ANALYSIS OF POINT PROCESSES

By

ALBERT EDWARD JONES, B.Sc.

A Thesis

Submitted to the Faculty of Graduate Studies

in Partial Fulfilment of the Requirements

for the degree

Master of Science

McMaster University

October, 1967

MASTER OF SCIENCE (1967)
(Mathematics)

McMaster University
Hamilton, Ontario

TITLE: Analysis of Point Processes

AUTHOR: Albert Edward Jones, B.Sc. (McMaster University)

SUPERVISOR: Dr. I. Z. Chorneyko

NUMBER OF PAGES: viii, 68

SCOPE AND CONTENTS: This thesis is concerned with investigating point processes and the numerous methods available for obtaining information about them.

PREFACE

Chapter I of this thesis introduces the definition of point processes in general and, in particular, Poisson processes. The Poisson processes is covered extensively with a test for randomness, various estimates for the expected rate of occurrence and techniques developed to deal with more than one process.

In the second chapter a process generated by superimposing periodic sequences of events is considered. Methods of determining the number of sources and the assigning of each event to its proper source are also investigated. The asymptotic properties of the pooled processes are also shown.

Non-periodic processes are considered in chapter III with particular emphasis given to the renewal process and the pooled outputs of renewal processes.

Graphical methods are outlined in chapter IV and the logarithmic transformation is considered.

In chapter V the correlation function is defined and the intensity function for any point process is estimated. The properties of the intensity function for the Poisson process are developed in some detail.

Spectral analysis of point processes is considered in chapter VI. Stationary processes are used to illustrate the properties

of the spectral distribution function.

Various estimates of the spectral density function are introduced and some rules for proceeding to estimate the spectrum from an observed sample function are considered.

ACKNOWLEDGMENTS

The author expresses his deepest appreciation to Dr. I. Z. Chorneyko of the Department of Mathematics for his encouragement and guidance in the writing of this thesis.

In addition, for the financial assistance received and the facilities made available, gratitude is expressed to the National Research Council, Ottawa, and to McMaster University, Hamilton.

TABLE OF CONTENTS

CHAPTER I

Point Processes and Poisson Processes

1.1 Introduction

Many stochastic processes have to deal with events which occur haphazardly in space or time. For example, there are many problems in physics and biology where we have to deal with particles or individuals distributed over an "infinity" of states. The events themselves may be electrical pulses in nerve fibers (the study of which motivated this work), fission in bacteria or in the nucleus, customers arriving at a queueing point, etc. The requirements for such a process are a continuum, space or time, which may or may not be one-dimensional, and a series of events which can be regarded as points in this continuum. Often each event will have some numerical quantity attached to it, which quantity, however, we shall ignore. We are primarily interested in the patterns or non-patterns exhibited by the points in the continuum. Such stochastic processes are called point processes.

One of the most important of the point processes is the so-called Poisson process. This process arises naturally in the study of queues, in the study of radioactive decay and in many other physical situations. In this chapter, we shall provide a detailed study of the methods used to analyze Poisson processes.

1.2 Stochastic Processes and the Poisson Process

Definition 1.2.1 A stochastic process is defined as a collection (ensemble, family) $\{X(t),\ t \in T\}$ of random variables, where t is a parameter running over a suitable index set T.

For our purposes, we shall consider a stochastic process completely specified once we prescribe:

(1) the state space, i.e., the space in which the possible values of each X(t) lie,

(2) the index set T,

(3) the family of joint distributions for any arbitrary finite set of possible values t_r (r=1,2, ... , n) of the parameter t.

Definition 1.2.2 Let $T = [0, \infty)$.

Let X(t) = number of specified events occurring during the time period from 0 to t.

Then $\{X(t),\ t \in T\}$ is said to be a point process.

It shall be convenient at some time to take in definition 1.2.2, $T = (-\infty, \infty)$. This enables us to study past history, if we so desire. In either case, we shall call the family $\{X(t),\ t \in T\}$ a point process. It should be pointed out that there are other ways of specifying a point process and we shall do so, when it is convenient.

Definition 1.2.3 Let $\{X(t),\ t \in T\}$ be a point process. Further, suppose the family $\{X(t),\ t \in T\}$ satisfies the following requirements:

(1) $X(0) = 0$

(2) Let $t_0 < t_1 \ldots < t_n$. Then:

$X(t_1) - X(t_0)$, $X(t_2) - X(t_1)$, . . ., $X(t_n) - X(t_{n-1})$ are mutually

independent random variables.

(3) The random variable $X(t_0 + t) - X(t_0)$ depends only on t and not on t_0, or on the value of $X(t_0)$.

(4) The probability of at least one event happening in a time period of duration h is

$$\Pr[\text{event in time } h] = \lambda h + o(h),\ h \to 0,\ \lambda > 0.$$

(5) The probability of 2 or more events happening in time h is o(h).

Then $\{X(t),\ t \in T\}$ is said to be a Poisson process.

Let $P_n(t) = \Pr[X(t) = n]$, $n = 0,1,2 \ldots$

Then, it can be shown [Parzen (29)] that

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

i.e., $P_n(t)$ has the Poisson distribution with parameter λt.

An equivalent interpretation of the Poisson process is one in which the n events are assumed to be uniformly and independently distributed between 0 and t and that the probability of two or more events occurring in the same interval of very short duration is negligible. Such a series of events is said to be a random series of events and $P_n(t) = \frac{(\lambda t)^n \exp(-\lambda t)}{n!}$. λt is the mean number of occurrences in time t.

Another way of specifying a point process is to find the distribution of the intervals between events. More specifically, let $\{X(t),\ t \in T\}$ be a Poisson process. Let Y be a random variable, the values of which are the lengths of the intervals between successive events in this process.

Now, since $\{X(t),\ t \in T\}$ is a Poisson process, we know that $P_0(y) = \exp(-\lambda y)$ is the probability of no events occurring in a time interval of length y or $P_0(y)$ is the probability that the time interval between events exceeds y. If f(y) is the probability density function of the random variable Y, then

$$\int_y^\infty f(y')dy' = \exp(-\lambda y).$$

Thus, $f(y) = \lambda \exp(-\lambda y)$, which is the exponential density function. Thus, we have proved the following Theorem:

Theorem 1.2.1 Let $\{X(t),\ t \in T\}$ be a Poisson Process. Let Y be the random variable, the values of which are the lengths of intervals between successive events. Then the time intervals between events are independently distributed with the exponential density function, i.e.,

$$f(y) = \lambda \exp(-\lambda y).$$

All statistical questions involving Poisson processes involve inferences about λ. Inferences about λ, of course, can be made from either of the characterizing distributions. In the next sections, we shall describe some of the procedures involved in making these inferences.

1.3 A Test for Randomness

Perhaps the first thing one would like to know about a point process is whether it is a Poisson process or not, i.e., are the events in the time interval $[0,t]$ uniformly and independently distributed. We shall assume that the process is homogeneous, i.e., that the average rate of occurrence per unit of time is independent of the time.

Let X be a random variable with a continuous distribution function $F(x) = Pr(X \leq x)$. Let $X_1, X_2, \ldots, X_n$ be a sample of size n ordered so that $X_1 \leq X_2 \ldots \leq X_n$.

Definition 1.3.1 Let

$$F_n(x) = 0 \quad \text{if } x < X_1$$
$$F_n(x) = {}^{k}/_{n} \quad \text{if } X_k \leq x < X_{k+1}$$
$$F_n(x) = 1 \quad \text{if } X_n \leq x.$$

$F_n(x)$ is called the empirical distribution function of the sample.

Definition 1.3.2 Let $D_n = \text{l.u.b}_x \left| F(x) - F_n(x) \right|$.

It has been shown (6) that

$$\lim_{n \to \infty} Pr\left[D_n < {}^{z}/_{n} \right] = 1 - 2 \sum_{j=1}^{\infty} (-1)^j \exp(-2j^2 z^2).$$

Birnbaum (6) has calulated the probability distribution of D_n for small values of n. These tables can be used to test for the randomness of a time series. If the series is random then, the n_t observations should be independently and uniformly distributed over the observation period $[0,T]$. In that case, we have that $nF(t) = \frac{nt}{T}$ and $nF_n(t) = n_t$. Therefore,

$$nD_n = n \max_{0 < t < T} \left| F_n(t) - F(t) \right|$$
$$= \max_{0 < t < T} \left| n_t - \frac{nt}{T} \right|.$$

We then compare the value of nD_n with the appropriate values in the tables for the chosen significance level. We reject the hypothesis of randomness if this value of nD_n is larger than the value in the table.

1.4 Point Estimation of the Parameter λ in the Poisson Process

Once we have determined that the series is random with constant parameter λ, we can then proceed to the important question of estimating λ. There are two methods for estimating λ, both of which are well known. For the sake of completeness, we list these methods very briefly.

(A) Maximum Likelihood Estimator. This estimator $\hat{\lambda} = \frac{n_t}{t}$ is derived in the standard way and has all the properties of Maximum Likelihood Estimators.

