THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

CERTAIN STATISTICS RELATED TO TIME-HOMOGENEOUS BIRTH AND DEATH PROCESSES

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

PETER WILLIAM MEREDITH JOHN

Norman, Oklahoma

1955

CERTAIN STATISTICS RELATED TO TIME-HOMOGENEOUS BIRTH AND DEATH PROCESSES

APPROVED BY

THESIS COMMITTEE

ACKNOWLEDGMENT

I would like to express my sincere gratitude to Professor Casper Goffman, who has been a continual source of inspiration, aid and encouragement. I wish also to thank the members of the committee for their constructive criticisms.

TABLE OF CONTENTS

CERTAIN STATISTICS RELATED TO TIME-HOMOGENEOUS BIRTH AND DEATH PROCESSES

CHAPTER I

INTRODUCTION

In the classical theory of population growth, the size of the population is considered as a continuous variable throughout the whole range of ages, and the changes that occur in this variable through births, deaths or other agencies are considered deterministically. A short sketch of the deterministic theory is given by D. G. Kendall [13].[1]

The deterministic approach is most readily illustrated by the example of the simple birth process; the rate at which the population is increasing is directly proportional to the size n. This leads at once to the linear differential equation,

$$\frac{dn}{dt} = \lambda n ,$$

where λ is the birth rate per capita per unit time. When λ is constant, the solution is the law of Malthus

$$n(t) = n(0) e^{\lambda t}. \qquad (1)$$

[1] Numbers in brackets refer to the list of references. When two numbers are used, the second number is the page number in the reference.

It is clear that so categorical a statement as (1) cannot be a satisfactory answer. Not only is the solution (1) not even a rational number, let alone an integer, although that objection is perhaps more philosophical than practical, but it takes no account of the role of chance effects upon the size of the population.

In the stochastic approach, the probabilities $P_n(t)$ that the population size will be n at time t are calculated. For the example above, the simple birth process, these probabilities were first obtained by Yule [21]. For this reason, the process is occasionally called a Yule process. It is, however, more usual nowadays to call it a Furry process, since Furry considered the application of such a process to the physical problem of the passage of radioactive particles through a slab of lead [9]. The first paper published with a systematic mathematical approach to this problem is due to Feller [5], 1939.

If, in addition, the population may diminish with a death rate μ, we have a birth and death process. In the case where the rate of change of the population size is directly proportional to the size itself, the simple process, complete results for the $P_n(t)$ may be obtained by the method of generating functions. This method is attributed by Arley and Borchsenius [1] to C. Palm (unpublished), and may be used whether the parameters λ and μ are constants or functions of t. Kendall obtains the results for the birth and death process using this method [12].

When the rate of change is not directly proportional to n, and when the birth and death rates are not necessarily the same functions for every n, we have a generalized process, which is discussed by Feller [7]. In the simple case, the population size obtained by the deterministic approach appears as the expected size of the population in the stochastic approach; this is not, however, true in general.

It may happen in the general case that the population size becomes infinite in finite time. This is the Feller-Lundberg phenomenon, and in terms of probabilities it may be expressed as

$$\sum_{n=1}^{\infty} P_n(t) < 1 \quad .$$

The second chapter of this paper is concerned with this phenomenon. The problem is treated by considering the expected time of such a transition to infinity. By this means, the Feller-Lundberg result for birth processes is obtained in a direct and straightforward manner. Feller [7] obtains the result as a consequence of a more general theorem, the proof of which is by no means readily comprehensible. Feller also gives an elementary proof in [7] which is incorrect, since he proves the sufficiency condition twice; he repeats this in his book [8, 370].

We also obtain by this method necessary and sufficient conditions for the occurrence of the phenomenon in birth and death processes; our conditions are similar to those obtained by Reuter and Ledermann [15], [19], although they differ in detail. We obtain new information concerning this phenomenon. In particular, we find that if the expected time of the occurrence of an infinite number of events is finite, then the phenomenon occurs with probability 1, and if the expected time is infinite the phenomenon occurs with probability 0.

The purpose of the third chapter is to lay the foundation for the final two chapters. It comprises a short survey of the present state of knowledge of estimation of the parameters of these processes, and introduces the concept of proper, improper and strictly improper statistics. In it, also, the Poisson and Furry processes are defined in terms of the stochastic theory. The Furry process has been mentioned above; for the Poisson process the probability $P_n(t)$ that n events (births) occur in time t is given, for all n, by

$$P_n(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!} ,$$

It is characterized by the fact that the number of events occurring in any time interval is independent of the population size.

In Chapter IV, tests are obtained for the hypothesis that a given process is actually Poissonian. It is necessary for this purpose to obtain statistics whose distributions are independent of the value λ of the parameter. Such statistics are called strictly improper, and a family of them is obtained, each one of which provides a test for the hypothesis. These statistics are of the form $\frac{ax + by}{cx + dy}$, where a, b, c and d are constants such that $ad > bc$, and $cd > 0$, and x, y are the lengths of time intervals between events. The simplest of these is the statistic $\frac{x - y}{x + y}$, which we show to have the rectangular distribution over the range $(-1, 1)$. In general, the range is $\left(\frac{b}{d}, \frac{a}{c}\right)$, but the distribution is only rectangular if $c = d$. These statistics are also adaptable to the Furry process.

In Chapter V, we obtain the distribution of the statistic $\frac{x - y}{x + y}$, where x, y are lengths of time intervals between births in a Furry process. We show that, if x is the length of the i^{th} interval and y is the length of the ki^{th} interval, then the distribution of the statistic depends upon k, but not upon i. This result leads to a test for discriminating between the hypothesis that a process is Poissonian and the hypothesis that it is a Furry process. When k is an integer, the test is quite simple, and, for this case, we give a table showing, for various values of k, the critical value of the statistic and the number of observations required at the 0.05 level of significance.

CHAPTER II

DIVERGENT PROCESSES

In a generalized birth and death process we consider a population, the size of which is given by an integer-valued random variable $n(t)$. We shall denote by $P_{ij}(\tau, t)$ the probability that the population size will be j at time t, subject to the condition that it is i at time τ. The process is a Markov process in which the transitions are governed by the following rules.

In an element of time Δt possible transitions and their associated probabilities are given as follows. If $n(t)$ denotes the population size at time t, then the probability that

$$
(1)\left\{
\begin{array}{ll}
n(t+\Delta t) = n(t) \pm k,\ k > 1 & \text{is } \sum P_{n,n\pm k}(t, t+\Delta t) = o(\Delta t), \\
\text{that } n(t+\Delta t) = n(t)+1 & \text{is } P_{n,n+1}(t,t+\Delta t) = \lambda_n(t) \ t+o(\Delta t), \\
\text{that } n(t+\Delta t) = n(t) & \text{is } P_{n,n}(t,t+\Delta t) = 1-\{\lambda_n(t)+\mu_n(t)\}\Delta t+o(\Delta t), \\
\text{and that } n(t+\Delta t) = n(t)-1 & \text{is } P_{n,n-1}(t,t+\Delta t) = \mu_n(t)\Delta t+o(\Delta t),
\end{array}
\right.
$$

where $o(\Delta t)$ has its usual meaning, i.e.,

$$\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0.$$

The functions $\lambda_n(t)$ and $\mu_n(t)$ are in general functions of both n and t. If all the $\lambda_n(t)$, $\mu_n(t)$ are independent of t, the process is said to be time-homogeneous.

An increase of one in the population size is called a birth; a decrease of one is called a death; the occurrence of either will also occasionally be called an event.

We recall that a process is called a Markov process if the probability distribution of the random variable, in this case n(t), at time $t+\Delta t$ is completely determined by the value of the variable at time t, and is independent of the previous development of the process. For a Markov process, it is well known that the probabilities satisfy the following relations, the so-called Chapman-Kolmogorov equations:

$$(2) \qquad P_{ij}(\tau,t) = \sum_k P_{ik}(\tau,s)\, P_{kj}(s,t) \quad \text{for all} \quad \tau < s < t.$$

Thus, in connection with the birth and death process, we must find a set of functions satisfying both (1) and (2).

