, 1972). (Compare also Subsections 2.C and D above.)

7. SELECTION, SELECTIVITY, AND SELECT ACTUARIAL TABLES. A DISCUSSION OF TERMINOLOGY

7.A. So far, we have stuck mostly to terminology geared to demography. For various reasons, actuarial modes of expression differ from this in certain respects. Thus, while a demographer is apt to talk of duration-dependence, open and closed birth intervals, and the like, an actuary would probably speak of select tables or models, select forces of fertility, and so on. It is perhaps unfortunate that what is essentially the same phenomenon should have different names in different contexts, but when this is the case anyway, it may be useful to have correspondences pointed out. That is one purpose of the present, final Section of this paper.

7.B. Duration-dependent forces of transition are called select by actuaries. It is apt to be confusing when words like "selected", "selection" and "selectivity" are used (in actuarial science, demography, and elsewhere) also for completely different phenomena, viz. in connection with heterogeneity of the subpopulation with a given status, as well as with the fact that mortality and other rates may depend on demographic status. A second purpose of this Section is to throw some light on the multiplicity of uses of the same set of words.

We shall give some examples. These will be for illustration only. The intention is not to be exhaustive.

7.C. Speaking of mortality, Hooker and Longley-Cook (1953, p. 22) give the following explanation of what "select" means.

"It is now proposed to consider [a] type of table relating to a specific class of lives, in which the functions vary not only with the age but also with the period which has elapsed since entry into the class of lives to which the table relates. Such tables are known as select tables, and the process which makes it necessary to introduce functions which vary with the duration as well as the age is referred to as selection. A life who has just been selected is said to be select. (The word 'selection' has been used here to denote the process of choice of the lives, as opposed to random sampling; in some actuarial literature the word 'selection' is used to denote the subsequent effect on the mortality rates of the process here called selection.)"

Seal (1959) has given a critical review of the evidence for the existence of such selection, as well as of explanations given.

There are a number of reasons why insurance companies would observe select mortality in, say, studies of cohorts of assured lives. We shall list the following three, which are classical suggestions.

(i) There are continuing effects of an initial selection on the part of the insurer or by the assured life (self selection).

(ii) The stock of lives with a given age at issue of insurance is heterogeneous with respect to mortality, and the higher mortality risks get weeded out through death. (This would give an effect towards making mortality for equally old insured lives decrease with duration.)

(iii) The stock of lives with a given age at issue is heterogeneous as stated, and there is a gradual withdrawal from life assurance of healthy lives. (This would give a contrary effect towards making mortality for equally old assured lives increase with duration.)

(In studies of period mortality, one would get an additional effect due to secular mortality improvement, which we want to avoid discussing here.)

7.D. Evidently, these ideas are easily extended to functions other than forces of mortality, and this is done regularly. The main point is that the forces of transition may turn out to depend on, say, age at issue and duration since issue, separately, not only on their sum, which is age attained.

Fortunately, it is not necessary to use the apparatus of semi-Markovian models to give a satisfactory account of this phenomenon. Age at issue will, of course, be constant throughout the entire period of insurance. Duration since issue is then the only important changing time variable involved. As there is no separate duration variable in addition, the theory of inhomogeneous Markov chains with a continuous time parameter, or, in many cases, much simpler mathematics, can therefore easily cope with the situation.

On the other hand, actuarial forces of transition may also depend on current-state-duration, as illustrated by the disability example of Subsection 2.A. Such forces are also called select, and the quotation from Hooker and

Longley-Cook essentially covers this case as well. We see, therefore, that select forces of transition can arise in two ways, viz. through dependence on duration-since-issue and through dependence on duration-in-current-state. From a mathematical point of view, these are quite different phenomena, and the latter one necessitates the use of semi-Markovian machinery.

7.E. So much about select forces and select tables. We now turn to the second type of use of similar words.

When there is a withdrawal from assurance of healthy lives, as suggested in point (iii) of Subsection 7.C, this is described by saying that "withdrawals are selective from a mortality standpoint" (Seal, 1959, p. 175; see also p. 167). This must not be confused, then, with the fact that withdrawal rates are frequently select, i.e. they depend on time since entry. (For the latter empirical fact, compare, e.g., Hooker and Longley-Cook, 1957, pp. 13 and 157.)