(B) Suppose the Poisson process is observed until k events have occurred. Let $X_1, X_2, \ldots, X_k$ be the random variables representing the interval lengths between events. Then, the likelihood function is

$$L(x_1, \ldots, x_n) = (\tfrac{1}{\gamma})^k \exp\left[-\frac{1}{\gamma}\sum_{i=1}^{K} x_i\right], \gamma = \frac{1}{\lambda}.$$

$$\text{Then, } \frac{\partial \ln L}{\partial \gamma} = -\frac{k}{\gamma} + \frac{\sum x_i}{\gamma^2}.$$

Thus, the maximum likelihood estimate of γ is $\frac{\sum x_i}{k}$ and has variance $\frac{k}{\gamma^2}$.

1.5 Confidence Intervals for λ

If one observes a Poisson process for a predetermined observation time t, then the number of occurrences n_t can be used to form point estimates and confidence intervals for λ, using the fact that n_t is Poisson with mean λt. On the other hand, if the observation of the Poisson process is continued until a fixed number of events m

have been counted then the amount, T_m, of observation time required can be used to form confidence intervals for λ. Since T_m has the gamma distribution with parameters m and λ, $2\lambda T_m$ has the χ^2 distribution with 2m degrees of freedom. Let C and D be such values that, if Z has a χ^2 distribution with 2m degrees of freedom, then $\Pr(Z < C) = \Pr(Z > D) = \alpha/2$. Then $1 - \alpha = \Pr(C \leq 2\lambda T_m \leq D) = \Pr(\frac{C}{2T_m} \leq \lambda \leq \frac{D}{2T_m})$. Consequently $\left[\frac{C}{2T_m}, \frac{D}{2T_m}\right]$ is a confidence interval for λ with confidence coefficient $1-\alpha$.

1.6 Approximate Confidence Interval for λ

If t and n are large, so that the number of intervals between events is large, then we can derive approximate confidence intervals from the large sample theory of maximum likelihood estimators.

Let $X_1, X_2, \ldots, X_k$ be random variables representing the interval lengths from a series generated by a Poisson process with parameter λ. Let $\gamma = \frac{1}{\lambda}$.

Then $$L(x_1, x_2, \ldots, x_k) = \left(\frac{1}{\gamma}\right)^k \exp\left[-\frac{1}{\gamma}\sum x_i\right].$$

Now

$$\frac{\partial \ln L}{\partial \gamma} = -\frac{k}{\gamma} + \frac{\sum x_i}{\gamma^2}$$

and $$\hat{\gamma} = \frac{\sum x_i}{k}.$$

It has been shown in Mood (25) that the maximum likelihood estimator $\hat{\gamma}$ is, for large samples, approximately normally distributed with mean γ and variance $\frac{1}{k\sigma_\mu^2}$

where

$$\sigma_{\mu}^{2} = -E\left[\frac{\partial^2}{\partial \gamma^2}\left(\ln \frac{1}{\gamma} \exp(-\frac{x}{\gamma})\right)\right]$$
$$= -E\left[(\gamma^{-2} - 2x\gamma^{-3})\right]$$
$$= \gamma^{-2}.$$

Then

$$W = \frac{(\hat{\gamma} - \gamma)\sqrt{k}}{\gamma}$$
$$= (\hat{\gamma}\lambda - 1)\sqrt{k}$$

is approximately normally distributed with mean 0 and variance 1.

Therefore $\Pr\left[-Z_{\frac{\alpha}{2}} \leq (\hat{\gamma}\lambda - 1)\sqrt{k} \leq Z_{\frac{\alpha}{2}}\right] = 1-\alpha$

where $Z_{\frac{\alpha}{2}}$ is that point of the N(0,1) distribution such that

$$\int_{Z_{\frac{\alpha}{2}}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{Z^2}{2})dZ = \frac{\alpha}{2}.$$

By converting the above inequalities, we obtain for a $1 - \alpha$ confidence interval for λ

$$\left[\frac{-Z_{\frac{\alpha}{2}}\sqrt{k} + k}{\sum x_i}, \frac{Z_{\frac{\alpha}{2}}\sqrt{k} + k}{\sum x_i}\right].$$

1.7 Estimating λ with Prescribed Precision

We shall now obtain an estimate λ' of λ such that $|\lambda' - \lambda|$ is within bounds independent of t and n, i.e., of the observation time and the number of events observed. The procedure is as follows:

(1) Observe the process until n events are observed. Let the observation time be T_n.

(2) Choose $\alpha > 0$ and $\varepsilon > 0$.

(3) Let $C = \frac{\alpha \varepsilon^2}{2n}$.

(4) Perform the experiment for an additional $\frac{1}{2CT_n}$ units of time during which, say, $\bar{n}$ events are observed.

(5) Let $\lambda' = 2CT_n\bar{n}$.

Theorem 1.7.1 Let λ' be the estimate of λ obtained from the above procedure. Then

$$\Pr(|\lambda' - \lambda| \leq \varepsilon) \geq 1 - \alpha.$$

Proof:

$$E(\lambda') = 2C\, E(T_n\bar{n}),$$

$$f(T_n\bar{n}) = f(T_n)\, f(\bar{n} \mid T_n) \qquad \text{where}$$

$$f(T_n) = \frac{1}{\gamma^n \Gamma(n)} \exp(-\frac{T_n}{\gamma}) T_n^{n-1}, \quad (\gamma = \frac{1}{\lambda})$$

$$\text{and } f(\bar{n} \mid T_n) = \frac{\exp(-\frac{1}{\gamma 2CT_n})\,(\frac{1}{\gamma 2CT_n})^{\bar{n}}}{\bar{n}!} .$$

$$\therefore E(T_n\bar{n}) = \int_0^\infty \sum_{\bar{n}=0}^{\infty} T_n\bar{n}\, f(T_n)\, f(\bar{n} \mid T_n)\, dT_n$$

$$= \int_0^\infty T_n\, f(T_n)\, E(\bar{n} \mid T_n)\, dT_n$$

$$= E\left[T_n(E(\bar{n} \mid T_n))\right]$$

$$= \lambda/2C .$$

Also

$$E(T_n\bar{n})^2 = \int_0^\infty T_n^2\, f(T_n) \sum_{\bar{n}=0}^{\infty} \bar{n}^2\, f(\bar{n} \mid T_n)\, dT_n$$

$$= E\left[T_n^2\, E(\bar{n}^2 \mid T_n)\right].$$

$$\therefore \text{Var}(T_n\bar{n}) = E\left[T_n^2\, E(\bar{n}^2 \mid T_n)\right] - E^2(T_n E(\bar{n} \mid T_n)$$

$$= E\left[T_n^2\, \text{Var}(\bar{n} \mid T_n)\right] + E\left[T_n^2\, E^2(\bar{n} \mid T_n)\right] - E^2\left[T_n E(\bar{n} \mid T_n)\right]$$

$$= E\left[T_n^2\, \text{Var}(\bar{n} \mid T_n)\right] = {}^{n}/_{2C} .$$

$$\text{Hence} \qquad E(\lambda') = 2C\,\frac{\lambda}{2C} = \lambda$$

$$\text{and} \qquad \text{Var}(\lambda') = 4C^2\,\text{Var}(T_n\bar{n}) = 2nC .$$

By Tchebycheff's Inequality

$$\Pr\left(|\lambda' - \lambda| \leq \xi\right) \geq 1 - \frac{\text{Var}(\lambda')}{\xi^2} = 1 - \alpha .$$

Therefore, given the prescribed precision of the estimate ξ and the prescribed probability α, we can always find an unbiased estimator λ' within the required bounds.

1.8 The Sequential Probability Ratio Test for Poisson Processes

In this section we shall develop a sequential probability ratio test for testing

$$H_0 : \lambda = \lambda_0 \text{ versus}$$

$$H_1 : \lambda = \lambda_1 .$$

To use Wald's sequential test, we must first determine

$$L_m = \prod_{i=1}^{m} \frac{f(n_i \mid \lambda_1)}{f(n_i \mid \lambda_0)} .$$

We can determine L_m by considering m samples of size $t_1, t_2, \ldots, t_m$ where $\sum_{i=1}^{m} t_i = t$. This gives

$$L_m = \exp(-\lambda_1 t + \lambda_0 t)\left(\frac{\lambda_1}{\lambda_0}\right)^n .$$

Therefore, $\ln L_m = -t(\lambda_1-\lambda_0) + n \ln \frac{\lambda_1}{\lambda_0}$

Hence, by Wald's approximation, we have

$$\ln \frac{1-\beta}{\alpha} < -t(\lambda_1-\lambda_0) + n \ln (\frac{\lambda_1}{\lambda_0}) < \ln \frac{\beta}{1-\alpha} ,$$

where α = Pr (rejecting H_0 given H_0 is true) and

β = Pr (accepting H_0 given H_1 is true)

are specified in advance.