Equations (1) lead formally to an infinite set (3) of differential equations for the $P_{ij}(\tau,t)$ as follows:

$$P_{ij}(\tau,t+\Delta t) = P_{i,j-1}(\tau,t)\{\lambda_{j-1}(t)\Delta t+o(\Delta t)\} + P_{i,j+1}(\tau,t)\{\mu_{j+1}(t)\Delta t+o(\Delta t)\}$$

$$+P_{ij}(\tau,t)\{1-\lambda_j(t)\Delta t-\mu_j(t)\Delta t+o(\Delta t)\} + \sum_k P_{i,j+k}(\tau,t)o(\Delta t) .$$

$$\frac{P_{ij}(\tau,t+\Delta t)-P_{ij}(\tau,t)}{\Delta t} = \lambda_{j-1}(t)P_{i,j-1}(\tau,t)+\mu_{j+1}(t)P_{i,j+1}(\tau,t)$$

$$-\{\lambda_j(t)+\mu_j(t)\}\; P_{ij}(\tau,t)$$

$$+\frac{1}{\Delta t}\{P_{ij}(\tau,t)+P_{i,j\pm 1}(\tau,t)+\sum_k P_{i,j\pm k}(\tau,t)\}\, o(\Delta t).$$

Taking the limit as $\Delta t \to 0$ we have

$$(3)\quad \frac{\partial}{\partial t}(P_{ij}(\tau,t)) = \lambda_{j-1}(t)P_{i,j-1}(\tau,t)-\{\lambda_j(t)+\mu_j(t)\}P_{ij}(\tau,t)$$

$$+\mu_{j+1}(t)P_{i,j+1}(\tau,t).$$

Feller [7] has shown that under very general conditions the solutions to the system (3) exist and satisfy (2), so that since they obviously satisfy (1), they furnish a solution to the problem. The uniqueness question is more delicate, and has been discussed by Doob [4] and by Reuter and Ledermann [15], [19]. It bears a certain relation to the Feller-Lundberg phenomenon which will be discussed below.

We now prove a lemma which will be needed subsequently.

__LEMMA 2.1:__ The probability $P_{ii}^{(0)}(\tau,t)$ that the population size remains at the value i throughout the interval (τ,t) is

$$P_{ii}^{(0)}(\tau,t) = e^{-\int_\tau^t \{\lambda_i(s)+\mu_i(s)\}\, ds} .$$

__PROOF:__ We have

$$P_{ii}^{(0)}(\tau,t+\Delta t) = P_{ii}^{(0)}(\tau,t)\{1-\lambda_i(t)\Delta t-\mu_i(t)\Delta t+o(\Delta t)\},$$

so that

$$\frac{P_{ii}^{(0)}(\tau,t+\Delta t)-P_{ii}^{(0)}(\tau,t)}{\Delta t} = -P_{ii}^{(0)}(\tau,t)\{\lambda_i(t)+\mu_i(t)\}+\frac{1}{\Delta t}\{P_{ii}^{(0)}(t)o(\Delta t)\}$$

and

$$\frac{\partial}{\partial t}P_{ii}^{(0)}(\tau,t) = -P_{ii}^{(0)}(\tau,t)\{\lambda_i(t)+\mu_i(t)\},$$

giving

$$\log P_{ii}^{(0)}(\tau,t) = -\int_0^t \{\lambda_i(s)+\mu_i(s)\}ds+c(\tau).$$

But,

$$P_{ii}^{(0)}(\tau,\tau) = 1 \qquad \text{for all } \tau ,$$

so that

$$0 = -\int_0^\tau (\lambda_i(s)+\mu_i(s))ds+c(\tau) \qquad \text{for all } \tau ,$$

and

$$c(\tau) = \int_0^\tau (\lambda_i(s)+\mu_i(s))ds.$$

Therefore,

$$P_{ii}^{(0)}(\tau,t) = e^{-\int_\tau^t (\lambda_i(s)+\mu_i(s))ds}.$$

It is the purpose of this chapter to consider the relation $P_{ij}(\tau,t) \leqq 1$ for time-homogeneous processes. It is a remarkable fact observed by Feller [7] and apparently independently by Lundberg [16] that the inequality sign cannot in general be replaced by an equality. Feller and Lundberg show that a necessary and sufficient condition for the inequality to hold for a pure birth process (i.e., a process for which all the μ_n are zero) is that $\sum_{n=1} \frac{1}{\lambda_n}$ should converge. Further work has been done by Reuter and Ledermann [15] and [19], who have obtained conditions for the inequality $\sum_j P_{ij}(\tau,t) < 1$ for a time-homogeneous birth and death process. We obtain a new result regarding this phenomenon which gives further insight into its nature, and from which the results of the above authors follow. In particular, we show that, if there is a positive probability that an infinite number of events occur in a finite time, then that probability must equal 1.

For convenience we shall assume that the population size is 1 when t is zero, and we shall write $p_n(t)$ for the probability that the population size is n at time t.

The inequality $\sum_1^\infty p_n(t) < 1$ means that there is a positive probability that n is infinite at time t, i.e., that infinitely many transitions have taken place in finite time. The results will be obtained by considering the expected value of the time taken for such a transition to infinity.

We shall call a process for which $\sum_{n=1}^{\infty} p_n(t) < 1$ for some value of t, a divergent process.

<u>THEOREM 2.1</u>: For a time-homogeneous pure birth process the expected value of the time $\tau_m = t_{m+1} - t_m$, where t_k is the time at which the population becomes k, is $\frac{1}{\lambda_m}$. ($\lambda_m > 0$.)

<u>PROOF</u>: By Lemma 2.1

$$p^{(0)}_{m,m}(s_1, s_2) = e^{-\lambda_m(s_2 - s_1)}.$$

Then the probability that $t \leq \tau_m \leq t + \Delta t$ is

$$P(t \leq \tau_m \leq t + \Delta t) = e^{-\lambda_m t} \lambda_m \Delta t + o(\Delta t),$$

so that

$$E(\tau_m) = \int_0^\infty \lambda_m t e^{-\lambda_m t} dt$$

$$= \left[-te^{-\lambda_m t}\right]_0^\infty + \int_0^\infty e^{-\lambda_m t} dt$$

$$= -\frac{1}{\lambda_m}\left[e^{-\lambda_m t}\right]_0^\infty = \frac{1}{\lambda_m} .$$

<u>COROLLARY 2.1</u>: If t_∞ is the time taken for the population size to pass from 1 to ∞, then $\bar{t}_\infty \equiv E(t_\infty) = \sum_{m=1}^\infty \frac{1}{\lambda_m}$.

<u>PROOF</u>: We have

$$t_\infty = \sum_{m=1}^\infty \tau_m .$$

Then

$$\bar{t}_\infty = E(\sum_{m=1}^\infty \tau_m) = \sum_{m=1}^\infty E(\tau_m) = \sum_{m=1}^\infty \frac{1}{\lambda_m} .$$

In a stationary birth and death process it is possible that the population size may never reach some given finite value for two reasons:

(i) it may happen that $\lambda_j = 0$ for some j, $1 \leqq j \leqq n$;

(ii) there may be an attainable lower absorbing barrier for the population size. An example which we may consider is the case in which $\mu_1 > 0$, $\lambda_0 = 0$; when once the population dies out, there is no re-birth.

In case (i) it is not possible for the inequality $\sum_1^\infty p_n(t) < 1$

to be satisfied, since $\sum_{1}^{j} p_n(t) = 1$. In case (ii) we may still have $\sum p_n(t) < 1$. As a simple example, consider the process in which $\mu_1 > 0$, $\mu_m = 0$, for all $m > 1$, $\lambda_0 = 0$ and λ_m, $m > 0$, satisfy the requirements for a divergent birth process.

We shall therefore denote by τ_m the time for n to increase from m to m+1, and we shall consider $E'(\tau_m)$, the conditional expected time of the change, by which is meant the expected time, conditional on the change not being prevented by passage to an absorbing barrier in the interim.

<u>THEOREM 2.2</u>: For a time-homogeneous birth and death process $E(\tau_m)$ is given by the recursion formula

$$E(\tau_m) = \frac{1}{\lambda_m} + \frac{\mu_m}{\lambda_m} E(\tau_{m-1}) .$$

<u>PROOF</u>: By Lemma 2.1,

$$P^{(0)}_{m,m}(s_1, s_2) = e^{-(\lambda_m+\mu_m)(s_2-s_1)} .$$

Hence, by the same method as was adopted in Theorem 2.1, the expected time for a further event to occur after the population size reaches m is $\frac{1}{\lambda_m+\mu_m}$.

Such an event has probability $\frac{\lambda_m}{\lambda_m+\mu_m}$ of being a birth, in which case the population has passed from m to m+1 as required, and probability $\frac{\mu_m}{\lambda_m+\mu_m}$ of being a death, in which case the desired increase requires further passage from m-1 to m and then from m to m+1.