In follow-up studies one will frequently postulate that losses are not selective, i.e., that individuals lost to follow-up are homogeneous with the observed population with respect to the phenomenon investigated. For instance, this is surely what Potter (1969, p. 465) means when he assumes that cases lost to follow-up in the k-th months of a study on the use-effectiveness of intrauterine contraception "are unselected relative to the subsample effectively observed during that month".

On the other hand, observed duration-dependence due to heterogeneity of the population is a well-known phenomenon in demography too. (See, e.g., Sheps, 1966, and her references.) Let me mention only that some women seem to have a larger risk of expelling an IUD than others (Tietze, 1968, p. 382). This should contribute towards making the expulsion rate a decreasing function of retention period, as one can observe. (This corresponds to the weeding out of unhealthy lives by mortality, as mentioned in point (ii) of Subsection 7.C.)

In Subsections 2.C and D we mentioned that observed duration-dependence could be due to real-life duration-dependence or could be an artifact caused by the observational plan. We are reminded now that it can also be a result of an inadequacy of the model, in that the heterogeneity of the population is not taken into account explicitly, but only through its consequences for the forces of transition.

7.F. The third use of words like "selectivity" is to express the dependence of forces of transition on the "delivering" state, say the dependence of $\mu_{ij}(s,u)$ on i. Take the following quotation from Hooker and Longley-Cook (1957, p. 21):

"..... let us suppose that a large organisation institutes a pension scheme for its employees and let us assume that, after a certain age, there are no withdrawals from service before the attainment of the normal pension age except by ill-health retirement. It will almost certainly be found that the observed mortality of the ill-health pensioners is heavier ..... . In these circumstances ill-health retirement is stated to be a selective decrement".

Thus, in terms of the three-state disability model of Subsection 2.A, one may expect $\eta(x,u)$ to be greater than $\mu(x)$ ("especially during the first few years after retirement", to continue quoting Hooker and Longley-Cook, 1957, p. 21).

To generalize, assume that there are three states, i, j, and k, such that direct transitions $i \to j$, $i \to k$, and $j \to k$ are possible. Then transition $i \to j$ is said to be a selective decrement with respect to k if $\mu_{ik} \neq \mu_{jk}$. Otherwise it is a non-selective decrement.

The particular constellation of the states i, j, and k is important in the definition of a selective decrement, as is seen from the following counterexample. If, in a certain model, an IUD user is said to be in state n if she has experienced n expulsions during a testing period, then $\mu_{01} < \mu_{12}$, since an observed expulsion has proved to be an indication of a high risk of further expulsion (Tietze, 1968, p. 382). One may perhaps say that there is a differential inclination to expell the IUD. Yet the language of selective decrements does not apply to states 0, 1, and 2.

If, on the other hand, the risk of conception with the IUD in situ is different for expellors than for others, then expulsion can be a selective decrement with respect to in-situ conception.

The fact that transitions are non-selective decrements can simplify many things. Suppose, for instance, that $\mathcal{K}$ in Subsection 6.D consists of a single state k, and suppose that all transitions are non-selective decrements with respect to k. Then μ_{ik} is independent of i. Suppose that it is independent of current duration u too, so that there is no differential tendency to transfer to k at all. Then the theory of Subsection 6.D applies. Thus, the purged probabilities, given non-entry into k, are equal to the partial probabilities arising from the elimination of k, and, therefore, retrospective studies will not be biased due to differential entry into k.

7.G. In summary, we see that words like select, selected, selective, and selection are used for a number of purposes in connection with models involving forces of transition. We have discerned three lines of interpretation which are distinct from each other, yet are closely interrelated, all reflecting important aspects of the models.

REFERENCES

[1] Bühlmann, Hans (1970): "Mathematical Methods in Risk Theory." Springer-Verlag.

[2] Chiang, Chin Long (1971): "A stochastic model for human fertility." Biometrics 27: 345-356.

[3] Chung, Kai Lai (1960): "Markov Chains with Stationary Transition Probabilities." Springer-Verlag.

[4] Çinlar, Erhan (1969): "Markov renewal theory." Advances in Applied Probability 1 (2): 123-187.

[5] de Smit, Jos H.A. (1971): Book review. Rev. Int. Statist. Inst. 39 (2): 248.

[6] Du Pasquier, L. Gustav (1912/13): "Mathematische Theorie der Invaliditätsversicherung." Mitteil. Verein. schweizer. Versich.-mathem. 7: 1-7, and 8: 1-153.

[7] Ginsberg, Ralph B. (1971a): "Semi-Markov processes and mobility." J. Math. Sociol. 1: 233-262.