The above inequality yields

$$\frac{\ln(\frac{1-\beta}{\alpha}) + t(\lambda_1-\lambda_0)}{\ln(\frac{\lambda_1}{\lambda_0})} < n < \frac{\ln(\frac{\beta}{1-\alpha}) + t(\lambda_1-\lambda_0)}{\ln(\frac{\lambda_1}{\lambda_0})} .$$

As long as n at time t remains between these bounds we keep on sampling. As soon as n exceeds the upper bound we accept H_1 and if n is less than the lower bound, we accept H_0.

1.9 The Comparison of Two Poisson Processes

Method I. In direct Poisson sampling in which the number of events occurring in a fixed time is recorded, we have

$$\text{Pr (no. of events} \geq n) = \sum_{r=n}^{\infty} \frac{e^{-\lambda t} (\lambda t)^r}{r!} = \text{Pr } (\frac{1}{2\lambda} \chi^2_{2n} < t)$$

and Pr (no. of events $\geq$ n+1) = Pr $(\frac{1}{2\lambda} \chi^2_{2n+2} < t)$.

If we wish to make an approximation to Pr (no. of events $>$ n) in which the number of events is treated as a continuous random variable, it seems reasonable to take

$$\text{Pr (no. of events} > n) \cong \text{Pr } (\frac{1}{2\lambda} \chi^2_{2n+1} < t)$$

i.e., we calculate probabilities as if

$$2\lambda t \text{ is distributed as } \chi^2_{2n+1} .$$

Thus, if we sample two populations with rates of occurrence λ_1 and λ_2 and in times t_1, t_2 observe n_1 and n_2 events, then

$$\frac{t_1(n_2 + \frac{1}{2})\lambda_1}{t_2(n_1 + \frac{1}{2})\lambda_2}$$

is distributed approximately as the F distribution with $(2n_1 + 1, 2n_2 + 1)$ degrees of freedom. Thus, we may test the hypothesis that $\lambda_1 = \lambda_2$ by referring $F = \frac{t_1(n_2 + \frac{1}{2})}{t_2(n_1 + \frac{1}{2})}$ to the F tables with $(2n_1 + 1, 2n_2 + 1)$ degrees of freedom. Also, a $(100-2\alpha)\%$ confidence interval for $\frac{\lambda_1}{\lambda_2}$ may be obtained from

$$\frac{t_2(n_1 + \frac{1}{2})}{t_1(n_2 + \frac{1}{2})} F_- < \frac{\lambda_1}{\lambda_2} < \frac{t_2(n_1 + \frac{1}{2})}{t_1(n_2 + \frac{1}{2})} F_+ ,$$

where F_- and F_+ are the lower and upper $\alpha\%$ points of F with $(2n_1 + 1, 2n_2 + 1)$ degrees of freedom.

This test may be expected to be accurate for large samples. For small samples, the accuracy was investigated by D. R. Cox (11). The general conclusions are that, except when the population means are very small, the approximate F test gives probability of errors of the first kind sufficiently accurately for all practical purposes. For samples of the same size, the test may be considered satisfactory at the 5% level if the true mean exceeds one and satisfactory at the 1% level if the true mean exceeds two.

Method 2. Suppose we observe two Poisson processes with rates of occurrence λ_1 and λ_2 simultaneously in the time interval $(0,T)$. Then, the probability that any event observed is from the first process is just

$$\rho = \frac{\lambda_1}{\lambda_1 + \lambda_2} = \frac{1}{1 + \frac{\lambda_2}{\lambda_1}}$$

$$= \frac{1}{1+\pi} .$$

To test H_0: $\pi = \pi_0$ (i.e., $\lambda_2 = \pi_0\lambda_1$), we consider the corresponding test

$$H_0': \rho = \frac{1}{1+\pi_0} = \rho_0 .$$

The probability that of the n events observed, k of these came from the first process is

$$\binom{n}{k} \rho^k (1-\rho)^{n-k} .$$

Suppose samples are obtained from two Poisson processes and k is the number of events observed from the first process. The maximum likelihood estimator of ρ is $\hat{\rho} = k/n$, where n is the total number of events occurring. Let K be the random variable that counts the number of occurrences of the first process. Then, we can determine a number $l(\hat{\rho})$ such that $\Pr(k \leq l(\hat{\rho}) - 1 \mid \hat{\rho}) \leq \alpha/2$ and a number $\mu(\hat{\rho})$ such that

$$\Pr(k \geq \mu(\hat{\rho}) + 1 \mid \hat{\rho}) \leq \alpha/2 .$$

Since $$\Pr(AB) \leq \Pr(A) + \Pr(B),$$

we have that

$$\Pr(k \leq l(\hat{\rho}) - 1 \text{ and } k \geq \mu(\hat{\rho}) + 1 \mid \hat{\rho}) \leq \alpha$$

and hence

$$\Pr(l(\hat{\rho}) \leq k \leq \mu(\hat{\rho}) \mid \hat{\rho}) \geq 1 - \alpha .$$

Hence $$\Pr\left(\frac{l(\hat{\rho})}{n} \leq \frac{k}{n} \leq \frac{\mu(\hat{\rho})}{n} \mid \hat{\rho}\right) \geq 1 - \alpha .$$

The confidence interval for ρ is computed as follows. We observe the number of events k from the first process in the first n occurrences of the two processes. We form the estimate $\hat{\rho}$ and find the numbers $l(\hat{\rho})$ and $\mu(\hat{\rho})$ from the appropriate Tables of Binomial Probabilities

or Tables of the Incomplete Gamma Function for the confidence coefficient α. Dividing by n, gives the appropriate $100(1-\alpha)\%$ confidence interval for ρ .

1.10 Comparisons in Terms of Differences in Expected Rates of Occurrence

Rather than make inferences about λ_1 and λ_2 for two independent, Poisson processes, sometimes it is advantageous to make inferences about $\Delta = \lambda_2 - \lambda_1$. In this section, we suppose that the two processes are observed simultaneously and that we can identify the sources.

Let T_m be the waiting time until the m^{th} event has been observed from the two processes occurring simultaneously. Let $\beta > 0$, $\eta > 0$ be two fixed, arbitrary constants. Let $d = \frac{\beta \eta^2}{2m}$. Perform additional observations on each process for $\frac{1}{2dT_m}$ units of time. Let n_1 and n_2 be the respective number of events observed for the two processes during this time period.

Theorem 1.10.1 $\Delta' = 2dT_m(n_2 - n_1)$ is an estimate of $\Delta = \lambda_2 - \lambda_1$ such that

$$\Pr(|\Delta' - \Delta| \leq \eta) \geq 1 - \beta .$$

Proof: By a similar procedure as in (1.7), we can show that

$$E(T_m n_i) = E(T_m E(n_i \mid T_m)) = \frac{\lambda_i}{2d}, i = 1,2$$

and

$$\begin{aligned} \mathrm{Var}(T_m n_i) &= E(T_m^2 \mathrm{Var}(n_i \mid T_m)) \\ &= \frac{m \lambda_i}{(\lambda_1 + \lambda_2)2d}, i = 1,2 . \end{aligned}$$

Therefore,

$$E(\Delta') = 2d(E(T_m n_2) - E(T_m n_1))$$

$$= 2d \frac{\lambda_2 - \lambda_1}{2d} = \Delta$$

and $$Var(\Delta') = 4d^2 \left[Var(T_m n_2) + Var(T_m n_1) \right]$$

$$= 4d^2 \frac{m}{2d} = 2md \ .$$

Hence, applying Tchebycheff's inequality

$$Pr(|\Delta' - \Delta| \leq \eta) \geq 1 - \frac{2md}{\eta^2} = 1 - \beta \ .$$

1.11 Comparison of Three or More Poisson Processes

Let T_i be the waiting time for n_i events to occur from the i^{th} process after an event has been observed from the i^{th} process. Then $2\lambda_i T_i$ has approximately the χ^2 distribution with $2n_i$ degrees of freedom. The problem of comparing the λ_i is equivalent to the problem of comparing variances of k normal processes on the basis of sample variances.

M. S. Bartlett (5) has shown that, for k samples from k independent normal populations, if s_i^2 is the usual estimate of σ_i^2, then

$$M = N \ln \left[\sum_{i=1}^{k} \frac{f_i s_i^2}{N} \right] - \sum_{i=1}^{k} f_i \ln s_i^2$$

is distributed approximately as χ^2 with k-1 degrees of freedom; f_i is the degrees of freedom associated with s_i^2 and $N = \sum_{i=1}^{k} f_i$. Tables at the .01 and .05 level of significance for M are available in Thompson and Merrington (30). If we now substitute T_i for s_i^2 and $2n_i$ for f_i, we can calculate M for k Poisson processes and test the hypothesis that $\lambda_1 = \lambda_2 = , \ldots , = \lambda_k$.

CHAPTER II

Superposition of Several Strictly Periodic Sequences of Events

2.1 Introduction

Suppose there are a number of sources at each of which events occur from time to time. Suppose the outputs are pooled so as to give one combined output. Such outputs occur for example, in the study of threshold activity at motor-nerve endings and in the study of neural nets. We assume that the events on any one source occur at exactly regular intervals, so that the times of occurrence are θ_i, $2\theta_i$, . . . where θ_i is the period of the i^{th} source. If N periodic events are superimposed, where N may be unknown, we would like to:

(1) differentiate the pooled series from a completely random series,

(2) determine the fundamental periods,

(3) determine N (the number of individual sources).