We thus have:

$$E(\tau_m) = \frac{\lambda_m}{\lambda_m+\mu_m} \cdot \frac{1}{\lambda_m+\mu_m} + \frac{\mu_m}{\lambda_m+\mu_m}\left(\frac{1}{\lambda_m+\mu_m} + E(\tau_{m-1}) + E(\tau_m)\right) ,$$

$$E(\tau_m)(\lambda_m+\mu_m)^2 = \lambda_m+\mu_m+\mu_m(\lambda_m+\mu_m)(E(\tau_{m-1})+E(\tau_m)) ,$$

$$(\lambda_m+\mu_m)E(\tau_m) = 1 + \mu_m E(\tau_{m-1}) + \mu_m E(\tau_m) ,$$

yielding

(4) $$E(\tau_m) = \frac{1}{\lambda_m} + \frac{\mu_m}{\lambda_m} E(\tau_{m-1}) .$$

In case (i) $E(\tau_j)$ is infinite if $\lambda_j = 0$, and all subsequent $E(\tau_j)$ are infinite.

In case (ii) $E(\tau_1)$ is infinite, there being a positive probability of passage to the absorbing barrier $n = 0$, and hence all the $E(\tau_m)$ are infinite.

We therefore use $E'(\tau_m)$, the conditional expected time of passage defined above, and our equation becomes

(4') $$E'(\tau_m) = \frac{1}{\lambda_m} + \frac{\mu_m}{\lambda_m} E'(\tau_{m-1}) .$$

(If there is no absorbing barrier, we have $E(\tau_m) = E'(\tau_m)$.)

Writing $\overline{\tau}_m$ for $E'(\tau_m)$ we have

$$\overline{\tau}_m = \frac{1}{\lambda_m} + \frac{\mu_m}{\lambda_m} \overline{\tau}_{m-1} ,$$

which is the equation obtained by Reuter and Ledermann [19].

Furthermore (if $n = 0$ is the absorbing barrier, or if $\mu_0 = 0$, when there is no absorbing barrier)

$$\overline{\tau}_1 = \frac{1}{\lambda_1}, \qquad \overline{\tau}_2 = \frac{1}{\lambda_2} + \frac{\mu_2}{\lambda_2\lambda_1},$$

$$\overline{\tau}_n = \frac{1}{\lambda_n} + \frac{\mu_n}{\lambda_n\lambda_{n-1}} + \cdots + \frac{\mu_n \cdots \mu_2}{\lambda_n \cdots \lambda_1},$$

which results differ in detail from those of the above authors. They, using ω where we have used $\overline{\tau}$, have begun their ω sequence with ω_i and have put "$\omega_i = \frac{1}{\lambda_i} + \frac{\mu_i}{\lambda_i}$ (for formal convenience)", which would seem tacitly to assume that $\omega_{i-1} = 1$.

The conditional expected time of passage to infinity is therefore given by $\overline{t}_\infty = \sum_{n=1}^{\infty} \overline{\tau}_n$.

We proceed to obtain our main results. The first is obtained as

<u>THEOREM 2.3</u>: If $\overline{t_\infty}$ is finite, there are values of t for which $\sum_{n=0}^{\infty} p_n(t) < 1$.

<u>PROOF</u>: $\sum_{n=0}^{\infty} p_n(t) = 1$ implies that the probability that $t_\infty \leqq t$ is zero, which in turn implies that

$$P(t_\infty > t) = 1 .$$

Using Cramér's generalization of the Tchebycheff inequality [3,182] we have, for all t,

$$P(t_\infty \geq t) \leq \frac{E(t_\infty)}{t} = \frac{\bar{t}_\infty}{t},$$

so that for $t > \bar{t}_\infty$

$$\sum_{n=0}^{\infty} p_n(t) = P(t_\infty \geq t) \leq \frac{\bar{t}_\infty}{t} < 1,$$

and indeed, by taking t large enough, $\sum_{n=1}^{\infty} p_n(t)$ may be made as small as we wish.

Thus if $\bar{t}_\infty$ is finite, then for all $t > \bar{t}_\infty$, $\sum_{n=0}^{\infty} p_n(t) < 1$.

The second main result is

THEOREM 2.4: If there is a finite time τ such that $\sum p_n(t) < 1$, then $\bar{t}_\infty$ is finite.

In case (i) for a birth and death process the requirement of this theorem is never satisfied. In case (ii) the probabilities below are to be taken as conditional probabilities, conditional upon non-absorption.

PROOF: Suppose that $p_{1\infty}(\tau) = 1 - \sum_0^\infty p_n(\tau) = \alpha > 0$;

then

$$P[n(\tau) < \infty] = 1 - \alpha \quad \text{and} \quad P_{i\infty}(\tau) \geq \alpha, \qquad i \geq 1.$$

$$P[n(m\tau) < \infty] \leq (1 - \alpha)^m;$$

so that

$$P[n(m\tau) < \infty,\ n((m+1)\tau) = \infty] \leq (1 - \alpha)^m;$$

thus

$$\bar{t}_\infty \leq \sum_{m=0}^{\infty} (m+1)\tau \cdot P[n(m\tau) < \infty,\ n((m+1)\tau) = \infty]$$

$$\leq \sum_{m=0}^{\infty} (m+1)\tau(1-\alpha)^m = \tau \sum_{m=0}^{\infty}(m+1)(1-\alpha)^m.$$

But the series $\sum(m+1)x^m$ converges for $|x| < 1$, being the derivative of the power series $\sum x^m$.

Therefore, $\bar{t}_\infty$ is finite.

<u>COROLLARY 2.2</u>: A necessary and sufficient condition for a process to be divergent is that $\bar{t}_\infty$ shall be finite.

For a pure birth process this condition becomes $\sum \frac{1}{\lambda_n} < \infty$, and for a birth and death process we have $\sum \bar{\tau}_n < \infty$.

<u>COROLLARY 2.3</u>: For a birth and death process with no lower absorbing barrier, the probability that an infinite number of events will take place in finite time is either zero or 1.

<u>PROOF</u>: Suppose $\bar{t}_\infty$ is finite, then from Theorem 2.3 we have for all $t > \bar{t}_\infty$

$$P(t_\infty > t) \leq \frac{\bar{t}_\infty}{t}.$$

But, $\frac{\bar{t}_\infty}{t} \to 0$ as $t \to \infty$, so that

$$P(t_\infty < t) \to 1,$$

i.e., if $\bar{t}_\infty$ is finite, the probability that an infinite number of events occurs in finite time is 1.

It follows immediately from Theorem 2.4 that if the probability of an infinite number of events in finite time is not zero, then $\bar{t}_\infty$ is finite, so that the probability must therefore be 1.

COROLLARY 2.4: In any birth and death process the $\bar{t}_{i\infty}$ are either finite for all i or infinite for all i, where $\bar{t}_{i\infty}$ is the expected time of passage from $n = i$ to $n = \infty$.

CHAPTER III

GENERAL REMARKS ON ESTIMATION AND TESTING HYPOTHESES

With regard to the general theory of estimation and testing hypotheses, it seems desirable to classify statistics into three categories, <u>proper</u>, <u>improper</u> and <u>strictly improper</u>. Since in the sequel we shall be concerned mainly with Poisson and Furry processes, it is appropriate here to define these processes and give a short account of the work already done in the theory of estimation of their parameters. We shall also distinguish below between these three classes of statistics in terms of the Poisson process.

A Poisson process is a pure birth process in which $\lambda_n = \lambda$, a constant for all n; λ is called the parameter of the distribution. It can readily be shown that if $n = 1$, when $t = 0$, $p_n(t)$ is given by

$$p_n(t) = \frac{(\lambda t)^n e^{-\lambda t}}{n!} \qquad \text{for all } n.$$

The Furry process is a pure birth process in which $\lambda_n = n\lambda$ for all n. If $n(0) = 1$, we have

$$p_n(t) = e^{-\lambda t}(1 - e^{-\lambda t})^{n-1} \quad .$$

There are two approaches to the problem of estimation of the parameters of these processes, depending upon different stop rules.

The first method is to observe the number of events occurring in a given time. If for the Poisson process n events occur in the time interval (0, T), the maximum likelihood estimator of λ is

$$\hat{\lambda} = \frac{n}{T} .$$

Kendall [13] estimates the parameter of a Furry process in this way. The process is observed at times $t = 0, \tau, 2\tau, \ldots, k\tau = T$, the observed population numbers being respectively $N_0, N_1, \ldots, N_k$.

He obtains the maximum likelihood estimator of $e^{\lambda t}$

$$e^{\hat{\lambda}\tau} = \frac{N_1+N_2+\ldots+N_k}{N_0+N_1+\ldots+N_{k-1}} ,$$

understanding the number of replications to be very large.

A similar estimator of the parameter for a simple death process has been obtained by Immel [11].

The second approach considers the time intervals between successive events, and, for the Poisson process, has been discussed by Maguire, E. S. Pearson and Wynn [17], and also by Moran [18].