[8] Ginsberg, Ralph B. (1971b): "On the use and interpretation of probabilistic models: Application of the semi-Markov model to migration." Unpublished manuscript.

[9] Hoem, Jan M. (1969a): "Some notes on the qualifying period in disability insurance. I. Actuarial values." Mitteil. Verein. schweizer. Versich.-mathem. 69 (1): 105-116.

[10] Hoem, Jan M. (1969b): "Markov chain models in life insurance." Blätter Deutsch. Ges. Versich.-mathem. 9 (2): 91-107.

[11] Hoem, Jan M. (1969c): "Purged and partial Markov chains." Skand. Akt. tidskr. 52: 147-155.

[12] Hoem, Jan M. (1970a): "Probabilistic fertility models of the life table type." Theor. Popul. Biol. 1 (1): 12-38.

[13] Hoem, Jan M. (1970b): "A probabilistic approach to nuptiality." Biomētrie-Praximētrie 11 (1): 3-19.

[14] Hoem, Jan M; Jan Riis, and Rolf Sand: "Disability income benefits in group life insurance." Skand. Akt. tidskr., to appear.

[15] Hooker, P.F. and L.H. Longley-Cook (1953/1957): "Life and Other Contingencies." Two Volumes. Cambridge University Press.

[16] Land, Kenneth C. (1971): "Some exhaustible Poisson process models of divorce by marriage cohort." J. Math. Sociol. 1: 213-232.

[17] Lundberg, Filip (1903): "I. Approximerad framställning af sannolikhetsfunktionen. II. Återförsäkring af kollektivrisker." Uppsala: Almqvist & Wiksell.

[18] Lundberg, Ove (1940): "On Random Processes and their Application to Sickness and Accident Statistics." Uppsala: Almqvist & Wiksell.

[19] McGinnis, Robert (1968): "A stochastic model of social mobility." Am. Sociol. Rev. 33 (5): 712-722.

[20] Menken, Jane A. and Mindel C. Sheps (1970): "On relationships between longitudinal characteristics and cross-sectional data." Am. J. Publ. Health 60 (8): 1506-1514.

[21] Neuts, Marcel F. (1968): "A working bibliography on Markov renewal processes and their applications." Purdue University, Department of Statistics, Mimeograph Series No. 140.

[22] Perrin, Edward B. and Mindel C. Sheps (1964): "Human reproduction: A stochastic process." Biometrics 20 (1): 28-45.

[23] Potter, Robert G. (1969): "Use-effectiveness of intrauterine contraception as a problem in competing risks." Pp. 458-484 in "Family Planning in Taiwan. An Experiment in Social Change" (by R. Freedman and J.Y. Takeshita). Princeton University Press.

[24] Potter, Robert G. (1971): "Additional births averted when post conceptive methods are added to pre conceptive methods of fertility control." Paper prepared for the IASPS Satellite Symposium on Statistical Problems of Population Research, East-West Center, Honolulu, August 3-6, 1971.

[25] Pyke, Ronald and Ronald Schaufele (1964): "Limit theorems for Markov renewal processes." Ann. Math. Statist. 35 (4): 1746-1764.

[26] Rowntree, Griselda and Norman H. Carrier (1958): "The resort to divorce in England and Wales, 1858-1957." Population Studies 11 (3): 188-233.

[27] Schoenbaum, E. (1924/25): "Anwendung der Volterra'schen Integralgleichungen in der mathematischen Statistik." Skand. Akt. tidskr. 7: 241-265, and 8: 1-22.

[28] Seal, Hilary L. (1959): "A statistical review of the evidence for the existence of temporary selection." J. Inst. Act. 85: 165-210 (with discussion).

[29] Seal, Hilary L. (1970): "Probability distributions of aggregate sickness durations." Skand. Akt. tidskr. 53: 193-204.

[30] Sheps, Mindel C. (1966): "On the person years concept in epidemiology and demography." Milbank Mem. Fund. Quart. 44 (1): 69-91.

[31] Sheps, Mindel C. (1971): "A review of models for population change." Rev. Int. Statist. Inst. 39 (2): 185-196.

[32] Sheps, Mindel C. and Edward B. Perrin (1964): "The distribution of birth intervals under a class of stochastic fertility models." Population Studies 17 (3): 321-331.

[33] Sheps, Mindel C.; Jane A. Menken; and Annette P. Radick (1969): "Probability models for family building: An analytical review." Demography 6 (2): 161-183.