2.2 Weyl's Theorem and Its Application to Pooled Outputs

We assume that the numbers θ_i are all positive.

Definition 2.2.1 A set of numbers $\theta_1, \theta_2, \ldots, \theta_N$ are mutually irrational if there exists no set of integers n_i, not all zero, such that $\sum_{i=1}^{N} n_i\theta_i = 0$.

If $\{\theta_i\}$ is mutually irrational, then $\left\{\frac{\theta_i}{\theta_j}\right\}, i \neq j$, is also

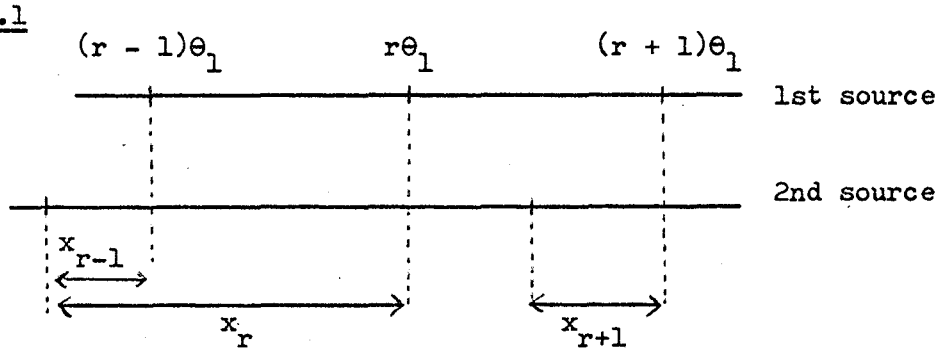
mutually irrational. We suppose that the sources are so numbered so that θ_N is the smallest of the θ_i.

The simplest form of Weyl's Theorem states that if α is irrational and if $\{x\}$ denotes the fractional part of x, then the sequence $(\{n\alpha\})$ $(n = 1, 2, \ldots)$ is uniformly distributed over (0,1). More precisely, if I_ℓ is any interval of length ℓ in (0,1) and if $p_m(I_\ell)$ is the proportion of $\{\theta\}, \{2\theta\}, \ldots, \{m\theta\} \ldots$ falling in I_ℓ, then $\lim_{m\to\infty} p_m(I_\ell) = \ell$.

To reformulate, the above result in terms of sources, consider any two sources, say the first two.

Associate with the r^{th} event on the first source a quantity x_r equal to the time between that event and the immediately preceding event on the second source (see Fig. 2.2.1).

Figure 2.2.1



Then $x_r = \theta_2 \left\{ \frac{r\theta_1}{\theta_2} \right\}$. It follows from Weyl's theorem, that since $\frac{\theta_1}{\theta_2}$ is irrational, the sequence (x_r) is uniformly distributed over $(0,\theta_2)$.

The generalized form of Weyl's Theorem states that if α_1, $\alpha_2, \ldots, \alpha_k$ are irrational numbers, themselves mutually irrational, then the sequences

$$(\{n\alpha_1\}), (\{n\alpha_2\}), \ldots, (\{n\alpha_k\}), n = 1,2, \ldots$$

are independently, uniformly distributed over (0,1). More precisely, if I_v is any portion of the volume v of the k-dimensional unit cube and if $p_m(I_v)$ is the proportion of the m k-plets $\{\alpha_1\}, \ldots, \{\alpha_k\}$; $\ldots$ $\{m\alpha_1\}, \ldots, \{m\alpha_k\}$ in I_v, then $\lim_{m\to\infty} p_m(I_v) = v$.

We define for the r^{th} event on the i^{th} source a set of (N-1) quantities, $x_r^{(1)}, \ldots, x_r^{(i-1)}, x_r^{(i+1)}, \ldots, x_r^N$ analogous to x_r above. Forexample, $x_r^{(1)}$ is the interval between the r^{th} event on the i^{th} source and immediately preceding event on the 1^{th} source. Then the generalized form of Weyl's theorem shows that the sequences $(x_r^{(1)}), \ldots, (x_r^{(i-1)}), (x_r^{(i+1)}), \ldots, (x_r^N)$ are independently, uniformly distributed over $(0,\theta_1), \ldots, (0,\theta_N)$.

2.3 Frequency Distribution of Intervals Between Successive Events

Let $q_i(y)$ be the frequency function of the length of intervals ending with an event from source i. The interval is between $(y,y+\Delta y)$ if one source, say the j^{th}, has its x-value between $(y,y+\Delta y)$ and if all the other sources have their x-values greater than y. The frequency of the first event is $\frac{\Delta y}{\theta_j}$ and of the second is $\prod_{k\neq i,j} \frac{\theta_k - y}{\theta_k}$. Since the x values are independent

$$q_i(y) = \sum_{j\neq i} \frac{1}{\theta_j} \prod_{k\neq i,j} \frac{\theta_k - y}{\theta_k} \quad (0 \leq y \leq \theta_N).$$

When i = N there is a point concentration at $y = \theta_N$ given by

$$Q_N = \prod_{k=1}^{N-1} \frac{\theta_k - \theta_N}{\theta_k}.$$

If $i \neq N$, there is no point frequency.

The overall frequency distribution is defined by a frequency function q(y) and by a point frequency Q obtained by taking a weighted

average of the $q_i(y)$. In a long time, the number of events from the i^{th} source is proportional to $\frac{1}{\theta_i}$, so that

$$q(y) = \sum_{i=1}^{N} \frac{q_i(y)}{\theta_i} \Big/ \sum_{i=1}^{N} \frac{1}{\theta_i}$$

$$= \sum_{i=1}^{N} \frac{\frac{1}{\theta_i} \sum_{j \neq i} \frac{1}{\theta_i} \prod_{k \neq i,j} \frac{\theta_k - y}{\theta_k}}{\sum_{i=1}^{N} \frac{1}{\theta_i}}, (0 \leq y \leq \theta_N),$$

$$Q = \frac{\frac{1}{\theta_N} \prod_{k=1}^{N-1} \frac{\theta_k - \theta_N}{\theta_k}}{\sum_{i=1}^{N} \frac{1}{\theta_i}} .$$

We wish the above forms more suitable for computational purposes. To do this, introduce the symmetric functions

$$\sigma_1^{(N)} = \theta_1 + \ldots + \theta_N$$
$$\sigma_2^{(N)} = \sum_{i_1 > i_2} \theta_{i_1} \theta_{i_2}$$
$$\vdots$$
$$\sigma_s^{(N)} = \sum_{i_1 > i_2, \ldots, > i_s} \theta_{i_1} x \ldots x \theta_{i_s}$$
$$\vdots$$
$$\sigma_N^{(N)} = \theta_1 x \ldots x \theta_N .$$

Then $q(y) = \frac{1}{\sigma_{N-1}^{(N)}} \sum_{i=1}^{N} \sum_{j \neq i} \prod_{k \neq i,j} (\theta_k - y)$ and $Q = \frac{\prod_{k=1}^{N-1} (\theta_k - \theta_N)}{\sigma_{N-1}^{(N)}}$.

The distribution consists of a discrete frequency Q at the smallest period θ_N and a continuous frequency curve between 0 and θ_N defined by a polynomial of degree N-2. When N is small, the above

distributions are easily computed numerically. When N is large we shall give the asymptotic distribution. It must be pointed out that if the periods are close together, even for small N, the amount of computation required is fairly large.

The mean of the distribution is easily obtained, since in a very long time T, there are asymptotically T/θ_i events from the i^{th} source and hence, asymptotically, $T \sum_{i=1}^{N} \frac{1}{\theta_i}$ events altogether. Therefore,

$$E(y) = \left[\sum_{i=1}^{N} \frac{1}{\theta_i} \right]^{-1} .$$

2.4 Asymptotic Properties

In this section we shall show that, under certain conditions on the θ_i, the distribution of intervals between successive events tends to the exponential distribution as N tends to infinity.

Let $\chi_i = \frac{1}{\theta_i}$. Let $Z = y \sum_{i=1}^{N} \chi_i$. The frequency function r(Z) of Z is then

$$r(Z) = \frac{\sum_{i=1} \chi_i \sum_{j \neq i} \chi_j \prod_{k \neq i,j} \left(1 - \frac{Z\chi_k}{\sum \chi_e}\right)}{\left[\sum \chi_e\right]^2}$$

since $dz = dy \sum_{i=1}^{N} \chi_i$.

$$\therefore r(Z) = \frac{\sum_{i \neq j} \prod_{k \neq i,j} \left(1 - \frac{Z\chi_k}{\sum \chi_e}\right) \chi_i \chi_j}{(\sum \chi_e)^2} .$$

Let $\rho(Z) = \prod_{i=1}^{N} (1 - \frac{Z\chi_i}{\Sigma\chi_e})$.