Instead of fixing the time of the experiment, we observe the occurrence of $N + 1$ events, noting the N intervals between them. These intervals have the probability distribution

$$dp = \lambda e^{-\lambda t} dt = e^{-\lambda t} d(\lambda t) ,$$

so that λt is distributed as a $\Gamma(1)$ variate, or $2\lambda t$ is distributed as χ^2 with 2 degrees of freedom.

If the sum of the intervals is T, then, from the additive property of χ^2, it follows that $2\lambda T$ is distributed as χ^2 with 2N degrees of freedom. The estimators are then readily obtained. If χ^2_α and $\chi^2_{1-\alpha}$ are the upper and lower α % points of the distribution of $\chi^2_{[2N]}$, there is probability $1-2\alpha$ that

$$\frac{\chi^2_{1-\alpha}}{2T} < \lambda < \frac{\chi^2_\alpha}{2T} .$$

The maximum likelihood point estimator of λ is

$$\hat{\lambda} = \frac{N}{T} .$$

By a similar method Moran [18] has obtained an estimator for the parameter of the Furry process.

In order to clarify the classification of statistics which was made at the beginning of this chapter, we consider the case of a Poisson process. The above methods of estimation are proper, since they always provide a confidence interval $\lambda_0 < \lambda < \lambda_1$ for λ. A statistic is improper if it sometimes yields a confidence interval for λ and sometimes rejects all values of λ. A strictly improper test is one which either accepts or rejects the hypothesis that the process is Poissonian, but yields no information regarding the value of λ.

The remainder of this paper is concerned with strictly improper statistics only, and it therefore seems appropriate to say a few words here about improper statistics. In regard to the Poisson process with parameter value λ, the mean value and variance of the time for an event are $\frac{1}{\lambda}$ and $\frac{1}{\lambda^2}$ respectively. This suggests that the number pair

$$\left(\frac{1}{n} \sum_{i=1}^{n} x_i \, , \; \frac{1}{n} \sum_{i=1}^{n} \left(x_i - \frac{1}{n}\sum x_i\right)^2 \right)$$

be considered as a point in the plane and that a confidence interval for λ be given in terms of a portion of the parabola $y = x^2$ which is sufficiently near to this point. On the other hand, if the above point is far enough away from the parabola we infer that the process is not Poissonian. This idea, which is intuitively simple and appealing, seems difficult to put on a rigorous basis, and for this reason we do not discuss it further. However, crude tests can easily be derived.

CHAPTER IV

HYPOTHESIS TESTING FOR THE POISSON PROCESS

In this chapter we shall consider statistics which lead to strictly improper tests for the Poisson process.

Before treating our main problem we give a test due to Maguire, E. S. Pearson and Wynn [17]. Following their estimate of λ to which reference was made above, they make use of the fact that if χ_1^2 and χ_2^2 are two independent values of χ^2 having, respectively, ν_1 and ν_2 degrees of freedom, then the variance ratio is defined as

$$F(\nu_1, \nu_2) = \frac{\chi_1^2}{\nu_1} \Big/ \frac{\chi_2^2}{\nu_2} ,$$

which is a distribution well known in analysis of variance.

If, then, we are given a set of N intervals between events and wish to test the hypothesis that the process is Poissonian, we may divide the set into two subsets, one containing n_1 intervals with a total time T_1, and the other containing n_2 intervals with total time T_2.

Then, if the hypothesis is true, the ratio

$$\frac{2\lambda T_1}{2n_1} \Big/ \frac{2\lambda T_2}{2n_2} = \frac{T_1 n_2}{T_2 n_1}$$

is distributed as F with $2n_1$ and $2n_2$ degrees of freedom. In particular if $n_1 = n_2 = \frac{1}{2}N$, the statistic becomes T_1/T_2 to be distributed as $F(N,N)$.

In a note on the paper by Maguire, Pearson and Wynn, Barnard [2] has suggested the use of Kolmogorov's D statistic [13], [6][1] to test the hypothesis that the process is Poissonian.

As we have shown, the probability density of the length of an interval t between two consecutive events in a Poisson process is

$$dp = \lambda e^{-\lambda t} dt .$$

This is the exponential distribution. The hypothesis that a process is Poissonian is equivalent to the hypothesis that the random variable t has the exponential distribution; we therefore proceed to develop some of the properties of this distribution.

The characteristic function of the exponential distribution is

$$\int_0^\infty e^{-(\lambda - i\theta)t} dt = \frac{\lambda}{\lambda - i\theta} = \frac{1}{1 - \frac{i\theta}{\lambda}} .$$

The cumulant generating function is therefore

$$\kappa(\theta) = -\log\left(1 - \frac{i\theta}{\lambda}\right) = \sum_r \left(\frac{i\theta}{\lambda}\right)^r \cdot \frac{1}{r} ,$$

so that

$$\kappa_r = \frac{(r-1)!}{\lambda^r} .$$

[1] The reference is made to the paper by Feller because the paper by Kolmogorov is not readily available.

We calculate the following population moments:

$$\mu_1' = \kappa_1 = \frac{1}{\lambda} ,$$

$$\mu_2 = \kappa_2 = \frac{1}{\lambda^2} ,$$

$$\mu_3 = \kappa_3 = \frac{2}{\lambda^3} ,$$

$$\mu_4 = \kappa_4 + 3\kappa_2^2 = \frac{6}{\lambda^4} + \frac{3}{\lambda^4} = \frac{9}{\lambda^4} .$$

We now consider the coefficient of variation of a sample of size n defined by

$$V_n = \frac{\text{s.d.}}{\text{mean}} = \frac{\sqrt{m_2}}{\bar{t}} .$$

<u>THEOREM 4.1</u>: The distribution of V_n converges to the normal distribution with mean 1 and variance $\frac{1}{n}$.

<u>PROOF</u>: Cramer [3, 357 and 367] has shown that for a random variable which takes only positive values and whose first four moments exist, the sampling distribution of the coefficient of variation W_n, $n=1, 2, \ldots$, is asymptotically normal, with mean and variance given by

$$E(W_n) = \frac{\sqrt{\mu_2}}{\mu_1'} + O\left(\frac{1}{n}\right) ,$$

$$\text{Var}(W_n) = \frac{\mu_1'^2\left(\mu_4 - \mu_2^2\right) - 4\mu_1'\mu_2\mu_3 + 4\mu_2^3}{4\mu_1'^4\,\mu_2\, n} + O\left(\frac{1}{n^{3/2}}\right) .$$

In the case of the exponential distribution these formulae lead to

$$E(V_n) = \sqrt{\frac{1}{\lambda^2}}\Big/\frac{1}{\lambda} + O\left(\frac{1}{n}\right) = 1 + O\left(\frac{1}{n}\right),$$

$$\mathrm{Var}\ (V_n) = \frac{\frac{1}{\lambda^2}\left(\frac{9}{\lambda^4} - \frac{1}{\lambda^4}\right) - 4 \cdot \frac{1}{\lambda} \cdot \frac{1}{\lambda^2} \cdot \frac{2}{\lambda^3} + \frac{4}{\lambda^6}}{4 \cdot \frac{1}{\lambda^4} \cdot \frac{1}{\lambda^2} \cdot n} + O\left(\frac{1}{n^{3/2}}\right)$$

$$= \frac{1}{n} + O\left(\frac{1}{n^{3/2}}\right).$$

Thus to the given order of approximation both $E(V_n)$ and Var (V_n) are independent of the parameter λ, and to the same order of approximation V_n is asymptotically normal with mean 1 and variance $\frac{1}{n}$.

It follows that to this order of approximation V_n is a statistic which leads to a strictly improper test. From a practical standpoint, however, we are handicapped by the fact that the sizes of the terms $O(\frac{1}{n})$ and $O\left(\frac{1}{n^{3/2}}\right)$ are not precisely known and may indeed be quite large, so that we are unable to make calculations of the size of sample required for any given confidence limit. We conjecture that these approximating terms actually vanish, but have not proved this.

We now proceed to the main results of the chapter, which develop statistics for the Poisson process that are independent of λ and may accordingly be used to test the hypothesis that the process is Poissonian at all.

<u>THEOREM 4.2</u>: Let x and y be two independent observations from an exponential distribution

$$dp = \lambda e^{-\lambda t} dt.$$

The distribution of the random variable

$$w = \frac{ax + by}{cx + dy},$$

where a, b, c, d are constants, $cd > 0$, and $\Delta = ad-bc \neq 0$ is independent of the parameter λ.

There is no loss of generality in taking a, c, and d to be positive. We shall assume for convenience that the determinant

$$\Delta = \begin{vmatrix} a & b \\ c & d \end{vmatrix} > 0, \quad \text{i.e.,} \quad \frac{a}{c} > \frac{b}{d} ;$$

$\frac{a}{c} < \frac{b}{d}$ implies that $b > 0$, so that this case may be handled by the interchange of the independent variables x and y. The necessity for the restriction $cd > 0$ will be seen later.