[34] Sheps, Mindel C.; Jane A. Menken; Jeanne Clare Ridley; and Joan W. Lingner (1970): "The truncation effect in closed and open birth interval data." J. Am. Statist. Ass. 65 : 678-693.

[35] Sheps, Mindel C. and Jane A. Menken (1972): "Distribution of birth intervals according to the sampling frame." Theor. Popul. Biol. 3 (1): 1-26.

[36] Störmer, H. (1970): "Semi-Markoff-Prozesse mit endlich vielen Zuständen." Springer Verlag, Lecture Notes in Operations Research and Mathematical Systems, No. 34.

[37] Tietze, Christopher (1968): "Oral and intrauterine contraception: Effectiveness and safety." International J. Fertility 13 (4): 377-384.

[38] Tolba, Salah El-Din (1966): "On human reproduction" and "On estimating the effects of changing the values of some factors of fertility." Cairo University Press, Proc. Inst. Statist. Studies and Research 1 : 69-178.

[39] U.S. Bureau of the Census (1970): "Probabilities of marriage, divorce, and remarriage." Current Population Reports, Special Studies, Series P-23, No. 32.

[40] Venkatacharya, K. (1971): "A Monte Carlo model for the study of human fertility under varying fecundability." Social Biology 18 (4): 406-415.

Some further references, not mentioned in the text

[41] Janssen, Jacques (1966): "Application des processus semi-markoviens à un problème d'invalidité." Bull. Ass. Roy. Actuaires Belges 63 : 35-52.

[42] Franken, Peter (1968): "Erlangsche Formeln für semimarkowschen Eingang." Elektron. Informationsverarbeit. Kybernetik 4 : 197-204.

[43] Kshirsagar, A.M. and Y.P. Gupta (1968): "Mean variance of the number of renewals in certain Markov renewal processes." J. Indian Statist. Ass. 6 (1-2): 75-80.

[44] Teugels, J.L. (1968): "Exponential ergodicity in Markov renewal processes." J. Appl. Prob. 5 : 387-400.

[45] Hunter, Jeffrey J. (1969): "On the renewal density matrix of a semi-Markov process." Sankhyā A 31 : 281-308.

[46] Hunter, Jeffrey J. (1969): "On the moments of Markov renewal processes." Adv. Appl. Prob. 1 (2): 188-210.

[47] Keilson, J. (1969): "On the matrix renewal function for Markov renewal processes." Ann. Math. Stat. 40 (6): 1901-1907.

[48] Kshirsagar, A.M. and Y.P. Gupta (1969): "Some results in Markov renewal processes." Bull. Calcutta Statist. Ass. 18 (2): 61-72.

[49] Stone, Lawrence D. (1969): "On the distribution of the supremum functional for semi-Markov processes with continuous state spaces." Ann. Math. Statist. 40 (3): 844-853.

[50] Hawkes, A.G. (1970): "Bunching in a semi-Markov process." J. Appl. Prob. 7 (1): 175-182.

[51] Jacod, Jean (1970): "Chaînes semi-markoviennes transientes et recurrentes, chaînes positives." C.r. Acad. Sci., Paris, Sér.A. 270: 776-779.

[52] Kshirsagar, A.M. (1970): "Life estimation and renewal theory." Techn. Rep. No. 78, Department of Statistics THEMIS Contract, Southern Methodist University, Dallas, Texas.

[53] Kshirsagar, A.M. and Y.P. Gupta (1970): "Distribution of the number of Markovian renewals in an arbitrary interval." Austral. J. Statist. 12 : 58-63.

[54] Ross, Sheldon M. (1970): "Average cost semi-Markov decision processes." J. Appl. Probab. 7 : 649-656.

[55] Schäl, M. (1970): "Rates of convergence in Markov renewal processes with auxilliary paths." Z. Wahrscheinl. theorie 16 : 29-38.

[56] Schäl, M. (1970): "Markov renewal processes with auxilliary paths." Ann. Math. Statist. 41 (5): 1604-1623.

[57] Kurtz, Thomas G. (1971): "Comparison of semi-Markov and Markov processes." Ann. Math. Statist. 42 (3): 991-1002.

[58] Lippman, Stephen A. (1971): "Maximal average-reward policies for semi-Markov decision processes with arbitrary state and action space." Ann. Math. Statist. 42 (5): 1717-1726.

[59] Serfozo, Richard F. (1971): "Random time transformations of semi-Markov processes." Ann. Math. Statist. 42 (1): 176-188.