As is easily verified,

$$r(Z) = \frac{d^2\rho(Z)}{dZ^2} .$$

Now let

$$\mu = \frac{\Sigma\chi_i}{N}$$

$$\mu_s' = \frac{\Sigma\chi_i^s}{N} \qquad s = 2, 3, \ldots .$$

Then

$$\ln \rho(Z) = \sum_{i=1}^{N} \ln(1 - \frac{Z\chi_i}{\Sigma\chi_e})$$

$$= -Z - \frac{Z^2}{2}\frac{\mu_2'}{N\mu^2} - \frac{Z^3}{3}\frac{\mu_3'}{N^2\mu^3} - \ldots .$$

We now assume that, as N tends to infinity, the χ_i have both an upper bound and a non-zero lower bound. This assumption implies that as $N \to \infty$, $\frac{\mu_s'}{\mu^s}$ is bounded for each s.

Letting N tend to infinity above, and under our assumption, we obtain that

$$\rho(Z) \sim e^{-Z} \quad \text{and}$$

$$r(Z) = \frac{d^2\rho(Z)}{dZ^2} \sim e^{-Z} .$$

We can differentiate because $\rho(Z)$ for all finite N, and the limit e^{-Z}, are integral functions.

The continuous frequency curve r(Z) has range $0 \le Z \le \frac{1}{\chi_N}\Sigma\chi_i$. We shall also assume that the smallest period θ_N is large compared with the mean interval, i.e., $\lim_{N\to\infty} \frac{1}{\chi_N} \sum_{i=1}^{N} \chi_i = \infty$. Under this assumption,

the point frequency Q tends to zero. Thus, the continuous frequency curve r(Z) accounts asymptotically for all the frequency.

We shall now prove the following.

Theorem 2.4.1 If t is small compared with all the periods θ_i, then the number of events observed in time t is a Poisson random variable with mean $t(\sum \chi_i)$.

Proof: Consider first a single source. A time interval $t < \theta_i$ will include either no event or one event, and the proportion of intervals including one event is t/θ_i. Thus, the generating function of the frequencies is

$$G(s) = (1 - t/\theta_i) + st/\theta_i = 1 - (1-s)t\chi_i .$$

Since the sources are independent, the generating function for the pooled output is

$$G_p(s) = \prod_{i=1}^{N} (1 - (1-s)t\chi_i)$$

$$\therefore \ln G_p(s) = \sum_{i=1}^{N} \ln(1 - (1-s)t\chi_i) .$$

If we assume $t\chi_i$ is small for all i,

$$\ln G_p(s) \sim \sum_{i=1}^{N} -t\chi_i(1-s)$$

$$\therefore G_p(s) \sim \exp-(t(1-s) \sum \chi_i) .$$

But the above is the generating function for the Poisson distribution with mean $t \sum \chi_i$.

The above results are special cases of the following intuitively obvious principle. Suppose there are a number of sources and we are

interested in some property of the pooled output depending only on its behaviour over times small compared with the individual recurrence times. Then, the result is indistinguishable from that of a random series, no matter what the form of the individual outputs.

2.5 The Variance Time Curve and Methods of Determining the Sources

Several practical problems now can be answered. In particular, we first would like to determine whether the series is indeed a pooled series of periodic sources rather than a random series. The pooled output is distinguished from a random series by its behaviour over lengths of time comparable with the individual periods θ_i. The most convenient way of expressing this behaviour is by means of a variance-time curve, V(t).

Definition 2.5.1 The variance-time curve V(t) associated with a point process is defined as the variance of the number of events occurring in a time t, considered as a function of t.

For a random series $V(t) = \lambda t$ where $\frac{1}{\lambda}$ is the mean interval between successive events.

To find V(t) for a pooled output of periodic sources first consider a single source. Let $\chi_i t = n_i + \beta_i$ where n_i is an integer and $0 \leq \beta_i < 1$, i.e., $\beta_i = \{\chi_i t\}$. An interval of length t contains either n_i or n_i+1 events from this source and the limiting frequency of intervals containing n_i+1 events is β_i. The variance of this two point distribution is $\beta_i(1 - \beta_i)$. Since the different sources are independent then for N sources

$$V(t) = \sum_{i=1} \beta_i(1 - \beta_i).$$

Now if t is very small compared with θ_i, then

$$\beta_i(1 - \beta_i) \sim \beta_i = t\chi_i$$

so that $V(t) \sim t \sum \chi_i = t\lambda$ where $\frac{1}{\lambda}$ is the mean interval between successive events. Further, $\beta_i(1 - \beta_i) < t\chi_i$ so that

$$V(t) < t\lambda .$$

For t large compared with θ_i,

$$\beta_i(1 - \beta_i) << t\chi_i$$

and

$$V(t) << t\lambda .$$

As t increases, β_i takes each value between 0 and 1 equally often, (Weyl's theorem) giving $\beta_i(1 - \beta_i)$ an average value of $\frac{1}{6}$. Thus for large t, V(t) oscillates about an average of $\frac{1}{6}N$.

In summary, we have that V(t) is tangential at t=0 to the straight line $y=\lambda t$ representing a random series with the same number of events per unit time. V(t) falls below the line as soon as t is comparable with an appreciable number of periods θ_i, and finally V(t) oscillates about $\frac{1}{6}N$ as soon as t is large enough.

We can now answer the problems posed in the introduction. By constructing the variance-time curve, we can decide whether the series is random or a pooled output of periodic sources and at the same time we obtain an estimate of N, the number of sources. Obviously, any method based on frequency distribution of intervals will only distinguish between random and pooled periodic series if N is relatively small, because the frequency curve is exponential when N is large.

We now proceed to estimate the variance-time curve from the data, since the actual variance-time curve is practically always unknown. We can, of course, divide the observed series into units of

time t and estimate the variance in the usual way. Another method given by Cox (8) provides more economy in observation. First, divide the series into intervals of length τ so that only a small number of events occur in each interval. Next count the number of events occurring in each interval deriving a discrete time series $X_1, X_2, \ldots, X_m$ where X_i is the number of events in $(i-1)\tau \leq t < i\tau$. Add the X's together in blocks of r, giving

$$S_1^r = X_1 + X_2 + \ldots + X_r$$
$$S_2^r = X_2 + X_3 + \ldots + X_{r+1}$$
$$\vdots$$
$$S_{m-r+1}^r = X_{m-r+1} + X_{m-r+2} + \ldots + X_m .$$

The S^r's are the number of events occurring in intervals of length $r\tau$ so that an estimate of $V(r\tau)$ may be formed from the corrected sums of squares of the S^r's. Let $M = m-r+1$. Then, we let

$$\hat{V}(r\tau) = \left[\frac{3M}{3M^2 - 3Mr + r^2 - 1}\right] \sum_{i=1}^{M} \left(S_i^r - \overline{S}_i^r\right)^2$$

where $$\overline{S}_i^r = \frac{1}{M} \sum_{i=1}^{M} S_i^{(r)} .$$

The factor in front of the sums of squares has been so chosen so that if the series is random, i.e., X_i are independent Poisson random variables with mean $\lambda\tau$, then $E(\hat{V}(r\tau)) = V(r\tau) = r\tau\lambda$. We do the above for various values of r and plot $\hat{V}(r\tau)$ against $r\tau$.

Now, if it is determined that N is small and the series available for analysis is long, it is possible in principle to determine the θ_i exactly and to assign each event to its appropriate source. To do this, we first form the frequency distribution of the intervals

between successive events. This will be bounded by a point concentration of frequency whose position will denote the smallest period θ_N. Next, find an interval of length θ_N and from it build up the output of the N^{th} source by repeatedly adding and subtracting θ_N. Delete this set of events and analyze the remaining series to find the next smallest period and so on. This method ceases to become practical as soon as the point frequency becomes very small.

Several practical problems arise in the determination of sources of pooled periodic outputs. First of all, if the sources are close together, the point frequency will be too small to be of any value unless the series is observed for a very long time. Secondly, if the periods of the sources are close together, the distribution of intervals tends very rapidly to the exponential form as N increases. Therefore, the methods advocated are useful if the number of sources is relatively small, and if the periods of the sources are relatively far apart.

It must also be pointed out that these results lean very heavily on Weyl's theorem which demands that the periods be mutually irrational. In actual practice, the θ_i cannot be mutually irrational and even if they were, they could not so be determined. However, we may either regard them as being close approximations to irrationals or argue that since the lowest common multiple of the θ_i is extremely large, the results derived on the assumption of mutual irrationality are likely to be extremely good approximations to the behaviour for rational θ_i.

CHAPTER III

Superposition of non-periodic processes

3.1 Renewal Processes

In the previous chapter, we considered a pooled series of strictly periodic sources. We shall now consider the case where the intervals between successive events have a probability distribution and the random variables associated with different sources are independent. For the sake of illustration, we consider the case of a single source. The process thus formed is called a renewal process and has been extensively studied.