LEMMA 4.1:

$$\frac{b}{d} \leqq \frac{ax + by}{cx + dy} \leqq \frac{a}{c}$$

(the equalities occurring when either x or y is zero).

PROOF: Since $\Delta > 0$, $\frac{ad}{c} > b$.

So

$$\frac{ax + by}{cx + dy} \leqq \frac{ax + \frac{ady}{c}}{cx + dy} = \frac{acx + ady}{c(cx + dy)} = \frac{a}{c} .$$

Similarly

$$\frac{ax+by}{cx+dy} \geqq \frac{\frac{bcx}{d} + by}{cx + dy} = \frac{b}{d} ,$$

i.e.,

$$\frac{b}{d} \leqq \frac{ax + by}{cx + dy} \leqq \frac{a}{c} .$$

PROOF OF THEOREM 4.2: We compute the cumulative distribution function, $F(\theta)$.

$$P(w > \theta) = P(ax + by > \theta(cx + dy))$$

$$= P(x(a - c\theta) > y(d\theta - b)).$$

We see from the lemma that neither $(a - c\theta)$ nor $(d\theta - b)$ is negative.

Hence

$$P(w > \theta) = P(x > \frac{d\theta - b}{a - c\theta} y),$$

i.e.,

$$P(w > \theta) = \int_{y=0}^{\infty} \int_{x=\frac{d\theta-b}{a-c\theta}y}^{\infty} \lambda^2 e^{-\lambda x} e^{-\lambda y}\,dx\,dy$$

$$= \int_{y=0}^{\infty} \lambda e^{-\lambda y} \left[-e^{-\lambda x}\right]_{\frac{d\theta-b}{a-c\theta}y}^{\infty} dy$$

$$= \int_{0}^{\infty} \lambda\, e^{-\lambda y\left[1+\frac{d\theta-b}{a-c\theta}\right]} dy$$

$$= \frac{1}{1+\frac{d\theta-b}{a-c\theta}} = \frac{a-c\theta}{a-c\theta+d\theta-b} \,.$$

So

(5) $$P(w \leqq \theta) = 1-P(w > \theta) = \frac{d\theta-b}{(d-c)\theta+a-b} = F(\theta)\,.$$

Thus $F(\theta)$ is independent of λ and the theorem is proved.

<u>COROLLARY 4.1</u>: The probability density function is

$$f(\theta) = \frac{\Delta}{\{(d-c)\theta + (a-b)\}^2}\,.$$

<u>PROOF</u>:

$$f(\theta) = \frac{d}{d\theta} \cdot F(\theta) = \frac{d\{(d-c)\theta + (a-b)\} - (d-c)(d\theta - b)}{\{(d-c)\theta + (a-b)\}^2}$$

$$= \frac{d^2\theta - cd\,\theta + ad - bd - d^2\theta + cd\,\theta + bd - bc}{\{(d-c)\theta + (a-b)\}^2}$$

(6) $$= \frac{\Delta}{\{(d-c)\theta + (a-b)\}^2}\,.$$

<u>THEOREM 4.3</u>: The random variable w has the rectangular distribution if, and only if, $c = d$.

PROOF: In Corollary 4.1, let c=d. Then

$$f(\theta) = \frac{\Delta}{(a-b)^2} = \frac{ad-bc}{(a-b)^2} = \frac{c}{a-b} \; .$$

COROLLARY 4.2: The random variable $V = \frac{ax+by}{x+y}$ is distributed uniformly over the range (b, a), $a > b$.

COROLLARY 4.3: The random variable $V = \frac{x-y}{x+y}$ is distributed uniformly over the range (-1,1).

COROLLARY 4.4: The random variable $V = \frac{x}{x+y}$ is distributed uniformly over the range (0,1).

This special case could also be obtained from a theorem on gamma distribution. "If x and y are independent Gamma variates with parameters ℓ and m respectively, the quotient x/x+y is a Beta variate of the first kind with parameters ℓ and m." Weatherburn [20, 153].

Thus if x and y are distributed as $dp = x^{\ell-1} e^{-x} dx$ and $dp = y^{m-1} e^{-y} dy$ respectively, then $\frac{x}{x+y}$ has the probability density function

$$\phi(\theta) = \frac{\theta^{\ell-1}(1-\theta)^{m-1}}{B(\ell, m)} ,$$

which gives us, on putting $\ell = m = 1$, the density function

$$\phi(\theta) = \frac{1}{B(1,1)} = 1.$$

We return to the general case, in which $c \neq d$.

It is advisable at this stage to show that the denominator does not vanish within the range, $\left(\frac{b}{d}, \frac{a}{c}\right)$.

$$(d-c)\theta + (a-b) = 0 \text{ for } \theta = \frac{a-b}{c-d} \; .$$

Then

$$\frac{a}{c} - \frac{a-b}{c-d} = \frac{ac-ad-ac+bc}{c(c-d)} = -\frac{\Delta}{c(c-d)},$$

$$\frac{b}{d} - \frac{a-b}{c-d} = \frac{bc-bd-ad+bd}{d(c-d)} = -\frac{\Delta}{d(c-d)}.$$

These differences are of the same sign (i.e., $\frac{a-b}{c-d}$ lies outside the interval $\left[\frac{a}{c}, \frac{b}{d}\right]$) if, and only if, c and d have the same sign. The restriction $cd > 0$ in the statement of Theorem 4.1 assures this.

We now transform our variable to $u = \theta + \frac{a-b}{d-c}$. The density function becomes

$$\phi(u) = \frac{\Delta}{(d-c)^2 u^2},$$

while the limits of u are

$$\frac{b}{d} + \frac{a-b}{d-c} = \frac{\Delta}{d(d-c)} \qquad \text{and} \qquad \frac{a}{c} + \frac{a-b}{d-c} = \frac{\Delta}{c(d-c)}.$$

$$\phi\left(\frac{\Delta}{d(d-c)}\right) = \frac{d^2}{\Delta}; \quad \phi\left(\frac{\Delta}{c(d-c)}\right) = \frac{c^2}{\Delta}.$$

Thus, if $d > c$, the frequency curve is a portion of the right half of the curve

$$X = \frac{k}{Y^2}, \qquad \phi\left(\theta = \frac{a}{c}\right) < \phi\left(\theta = \frac{b}{d}\right),$$

while, if $d < c$, we have a portion of the left half of the curve.

It follows that the frequency curve is either monotone increasing or monotone decreasing as the case may be, and has no mode.

The moments of the distribution u are readily calculated. We have

$$\mu_1' = \frac{\Delta}{(d-c)^2}\Big[\log u\Big]_{\frac{\Delta}{d(d-c)}}^{\frac{\Delta}{c(d-c)}}$$

$$= \frac{\Delta}{(d-c)^2} \log \frac{d}{c},$$

so that the mean of w is

$$\bar{w} = \frac{\Delta}{(d-c)^2} \log \frac{d}{c} - \frac{a-b}{d-c};$$

also

$$\mu_2^1 = \frac{\Delta}{(d-c)^2}\left(\frac{\Delta}{c(d-c)} - \frac{\Delta}{d(d-c)}\right)$$

$$= \frac{\Delta^2}{cd(d-c)^2},$$

so that

$$\text{var}(w) = \frac{\Delta^2}{(d-c)^2}\left[\frac{1}{cd} - \frac{\left(\log \frac{c}{d}\right)^2}{(d-c)^2}\right].$$

Furthermore, for every $n \geq 2$

$$\mu_n^1 = \frac{\Delta}{(d-c)^2} \cdot \frac{1}{n-1}\left\{\frac{\Delta^{n-1}}{c^{n-1}(d-c)^{n-1}} - \frac{\Delta^{n-1}}{d^{n-1}(d-c)^{n-1}}\right\}$$

$$= \frac{\Delta^n}{(n-1)(d-c)^{n+1}} \cdot \frac{d^{n-1}-c^{n-1}}{(cd)^{n-1}}.$$

Substituting $a = c = 1$, and $b = d = k$ the following theorem is immediately obtained.

<u>THEOREM 4.4</u>: The distribution of the random variable

$$w = \frac{x-ky}{x+ky}$$

has mean

$$\bar{w} = \frac{1}{(k-1)^2}\left(2k \log k - (k^2-1)\right)$$

and variance

$$\frac{4k^2}{(k-1)^2}\left(\frac{1}{k}-\frac{(\log k)^2}{(k-1)^2}\right).$$

The distribution function in this case is

(7) $$f(\theta) = \frac{\theta(1+k)}{\left[(k-1)\theta + 1 + k\right]^2}.$$

It is clear that as $k \to 1$, this distribution function converges uniformly to the rectangular $(-1,1)$, so that the moments also converge to those of the rectangular distribution. For the rectangular distribution $(-1,1)$ the mean is 0 and the variance is $\frac{1}{3}$.

So far we have made the restriction that $cd > 0$; when either c or d is zero, one end point of the range $\left(\frac{b}{d}, \frac{a}{c}\right)$ becomes infinite.

Let us consider what happens when $c = 0$. We have the variable

$$V = \frac{ax+by}{dy}.$$

We may proceed exactly as in Theorem 4.1. It follows that the distribution function is

$$f(\theta) = \frac{ad}{\{d\theta - (a-b)\}^2},$$

while the mean and variance are both infinite.

In particular, if $a = d = 1$ and $b = 0$, we have the following theorem.

THEOREM 4.5: The random variable $V = \frac{x}{y}$ has the distribution function $f(\theta) = \frac{1}{(1+\theta)^2}$ over the range $(0, \infty)$.

This result could also be obtained from a general theorem on gamma variates, Weatherburn [20, 158], which states that the quotient of two independent gamma variates with parameters ℓ and m is a $B_2(\ell, m)$ variate.

Thus, if x and y have the distributions $dp = x^{\ell-1}e^{-x}dx$ and $dp = y^{m-1}e^{-y}dy$ respectively, then $\frac{x}{y}$ has the probability density function

$$\phi(\theta) = \frac{\theta^{\ell-1}}{B(\ell,m)(1+\theta)^{\ell+m}},$$

which gives us, on putting $\ell = m = 1$, the density function

$$\phi(\theta) = \frac{1}{(1+\theta)^2}.$$

In the case where $cd < 0$, our proof of Theorem 4.1 is no longer valid.

Consider the case in which $c > 0$, $d < 0$ and $\frac{a}{c} > \frac{b}{d}$. Then $\frac{ad}{c} < b$ and $a > \frac{bc}{d}$.

So that, if $cx + dy > 0$,

$$\frac{ax+by}{cx+dy} > \frac{ax + \frac{ady}{c}}{cx+dy} = \frac{a}{c}$$

and, if $cx + dy < 0$,

$$\frac{ax+by}{cx+dy} < \frac{\frac{bcx}{d} + by}{cx+dy} = \frac{b}{d}.$$

Thus the values of the variate lie outside the interval $\left(\frac{b}{d}, \frac{a}{c}\right)$.

We shall now give a proof of Theorem 4.2 for this case.

PROOF: We calculate the cumulative distribution function. There are two cases to be considered:

(i)

$$w > \theta > \frac{a}{c}.$$

This requires $cx + dy > 0$, i.e., $x > \frac{-dy}{c}$

and

$$\frac{ax+by}{cx+dy} > \theta,$$

i.e.,

$$ax+by > \theta(cx+dy),$$

giving

$$x < \frac{b-d\theta}{c\theta-a}.$$

Then

$$P(w > \theta \geq \frac{a}{c}) = \int_{y=0}^{y=\infty} \int_{x=-\frac{d}{c}y}^{x=\frac{(b-d\theta)y}{c\theta-a}} \lambda e^{-\lambda y} \cdot \lambda e^{-\lambda x} dxdy$$

$$= \int_{y=0}^{\infty} \lambda e^{-\lambda y} \left[-e^{-\lambda y \frac{b-d\theta}{c\theta-a}} + e^{-\lambda y \left(\frac{-d}{c}\right)} \right] dy$$

$$= \int_{y=0}^{\infty} -\lambda e^{-\lambda y \frac{c\theta-a+b-d\theta}{c\theta-a}} + \lambda e^{-\lambda y\left(\frac{c-d}{c}\right)} dy$$

$$= -\frac{c\theta-a}{c\theta-a+b-d\theta} + \frac{c}{c-d}.$$

Then

$$F_1(\theta) = P(\frac{a}{c} \leq w < \theta)$$

$$= 1 + \frac{c\theta - a}{(c-d)\theta - (a-b)} - \frac{c}{c-d}$$

$$= -\frac{d}{c-d} + \frac{c\theta - a}{(c-d)\theta - (a-b)} .$$

(ii)

$$w < \theta < \frac{b}{d} .$$

This requires $cx + dy < 0$, i.e., $x < \frac{-d}{c} y$

and

$$\frac{ax+by}{cx+dy} < \theta ,$$

i.e.,

$$ax + by > (cx + dy)\theta ,$$

giving

$$x > \frac{d\theta - b}{a - c\theta} y .$$

Then

$$P(w < \theta < \frac{b}{d}) = \int_{y=0}^{\infty} \lambda e^{-\lambda y} \int_{x=\frac{d\theta - b}{a-c\theta}}^{-\frac{d}{c} y} \lambda e^{-\lambda x} dxdy$$

$$= \int_{0}^{\infty} \lambda e^{-\lambda y} \left[-e^{-\lambda y(-\frac{d}{c})} + e^{-\lambda y \frac{d\theta - b}{a-c\theta}} \right] dy$$

$$= \int_{0}^{\infty} -\lambda e^{-\lambda y\left(\frac{c-d}{c}\right)} + \lambda e^{-\lambda y \left(\frac{a-c\theta + d\theta - b}{a-c\theta}\right)} dy$$

$$= -\frac{c}{c-d} + \frac{a-c\theta}{a-c\theta + d\theta - b}$$

$$= -\frac{c}{c-d} + \frac{c\theta - a}{(c-d)\theta - (a-b)} = F_2(\theta),$$

which proves the theorem.

We note that

$$F_1(\tfrac{a}{c}) = -\frac{d}{c-d} = 1 - \frac{c}{c-d} = F_2(\tfrac{b}{d}).$$

The cumulative distribution function is thus given by

$$\begin{aligned} F(\theta) &= F_2(\theta), & \theta &\leqq \tfrac{b}{d}, \\ &= \frac{-d}{c-d}, & \tfrac{b}{d} &\leqq \theta \leqq \tfrac{a}{c}, \\ &= F_1(\theta) & \tfrac{a}{c} &\leqq \theta . \end{aligned}$$

Differentiating, we have the density function

$$\begin{aligned} f(\theta) &= \frac{-\Delta}{\{(c-d)\theta - (a-b)\}^2}, & \theta &\leqq \tfrac{b}{d} \text{ or } \theta \geqq \tfrac{a}{c}, \\ &= 0, & \tfrac{b}{d} &< \theta < \tfrac{a}{c}. \end{aligned}$$

We now consider the use of the statistic w to test the hypothesis that a process is actually Poissonian.

From a single observation of the statistic we are led to a crude, but very simple test of significance, which is independent of k.

<u>THEOREM 4.6:</u> Let w_η be the value of $w = \frac{x-ky}{x+ky}$, for which $P(w < w_\eta) = \eta$, then $w > w_\eta$ if, and only if, $x > \frac{\eta}{1-\eta} \cdot y$, $0 \leqq \eta \leqq 1$.

<u>PROOF</u>: For the variate $w = \frac{x-ky}{x+ky}$

we have

$$F(\theta) = \frac{k(1+\theta)}{(k-1)\theta + k + 1},$$

so that

$$\frac{k(1+w_\eta)}{(k-1)\,w_\eta + (k+1)} = \eta ,$$

i.e.,

$$k + kw_\eta = (k-1)\eta w_\eta + (k+1)\eta ,$$

giving

$$w_\eta = \frac{(k+1)\eta - k}{k - (k-1)\eta} .$$

Then $w > w_\eta$ if, and only if,

$$\frac{x-ky}{x+ky} > \frac{(k+1)\eta - k}{k-(k-1)\eta},$$

i.e.,

$$xk - x(k-1)\eta - k^2 y + k(k-1)\eta\, y >$$
$$x(k+1)\eta - xk + k(k+1)\eta\, y - k^2 y,$$

so that

$$x(k - (k-1)\eta + k + (k+1)\eta) > y(k(k+1)\eta - k(k-1)\eta),$$
$$x(2k - 2k\eta) > y \cdot 2k ,$$

giving

$$x > \frac{\eta}{1-\eta}\, y, \text{ which is independent of } k.$$

It follows, therefore, that if a sample of two intervals is taken from a process, a simple test procedure for the hypothesis that the process is Poissonian is to reject the hypothesis if

$$x > \frac{1-\xi}{\xi}\, y, \qquad (\xi = 1 - \eta)$$

where ξ is the desired level of significance.

If ξ is 0.10, this leads to $x > 9y$. For $\xi = 0.05$, we have $x > 19y$, while for $\xi = 0.01$, we have $x > 99y$.