Let the intervals between successive events be denoted by a sequence X_n. We assume the X_n are independent, identically distributed random variables all with an absolutely continuous distribution function $F(x)$ and a frequency function $f(x)$. The n^{th} event occurs at time $Z_n = X_1 + \ldots + X_n$ and if $f_n(x)$ is the frequency function of Z_n, then $h(x) = \sum_{n=1}^{\infty} f_n(x)$ is the total density of events at x.

Now we fix a time t and define a delay time $Y(t)$ to be measured from t back to the immediately preceding event. The frequency function $g(y;t)$ of $Y(t)$ is called the delay function at t. Since the total density of events at x is $h(x)$, and the chance that the interval between events exceeds y is $1-F(y) = F_c(y)$, we have that $g(y;t) = h(t-y)\, F_c(y)$.

It has been shown that under certain weak conditions on f(x), the total density of events at x, h(x) has the property

$$\lim_{t \to \infty} h(t) = \frac{1}{\mu} .$$

Therefore, under these conditions we have that

$$\lim_{t \to \infty} g(y,t) = \frac{F_c(y)}{\mu} .$$

3.2 Probability Distribution and Variance of the Number of Events in Intervals of Given Length

Before giving a rigorous proof, we shall discuss heuristically the result that if V(t) is the variance of the number of events in time t, then $V(t) \sim \frac{C^2 t}{\mu}$ as $t \to \infty$ where C is the coefficient of variation of the parent distribution (F(x)) and $\mu = E(X_n)$. The result follows heuristically from two equations in sequential analysis. If $Z_n = X_1 + X_2 + \ldots + X_n$ is the cumulative sum when sampling stops after n events, then $E(Z_n) = E(n)\mu$ and $E(Z_n - n\mu)^2 = E(n)\sigma^2$ where σ^2 is the variance of each of the X_i. Now Z_n, in this case, is the time up to the n^{th} renewal. If sampling is stopped after a long time t, $Z_n \sim t$, where n is the number of events occurring in time t. Therefore

$$E(n) \sim \frac{t}{\mu} \quad \text{and} \quad E(t - n\mu)^2 \sim \frac{\sigma^2 t}{\mu}$$

$$\text{also} \quad \text{var}(n) = \frac{1}{\mu^2} E(n\mu - t)^2$$

$$\text{i.e.,} \quad V(t) = \text{var}(n) \sim \frac{C^2 t}{\mu} .$$

In our case we are interested in the equilibrium behaviour of the process, i.e., we are interested in the number of events in intervals a long way from the origin.

Let $p_k^{u}(t) = \Pr\left[k \text{ events occur in } (u, u+t)\right]$.

In particular, $p_k^{o}(t) = \Pr\left[k \text{ events occur in } (o,t)\right]$.

Now $p_k^{u}(t) = \int_o^t g(y;u)\, p_{k-1}^{o}(t-y)dy$.

If we assume that $\lim_{t\to\infty} g(y;t)$ exists (it can be shown that under weak conditions, the limit actually exists, Cox and Smith (13)) then $g(y;u)$ for sufficiently large u is bounded for $(0 \le y \le t)$; also $0 \le p_{k-1}^{o}(t) \le 1$. Therefore, by the theorem of bounded convergence, we have

$$\lim_{u\to\infty} \int_o^t g(y;u)\, p_{k-1}^{o}(t-y)dy = \int_o^t g(y)\, p_{k-1}^{o}(t-y)dy,$$

where we write $g(y) = \lim_{u\to\infty} g(y;u)$.

Let $p_k(t) = \lim_{u\to\infty} p_k^{u}(t)$.

Then $p_k(t)$ is called the equilibrium distribution of the number of events. It was shown in Cox and Smith (13) that $g(y) = \frac{1 - F(y)}{\mu} = \frac{F_c(y)}{\mu}$.

Theorem 3.2.1 The variance of the equilibrium distribution

$$V(t) \sim \frac{c^2 t}{\mu} + \frac{3\mu_2^2 - 2\mu\mu_3}{6\mu^4}.$$

Proof:
$$p_k(t) = \frac{1}{\mu} \int_o^t \left[1 - F(y)\right] p_{k-1}^{o}(t-y)dy.$$

Let $z = t-y$.

Then
$$p_k(t) = -\frac{1}{\mu} \int_t^o \left[1 - F(t-z)\right] p_{k-1}^{o}(z)dz$$

$$= \frac{1}{\mu} \int_o^t \left[p_{k-1}^{o}(z) - F(t-z)\, p_{k-1}^{o}(z)\right] dz.$$

But
$$\int_0^t p_k^{\,o}(z)dz = \int_0^t dz \int_0^z f(z-v)\, p_{k-1}^{o}(v)dv$$
$$= \int_0^t F(t-z)\, p_{k-1}^{o}(z)dz$$
$$\therefore p_k(t) = \frac{1}{\mu} \int_0^t \left[p_{k-1}^{o}(z) - p_k^{\,o}(z) \right] dz.$$

Let
$$M_\theta^o(t) = \sum_{k=0}^{\infty} e^{k\theta}\, p_k^{\,o}(t),$$
$$M_\theta(t) = \sum_{k=0}^{\infty} e^{k\theta}\, p_k(t).$$

By substitution, we obtain

$$M_\theta(t) = 1 - \frac{(1-e^{\theta})}{\mu} \int_0^t M_\theta^o(z)dz.$$

By differentiating $M_\theta(t)$ with respect to θ and setting θ equal to zero, we obtain that, when the appropriate quantities exist,

$$m_1(t) = \sum_{k=0}^{\infty} k\, p_k(t) = {}^{t}/_{\mu}$$
$$m_2(t) = \sum_{k=0}^{\infty} k^2\, p_k(t) = {}^{t}/_{\mu} + \frac{2}{\mu} \int_0^t m_1^{\,o}(z)dz.$$

Therefore
$$V(t) = \frac{1}{\mu} \int_0^t \varphi(z)dz \qquad \text{where}$$

$$\varphi(z) = 1 + 2m_1^{\,o}(z) - \frac{2z}{\mu}.$$

Now the Laplace transform of $\varphi(z)$;

$$\mathcal{L}\left[\varphi(z)\right] = \bar{\varphi}(s) = \frac{1}{s} + \frac{2\bar{f}(s)}{s[1 - \bar{f}(s)]} - \frac{2}{\mu s^2}$$

since

$$m_1^{o}(t) = \sum_{k=0}^{\infty} k\, P_k^{o}(t)$$

$$= \sum_{k=1}^{\infty} k \int_o^t f(t-u)\, P_{k-1}^{o}(u)du$$

$$= \int_o^t f(t-u) \sum_{k=1}^{\infty} k\, P_{k-1}^{o}(u)du$$

$$= \int_o^t f(t-u) \sum_{k=0}^{\infty} (k+1)\, p_k^{o}(u)du$$

$$= \int_o^t f(t-u)\, m_1^{o}(u)du + \int_o^t f(t-u)du$$

then

$$\bar{m}_1^{o}(s) = \bar{f}(s)\, \bar{m}_1^{o}(s) + \frac{1}{s}\, \bar{f}(s)$$

and

$$\bar{m}_1^{o}(s) = \frac{\bar{f}(s)}{s(1 - \bar{f}(s))} .$$

Let μ, μ_2, u_3 . . . be the moments of f(x) about zero and let C be the coefficient of variation

$$C^2 = \frac{(\mu_2 - \mu^2)}{\mu^2} .$$

Expanding $\bar{\varphi}(s)$ and after some tedious algebra, as $s \to o$ we obtain

$$\bar{\varphi}(s) \sim \frac{C^2}{s} + \frac{3\mu_2^{\,2} - 2\mu\mu_3}{6\mu^3} .$$

This suggests that as $t \to \infty$,

$$V(t) = \frac{1}{\mu} \int_o^t \varphi(z)dz \sim \frac{C^2 t}{\mu} + \frac{3\mu_2^{\,2} - 2\mu\mu_3}{6\mu^4} .$$

Proof of this is long and difficult (given in /3).

The above equation shows the behavior of V(t) for large t. For small t, we have

$$V(t) = \frac{1}{\mu} \int_0^t \left[1 + 2m_1^{\circ}(z) - \frac{2z}{\mu} \right] dz$$

$$= \frac{t}{\mu} + 2 \int_0^t \left[m_1^{\circ}(z) - \frac{z}{\mu} \right] dz .$$

If we assume that $\int_0^t f(x)dx = O(t^{\beta}), \beta > o$ as $t \to o$, then

$$V(t) = \frac{t}{\mu} + o(t).$$

Thus near t = o the variance timecurve behaves like that of a completely random series.

If f(x) is given explicitly, we can obtain more precise results. For example, if f(x) is χ^2 with 2V degrees of freedom, then $C^2 = \frac{1}{V}$ and

$$V(t) \sim \frac{t}{V\mu} + \frac{1}{6}(1 - \frac{1}{V^2}) \text{ as } t \to \infty .$$

3.3 Pooled Outputs of Renewal Processes

Suppose that N sources are pooled to form a pooled process. We shall assume that the outputs of the individual sources are of the type considered previously, are independent, and have the same absolutely continuous parent distribution F(x). As before, we shall only be interested in equilibrium behaviour a long time after the start of the process.

First, we compute the frequency distribution of the interval between successive events in the pooled output. For each source, we define the delay random variable Y_i, where Y_i is the time from a fixed epoch back to the immediately preceding event on the i^{th} source. If Y is the corresponding random variable for the pooled output, then

$$Y = \min (Y_1, \ldots, Y_N).$$

Since the N sources are independent,

$$\Pr[Y \geq y] = \left[\int_y^\infty g(x)dx\right]^N = \left[\int_y^\infty \frac{F_c(x)}{\mu}\, dx\right]^N.$$

The delay function for the pooled outputs is

$$\frac{d\Pr[Y \geq y]}{dy} = \frac{NF_c(y)}{\mu}\left[\int_y^\infty \frac{F_c(x)}{\mu}\, dx\right]^{N-1}$$

and the frequency function of the interval between successive events

$$f_p(y) = -\frac{d}{dy}\left[F_c(y)\left[\int_y^\infty \frac{F_c(x)}{\mu}\, dx\right]^{N-1}\right].$$

For example, if $f(y) = e^{-y}$, then $f_p(y) = Ne^{-Ny}$, expressing the obvious fact that if separate outputs are completely random so is the pooled output. If the parent distributions are rectangular over (0,1), then

$$f_p(y) = (2N-1)\,(1-y)^{2N-2}.$$

The equilibrium distribution for the number of events in an interval of length t may be found by convoluting the distributions for individual sources. Hence, the equilibrium variance of the number of events in an interval is

$$V_p(t) = NV(t).$$

Therefore $V_p(t) \sim \frac{NC^2t}{\mu} = C^2t\lambda$ as $t \to \infty$ where $\lambda^{-1} = \frac{\mu}{N}$ is the mean interval between events on the pooled output. Also $V_p(t) \sim t\lambda$ as $t \to 0$. To estimate N, further assumptions on f(x) are required. For example, if f(x) is χ^2, then the intercept of $V_p(t)$ is $\frac{N}{6}(1 - C^4)$ and N can thus be estimated.

CHAPTER IV

General Methods of Analyzing Point Processes

4.1 Introduction

In the last two chapters, we have considered special cases of point processes, namely the Poisson and superposition of independent processes. In this chapter, we will consider some methods of analyzing point processes generally, mostly without considering an underlying model for the process. Unfortunately, the problem as stated is too general and we shall have to consider models less general than the one suggested. In most cases, the problems arising are computational in nature.

4.2 Graphical Methods of Analysis

In most practical cases, the statistician is faced with a short record of a point process, i.e., he observes the process for a fixed period of time T. Suppose that in this time period T, n events have occurred at times $t_1, t_2, \ldots, t_n$ in that order. Let $X_{j,j+k} = t_{j+k} - t_j$ with $t_o = 0$. For the sake of convenience, we shall fix a time period τ and let $n_i(\tau)$ be the number of events in the time interval $[(i-1)\tau, i\tau]$. The main graphical methods are the following: (1) Time T is divided into equal intervals τ and $n_i(\tau)$ is plotted against i. If the rate of occurrence is constant, this graph should

approximate a straight line. However, if there are large variations in the rate of occurrence the restriction to a fixed τ is inconvenient.
(2) A random-walk diagram can be plotted for the total number of events up to a time t as a function of t, i.e., N(t) is plotted versus t. From this graph, the average rates of occurrence over any period can be read off. The disadvantages of this method are the difficulty of detecting sampling fluctuations in the cumulative diagram and the awkward shape of the curve, making it very difficult to plot extensive data concisely.
(3) Plot $N(t) - a^{-1}t$ against t, when a is a suitable constant approximately equal to the mean interval between events. A suitable choice of "a" can make the computations simpler. If required, the value of n can be marked at suitable points along the graph. This procedure will smooth the curve, making it easier to plot extensive data. From it, one should be able to read off local rates of occurrence. There is also the possibility of plotting $(t - b^{-1}n)$ against n for a suitable choice of b, but this procedure is, on the whole, less useful. However, in preliminary analyses, one should never discount any procedure, a priori, unless one has a fairly good idea of what the underlying mechanism of the process is.
(4) The intervals $X_{o,k}$; $X_{k;2k}$, . . . can be plotted against the serial number of the interval or against the time at the mid-point of the interval. This method adjusts τ to the rate of occurrence. If it is required to examine the relation between the rate of occurrence and some smoothly varying external variable, the abscissa can be taken as the average of the external variable over the interval.

(5) A histogram can be formed for the interval between successive events and a scatter diagram obtained for successive intervals. Then a distribution for the interval between successive events can be fitted, if desirable or necessary. It is sometimes useful to consider intervals between every pair of events or between every other event.

4.3 Transformations

We often have problems in which the rate of occurrence λ is not constant but is the product either of unknown parameters or of known constants with unknown parameters. This may occur in the following ways:

(1) The interval of observation is not constant, e.g., in observing the number of stops in machines it is usually not practical to arrange that the observed running time is the same for all machines.

(2) It may be intuitively reasonable or suggested by inspection of results that, say in a two-way arrangement, row and column effects are multiplicative, i.e., the λ appropriate to each observation is a product of a row constant and a column constant.

If it is possible to obtain intervals between successive sets of k events so that λ is approximately a constant on each interval, the logarithmic transformation can be applied and causes no difficulty. However, if we work with numbers of events in fixed time intervals there is more difficulty. There is, however, one case for which we can obtain a transformation having desired properties.

Suppose we make an analysis of transformed observations where we are only going to use "unweighted" linear combinations, i.e., all our estimates will be based on linear combinations of the transformed

variables (observations) with coefficients not depending on the observations. If our linear estimates are to be unbiased, and if we wish to find linear unbiased estimates of their sampling variance, we must try to find, corresponding to the Poisson variate n, a transformed value z_n and an estimated variance v_n so that

$$E(z_n) = \ln\lambda \qquad \text{and}$$

$$\mathrm{Var}(z_n) = E(v_n).$$

It is impossible to satisfy these exactly for all λ. For large n, we have that $z_n \sim \ln \frac{n}{T}$ and $v_n \sim \frac{1}{n}$ so that we try

$$z_n = \ln \frac{n + \alpha}{T} \quad \text{and} \quad v_n = \frac{1}{n + \beta}$$

where α and β are to be determined.

If $\theta = \lambda T$ and $\Delta n = n - \theta$, we have

$$\begin{aligned} z_n &= \ln \frac{(n + \alpha)\lambda}{\theta} \\ &= \ln\lambda + \ln \frac{\theta + n - \theta + \alpha}{\theta} \\ &= \ln\lambda + \ln(1 + \frac{\Delta n}{\theta} + \frac{\alpha}{\theta}) \\ &= \ln\lambda + (\frac{\alpha}{\theta} + \frac{\Delta n}{\theta}) - \frac{1}{2}(\frac{\alpha}{\theta} + \frac{\Delta n}{\theta})^2 + \frac{1}{3}(\frac{\alpha}{\theta} + \frac{\Delta n}{\theta})^3 + \ldots . \end{aligned}$$

Now $E(\Delta n) = 0$, since $E(n) = \lambda T = \theta$ and $E(\Delta n)^2 = \mathrm{Var}(n) = \theta$.

If we assume that Δn and α are small compared to θ, then,

$$\begin{aligned} E(z_n) &= \ln\lambda + \frac{\alpha}{\theta} - \frac{\theta}{2\theta^2} + O(\frac{1}{\theta^2}) \\ &= \ln\lambda + \frac{\alpha}{\theta} - \frac{1}{2\theta} + O(\frac{1}{\theta^2}) \\ &= \ln\lambda + \frac{\alpha - \frac{1}{2}}{\theta} + O(\frac{1}{\theta^2}) . \end{aligned}$$

Therefore we take $\alpha = \frac{1}{2}$ and we have

$$E(z_n) = \ln\lambda + O(\frac{1}{\theta^2}) .$$

With this value of α, we then have

$$\text{that} \qquad \mathrm{Var}(z_n) = \frac{1}{\theta} + \frac{1}{2\theta^2} + O(\frac{1}{\theta^3})$$

$$\text{while} \qquad E(\frac{1}{n+\beta}) = \frac{1}{\theta} + \frac{1-\beta}{\theta^2} + O(\frac{1}{\theta^3})$$

so that we choose $\beta = \frac{1}{2}$ in order that

$$\mathrm{Var}(z_n) = E(v_n) \ .$$

Thus, we have arrived at the transformation

$$z_n = \ln \frac{n+\frac{1}{2}}{T} \qquad \text{with} \qquad v_n = \frac{1}{n+\frac{1}{2}} \ .$$

Since the above was arrived by a series expansion in $\frac{1}{\theta}$, we can reasonably expect that $E(z_n) = \ln\lambda$ and $V(z_n) = E(v_n)$ should be satisfied to say $\theta \sim 5$. Computations made for various values of $\theta = 1,2,3,4,5$ suggest that the above is indeed the case.