We proceed now to a less crude test of significance.

It has been shown by Gorsline [10] that for random variables having the rectangular distribution the averages of sums of these variables approach normality very quickly; indeed, he has shown that for an average of as few as four observations the normal approximation is very good.

We therefore consider the statistic

$$W_n = \frac{1}{n}\left\{\frac{x_1-y_1}{x_1+y_1} + \frac{x_2-y_2}{x_2+y_2} + \dots + \frac{x_n-y_n}{x_n+y_n}\right\},$$

where the x_i and y_i are drawn at random from the population of intervals between events, with no interval appearing more than once.

<u>THEOREM 4.7</u>: For every n the statistic W_n is a test for the hypothesis that a process is Poissonian with the 0.05 probability level given approximately by $|W_n| > \frac{1.155}{\sqrt{n}}$ and the 0.01 level by $|W_n| > \frac{1.443}{\sqrt{n}}$.

<u>PROOF</u>: If the hypothesis that the population of intervals is exponential, i.e., that the process is Poissonian, is true, then W_n is asymptotically normally distributed with mean 0 and variance $\frac{1}{3n}$.

The standard deviation of the distribution is thus $\frac{1}{\sqrt{3n}}$. The 0.05 significance level is given approximately by a 2σ deviation, and so we would reject the hypothesis if $|W_n| > \frac{2}{\sqrt{3n}} = \frac{1.155}{\sqrt{n}}$.

Similarly, taking 2.5σ for the approximate 0.01 level, we have the criterion $|W_n| > \frac{1.443}{\sqrt{n}}$.

Conversely, we may calculate the sample size required in order for $|W_n| < Z$ to be the region of acceptance for the hypothesis by the formulae $n = \frac{4}{3Z^2}$ at the 0.05 level, and $n = \frac{25}{12Z^2}$ at the 0.01 level.

CHAPTER V

THE PROBLEM OF DISCRIMINATION BETWEEN FURRY AND POISSON PROCESSES

In the previous chapter we have developed a test of significance for the hypothesis that a given process is Poissonian; we now proceed in the same way to consider the Furry process.

THEOREM 5.1: For the Furry process, if x is the duration of the m^{th} interval (i.e., the interval in which the population size passes from m to $m + 1$), and y is the duration of the n^{th} interval, then the distribution of the random variable

$$w = \frac{ax + by}{cx + dy}, \qquad \text{where } cd > 0$$

is independent of λ.

PROOF: This is similar to the proof of Theorem 4.2.

From the proof of Lemma 4.1, we see that the range is again $\left[\frac{b}{d}, \frac{a}{c}\right]$. We shall again take $c, d > 0$ and $\frac{a}{c} > \frac{b}{d}$.

We have

$$P(w > \theta) = P\left(x > \frac{d\theta - b}{a - c\theta}\, y\right),$$

i.e.,

$$P(w > \theta) = \int_{y=0}^{\infty} \lambda n \int_{x=\frac{d\theta - b}{a - c\theta} y}^{\infty} \lambda m e^{-\lambda m x} \cdot e^{-\lambda n y}\, dx dy$$

$$= \int_{y=0}^{\infty} \lambda\, n e^{-ny} \left[-e^{-\lambda m x} \right]_{\frac{d\theta-b}{a-c\theta} y}^{\infty} dy$$

$$= \int_{y=0}^{\infty} \lambda\, n e^{-\lambda y \left[n - m\left(\frac{d\theta-b}{a-c\theta}\right)\right]} dy$$

$$= \frac{n}{n + m\,\frac{d\theta-b}{a-c\theta}}$$

$$= \frac{n(a-c\theta)}{n(a-c\theta) + m(d\theta-b)} ,$$

so that

$$P(w \leq \theta) = 1 - P(w > \theta)$$

$$= \frac{m(d\theta-b)}{(md-nc)\theta + (na-mb)} = F(\theta) .$$

Thus $F(\theta)$ is independent of λ and the theorem is proved.

This is the same function as was obtained for the Poisson process with the substitution of na, mb, nc and md for a, b, c and d. (5). It follows immediately that the mean of the distribution for the Furry process is given by

$$\frac{na-mb}{nc-md} + \frac{\Delta nm}{(md-nc)^2} \log \frac{md}{nc} .$$

The variance is

$$\frac{n^2 m^2 \Delta^2}{(md-nc)^2} \left(\frac{1}{nmcd} - \frac{\left(\log \frac{nc}{md}\right)^2}{(md-nc)^2} \right) .$$

The frequency distribution is given by

$$f(\theta) = \frac{mn\Delta}{\{(md-nc)\theta + (na-mb)\}^2} \ .$$

<u>COROLLARY 5.1</u>: For the above Furry process the random variable $w = \frac{x-y}{x+y}$ has density function

$$f(\theta) = \frac{2mn}{\{(m-n)\theta + (m+n)\}^2} \ ,$$

the mean is

$$\frac{n+m}{n-m} + \frac{2mn}{(n-m)^2} \log \frac{m}{n} \ ,$$

and the variance is

$$\frac{4m^2n^2}{(n-m)^2} \left[\frac{1}{mn} - \frac{\left(\log \frac{n}{m}\right)^2}{(n-m)^2} \right] .$$

These results follow immediately from those above by putting $a = c = d = 1$ and $b = -1$.

We now consider the above random variable with the restriction $n = km$, and note that the mean, the variance and the density function depend only on the ratio $k = \frac{n}{m}$.

We have

$$f(\theta) = \frac{2k}{((1-k)\theta + (1+k))^2} \ , \qquad -1 \leqq \theta \leqq 1,$$

the mean is

$$\frac{k+1}{k-1} - \frac{2k}{(k-1)^2} \log k = \mu[k] \ ,$$

the variance is

$$\frac{4k^2}{(k-1)^2} \left[\frac{1}{k} - \frac{(\log k)^2}{(k-1)^2} \right] = \sigma_k^2 \ .$$

These formulae are similar to, but not identical with, those obtained in Theorem 4.3. (They are actually the same if $\frac{1}{k}$ is substituted for k.)

THEOREM 5.2: For $k > 1$, $\mu_{[k]}$ is a strictly increasing function of k.

PROOF:
$$\mu_{[k]} = \frac{k^2 - 2k \log k - 1}{(k-1)^2} \; .$$

$$\frac{d\mu_{[k]}}{dk} = \frac{2(k-1)\left[k - 1 - \log k\right] - 2\left[k^2 - 2k \log k - 1\right]}{(k-1)^3}$$

$$= \frac{2}{(k-1)^3}\left(k^2 - k - k \log k - k + 1 + \log k - k^2 + 2k \log k + 1\right)$$

$$= \frac{2}{(k-1)^3}\left(k \log k - 2k + 2 + \log k\right) \; .$$

The denominator is positive. We complete the proof of the theorem by showing that the numerator $N(k) = k \log k - 2k + 2 + \log k$ is positive.

$$\frac{d\,N(k)}{dk} = 1 + \log k - 2 + \frac{1}{k} = \frac{k \log k - k + 1}{k} \; .$$

But for $k > 1$, $k \log k - k + 1$ is an increasing function, its derivative being $\log k$.

When $k = 1$, we have $k \log k - k + 1 = 0$, so that for $k > 1$ we have $\frac{d\,N(k)}{dk} > 0$. Thus, since $N(0) = 0$, $N(k) > 0$ for all $k > 1$ and the theorem is proved.

Since $\mu_{[1]} = 0$, it follows that $\mu_{[k]} > 0$ for all $k > 1$. This leads us to the following theorem.

THEOREM 5.3: If $(m_1, n_1), (m_2, n_2) \ldots (m_i, n_i), \ldots$ is a sequence of pairs of positive integers such that $k_i = \frac{m_i}{n_i} > k > 1$ for all i, and if x_{m_i}, y_{n_i} denote the lengths of the m_i th and n_i th intervals respectively, then the statistic

$$W_n = \frac{1}{N} \sum_{i=1}^{i=N} \frac{x_{m_i} - y_{n_i}}{x_{m_i} + y_{n_i}}$$

is asymptotically normally distributed with positive mean, provided that the series

$$\sum_i \sigma_{k_i}^2 = \sum_i \frac{4k_i^2}{(k_i-1)^2} \left(\frac{1}{k_i} - \frac{(\log k_i)^2}{(k_i-1)^2} \right)$$

diverges.

PROOF: W_n is the average of a sum of statistics of the form $\frac{x-y}{x+y}$ which are uniformly bounded and such that the sum of the variances diverges. Hence, by one of the forms of the central limit theorem, Feller [8, 203], the distribution is asymptotically normal, with mean

$$\bar{W}_n = \frac{1}{N} \sum_{i=1}^{i=N} \mu_{[k_i]} .$$

But

$$\mu_{[k_i]} \geqq \mu_{[k]} \quad \text{for all } k_i, \text{ so that} \quad \bar{W}_N \geqq \mu_{[k]}.$$

THEOREM 5.4: If, in Theorem 5.3, $k_i = k > 1$ for all i, then the statistic

$$W_N = \frac{1}{N} \sum_{i=1}^{i=N} \frac{x_{m_i} - y_{n_i}}{x_{m_i} + y_{n_i}}$$

is asymptotically normally distributed with mean

$$\bar{W}_N = \mu_{[k]} \quad \text{and variance,} \quad \frac{\sigma_k^2}{n} .$$

PROOF: The sequence

$$n \left[\frac{4k^2}{(k-1)^2} \left(\frac{1}{k} - \frac{(\log k)^2}{(k-1)^2} \right) \right] \qquad \text{diverges.}$$

We shall now make use of these theorems in the formulation of a test for discrimination between Poisson and Furry processes.

Consider the statistic

$$W_n = \frac{1}{n} \sum_i \frac{x_i - y_{ki}}{x_i + y_{ki}} ,$$

where no interval is used more than once and where k is an integer; for example, if $k = 3$, the sequence might include

$$\frac{x_1 - y_3}{x_1 + y_3} , \frac{x_2 - y_6}{x_2 + y_6} , \frac{x_4 - y_{12}}{x_4 + y_{12}} , \ldots$$

The third interval, having already been used as a 'y' could not be regarded as an 'x'.

THEOREM 5.5: For every n the statistic W_n is a test for discrimination between the hypothesis that a process is Poissonian, and the hypothesis that the process is a Furry process.

PROOF: If the process is actually Poissonian W_n is distributed asymptotically normally with mean 0 and variance $\frac{1}{3n}$.

For a Furry process W_n is asymptotically normal with mean $\mu_{[k]}$ and variance $\frac{\sigma_k^2}{n}$.

We discriminate as follows. Let n be chosen so that

$\mu_{[k]} = \frac{2}{\sqrt{n}}\left(\frac{1}{\sqrt{3}} + \sigma_k\right)$. The critical value of W_n is then $W_n = \frac{2}{\sqrt{3n}}$.

If $W_n < \frac{2}{\sqrt{3n}}$, we accept the Poissonian hypothesis. If $W_n > \frac{2}{\sqrt{3n}}$, we accept the Furry hypothesis. The probability that W_n assumes a value greater than $\frac{2}{\sqrt{3n}}$, if the hypothesis that the process is Poissonian is true, is approximately 0.025; the probability of a value less than $\frac{2}{\sqrt{3n}}$ for a Furry process is also 0.025. Thus 0.025 is the probability of an error of either the first or second kind in the terminology of Neyman.

The 2.5 σ limit, for which n is chosen so that

$\mu_{[k]} = \frac{2.5}{\sqrt{n}}\left(\frac{1}{\sqrt{3}} + \sigma_k\right)$, gives errors of 0.006.

In the following table values are given, for various k, of the mean $\mu_{[k]}$, and variance σ_k^2 of the variate $\frac{x_i - y_{ki}}{x_i + y_{ki}}$ for the Furry process. We also give values of n_k, the number of ratios $\frac{x_i - y_{ki}}{x_i + y_{ki}}$ needed for W_n to discriminate at the 0.05 level of significance, of

W_n^* the critical value of W_n, and of N_k, where N_k is the number of intervals required to give the number of n_k of ratios.

k	$\mu_{[k]}$	σ_k^2	n_k	W_n^*	N_k
2	.2274	.3127	100	.1155	398
3	.3521	.2844	40	.1826	177
4	.4344	.2593	25	.2310	132
5	.4941	.2382	19	.2653	115
6	.5400	.2205	16	.2888	114
11	.6725	.1616	9	.3850	99

Thus, for example, if k is 5, then $n_k = 19$, $N_k = 115$ so that

$$W_n = \frac{1}{19}\left(\frac{x_1-y_5}{x_1+y_5} + \dots + \frac{x_{23}-y_{115}}{x_{23}+y_{115}}\right).$$

We take 19 pairs of observations, which requires that 115 events occur.

We accept the Poisson hypothesis if $W_n < 0.2653$; we accept the Furry hypothesis if $W_n > 0.2653$.

The difficulty occurring in the calculation of N_k is indicated in the exmple for $k = 3$ given immediately above Theorem 5.5.

We may, however, calculate N_k by the following formulae:

(i) If $n_k \equiv 0(\text{mod } (k-1))$,

$$N_k = k\left(\frac{kn_k}{k-1} - 1\right),$$

(ii) If $n_k \equiv p(\text{mod } (k-1))$, $p \neq 0$,

$$N_k = k\left(k\left[\frac{n_k}{k-1}\right] + p\right),$$

where $\left[\frac{n_k}{k-1}\right]$ denotes the greatest integer not exceeding $\frac{n_k}{k-1}$.

LIST OF REFERENCES

1. N. Arley and V. Borchsenius, On the theory of infinite systems of differential equations and their application to the theory of stochastic processes and the perturbation theory of quantum mechanics, Acta Math., vol. 76 (1945), pp. 261-322 (esp. pp. 298-9).

2. G. A. Barnard, Time intervals between accidents--a note on Maguire, Pearson and Wynn's paper, Biometrika, vol. 40 (1953), pp. 212-3.

3. H. Cramér, Mathematical Methods of Statistics, Princeton University Press (1946).

4. J. L. Doob, Markoff chains - denumerable case, Trans. Amer. Math. Soc., vol. 58 (1945), pp. 455-73.

5. W. Feller, Die Grundlagen ver Volterraschen Theorie des Kampfes ums dasein in wahrscheinlichkeitstheoretischer Behandlung, Acta Biotheoretica, vol. 5 (1939), pp. 11-40, (Abstract in Math. Reviews, vol. 1 (1940), p. 22.)

6. W. Feller, On the Kolmogorov-Smirnov limit theorems for empirical distributions, Annals of Math. Stat., vol. 19 (1948), pp. 177-89.

7. W. Feller, On the integro-differential equations of completely discontinuous Markoff processes, Trans. Amer. Math. Soc., vol. 48 (1940), pp. 488-515.

8. W. Feller, Probability Theory and Its Applications, vol. 1, Wiley (1950).

9. W. H. Furry, On fluctuation phenomena in the passage of high energy electrons through lead, Phys. Rev., vol. 52 (1937), pp. 569-81.

10. D. A. Gorsline, Calculation of composite probability density functions (unpublished M. A. thesis, University of Oklahoma, 1950), chapter III, p. 26 et seq.

11. E. R. Immel, Problems of estimation and of hypothesis testing connected with birth and death Markov processes (unpublished Ph. D. dissertation, University of California, Los Angeles, 1951).

12. D. G. Kendall, On the generalized birth and death process, Annals of Math. Stat., vol. 19 (1948), pp. 1-15.

13. D. G. Kendall, Stochastic processes and population growth, Roy. Stat. Soc. Jour., B. vol. 11 (1949), pp. 230-64.

14. A. Kolmogoroff, Sulla determinatzione empirica di una legge di distribuzione, Ist. Ital. Attuari, Giorn., vol. 4 (1933), pp. 1-11.

15. W. Ledermann and G. E. H. Reuter, Spectral theory for the differential equations of simple birth and death processes, Phil. Trans. Roy. Soc. Lond., Series A, vol. 246 (12 Jan., 1954), pp. 321-69.

16. O. Lundberg, On random processes and their application to sickness and accident statistics (University of Stockholm thesis, Uppsala, 1940). (Abstract in Math. Reviews, vol. 2 (1941), p. 230).

17. B. A. Maguire, E. S. Pearson and A. H. A. Wynn, The time intervals between industrial accidents, Biometrika, vol. 39 (1952), pp. [illegible]8-80.

18. P. A. P. Moran, Estimation methods for evolutive processes, Roy. Stat. Soc. Jour., B, vol. 13 (1951), pp. 141-46.

19. G. E. H. Reuter and W. Ledermann, On the differential equations for the transition probabilities of Markov processes with enumerably many states, Proc., Camb. Phil. Soc., vol. 49 (1953), pp. 247-62.

20. C. E. Weatherburn, A First Course in Mathematical Statistics, Cambridge University Press (1947).

21. G. Udny Yule, A mathematical theory of evolution based on the conclusions of Dr. J. C. Willis, F. R. S., Phil. Trans. Roy. Soc. Lond., Series B, vol. 213 (1924), pp. 21-87.