CHAPTER V

Correlation Analysis of Point Processes

5.1 The Correlation Function

One of the most developed techniques in the theory of stochastic processes is correlation analysis. For any stochastic process $\{X(t),\ t \in T\}$, we define the correlation function

$$R(t,s) = E[X(t),\ X(s)].$$

It should be pointed out that other terminology is often encountered in the literature. The function $R(t,s)$ or else the centered function $R(t,s) - m(t)\ m(s)$, where $m(t) = E[X(t)]$, is sometimes called the covariance function and the term correlation function or autocorrelation function is reserved for a stationary covariance. In the stationary case, the normalized and centred correlation function is known as the "correlation" function. In spite of all the confusion of terminology, we shall denote the correlation function as above.

There are several historical reasons for correlation analysis even though it should be emphasized that the mean value and the correlation function do not specify the stochastic process $\{X(t),\ t \in T\}$ uniquely. First of all, the primary interest and the most highly developed theory is that of Gaussian or normal processes. For normal stochastic processes, the mean value and the correlation function

completely specify the process, i.e., they completely determine all the distribution functions. Therefore, in principle, correlation theory can answer any question pertaining to normal random functions. Even for non-normal processes, the second order moments, i.e., correlation function, tell us something about the process. For example, if the process is stationary $R(t,s)$ is a function of $|s-t|$. The behaviour of this function gives us some idea of the relationship of the process at time instants $|s-t|$ units apart. In the ordinary cases of random variables, we are mostly interested in the first and second moments of these variables. Similarly, in the study of stochastic processes (families of random variables), first and second moments are of equal importance.

5.2 Properties of the Correlation Function for Stationary Processes

If $\{X(t),\ t \in T\}$ is a stationary process then the correlation function is a function of the parameter difference. We can thus write the correlation function as $R(\tau) = E[X(t+\tau)\ X(t)]$. The following properties of $R(\tau)$ can be noted:

(1) $R(\tau) = R(-\tau)$. $R(\tau)$ is an even function of τ.

(2) $R(\tau)$ is continuous everywhere if it is continuous at $\tau = 0$.

In applications of probability theory, one ordinarily deals with phenomena that repeat themselves many times. Hence, as the mean value of a random variable $X(t)$, we can take a large number of realizations $X^{(1)}(t)$, $X^{(2)}(t)$, . . . , $X^{(n)}(t)$ and compute the arithmetic mean $\frac{1}{n}\sum_{j=1}^{n} X^{j}(t)$. Similarly, for the correlation function $R(t,s)$ we can take the mean value of $X^{(j)}(t)\ X^{(j)}(s)$ for every pair of values

of t and s. However, in practice, observation of a stochastic process and subsequent processing of data can be very complicated and therefore it is very desirable to get along with as few realizations as possible. The possibility of calculating these characteristics of a stationary stochastic process from a single realization is a consequence of the so-called ergodic theorem applicable to most stationary random functions encountered in practice. According to the ergodic theorem, the mathematical expectation of both $X(t)$ and $X(t+\tau)\,X(t)$, obtained by averaging over the space of outcomes can be replaced by averaging over T. More precisely, if $\{X(t),\ t \in T\}$ is a stationary stochastic process satisfying certain general conditions to be indicated below, then the following limiting (in the mean square) relations hold:

(1) $$E[X(t)] = m = \lim_{N\to\infty} \frac{1}{N+1} \sum_{t=0}^{N} X(t)$$ in the discrete parameter case

$$= \lim_{T\to\infty} \frac{1}{T} \int_0^T X(t)\,dt$$ in the continuous parameter case.

(2) $$R(\tau) = E[X(t+\tau)\,X(t)] = \lim_{N\to\infty} \frac{1}{N+1} \sum_{t=0}^{N} X(t+\tau)\,X(t)$$

in the discrete parameter case

$$= \lim_{T\to\infty} \frac{1}{T} \int_0^T X(t+\tau)\,X(t)\,dt$$

in the continuous parameter case.

The above relations suggest that for sufficiently large N (or T), we can approximate m and $R(\tau)$ by $\frac{1}{N+1} \sum_{t=0}^{N} X(t)$ and $\frac{1}{N+1} \sum_{t=0}^{N} X(t+\tau)\,X(t)$, where $X(t)$ is a single realization of the process.

In order to make use of the formulas for the continuous parameter case, we replace the integrals by their approximating sums,

noting that

$$\int_0^T X(t)dt = \lim_{N\to\infty} \frac{T}{N} \sum_{k=1}^{N} X(k\frac{T}{N})$$

$$\int_0^T X(t+\tau)\, X(t)dt = \lim_{N\to\infty} \frac{T}{N} \sum_{k=1}^{N} X(\frac{kT}{N}+\tau)\, X(\frac{kT}{N}) \ .$$

Hence, m and $R(\tau)$ can be calculated approximately using the formulas

$$m \sim \frac{1}{N} \sum_{k=1}^{N} X(k\Delta); \quad R(\tau) \sim \frac{1}{N} \sum_{k=1}^{N} X(k\Delta+\tau)\, X(k\Delta),$$

where Δ is a small time interval and N is so chosen so as to make $N\Delta = T$ sufficiently large. In practice, Δ should be chosen in such a way that the realization $X(t)$ does not change appreciably during time intervals of length Δ, while N should be such so that a further increase in the number of terms has only a slight influence on the value of the average being calculated. The above formulae, for most realizations, represent a formidable computational task. This task has been greatly simplified by using computers, some of which have been specifically designed to perform such calculations.

We now discuss the conditions that must be imposed on the stochastic process $\{X(t),\ t \in T\}$ for the ergodic theorem stated above to hold. Let

$$B(\tau) = E[(X(t+\tau) - m)(X(t) - m)] = R(\tau) - m^2.$$

$B(\tau)$ is the centred correlation function and differs from the correlation coefficient of $X(t+\tau)$ and $X(t)$ by a factor $B(0)$. It was shown by Slutski (35) that (1) holds, if and only if,

(a) $$\lim_{T\to\infty} \frac{1}{T} \int_0^T B(\tau)d\tau = 0 \qquad \text{and}$$

(b) $$\lim_{N\to\infty} \frac{1}{N+1} \sum_{\tau=0}^{N} B(\tau) = 0 \qquad \text{respectively .}$$

For (2) to hold, the function

$$B_1(\tau) = E[X(t+\tau_o + \tau)\, X(t+\tau) - R(\tau_o)][X(t+\tau_o)\, X(t) - R(\tau_o)]$$

must satisfy condition (a) or (b).

In practice, the functions $B(\tau)$ and $B_1(\tau)$ usually approach 0 as $\tau \to \infty$. In this case, the above conditions are met. It should be noted, that $B_1(\tau)$ involves fourth-order moments of $X(t)$ and hence, the conditions for the validity of the ergodic theorem cannot be expressed in terms of the correlation function.

5.3 Correlation Analysis of Point Processes

There are, generally speaking, two ways in which one can describe a stationary point process. We can consider the sequence of intervals $(X_1, X_2, \ldots)$ between successive events or we may divide the time axis into a large number of narrow intervals of width Δt and count the number of events in each interval. The two specifications are equivalent, but lead of course to different correlation functions.

First, let us consider the sequence of intervals $(X_1, X_2, \ldots)$. In order that the sequence should be stationary, we suppose that an event occurs at $t = 0$. Let $m_x = E(X_i)$ and $\sigma_x^2 = \mathrm{Var}(X_i)$. If we consider the sequence as a stationary real-valued process, we can define

$$B_x(\tau) = \mathrm{Cov}(X_{i+\tau}, X_i) = E(X_{i+\tau}, X_i) - (m_x)^2.$$

We can also consider a variance function $\psi_x(\tau) = \mathrm{Var}(X_1 + X_2 + \ldots + X_\tau)$, specifying the variance of the time interval from an arbitrary event to the τ^{th} successive event.

We have that

$$\Psi_x(\tau) = \mathrm{Var}(X_1 + X_2 + \ldots + X_\tau)$$

$$= \sum_{i=1}^{\tau} \mathrm{Var}(X_i) + 2 \sum_{j=1}^{\tau-1} \sum_{i=1}^{j} B_x(i)$$

$$= \tau B_x(o) + 2 \sum_{j=1}^{\tau-1} \sum_{i=1}^{j} B_x(i) .$$

The functions $\Psi_x(\tau)$ and $B_x(\tau)$ are mutually equivalent. For some purposes, it is convenient to define the index of dispersion

$$C_x(\tau) = \frac{\Psi_x(\tau)}{\tau m_x^2} .$$

Let us now consider, formally, the process $\Delta N(t) = N(t+\Delta t) - N(t)$ [the number of events in $(t, t + \Delta t)$]. Suppose the mean rate of occurrence is λ, so that $\Pr[\text{one event in } (t, t + \Delta t)] = \lambda\Delta t + o(\Delta t)$ for all t. We are here assuming that multiple occurrences do not arise. We define the covariance density for $\tau > o$ of the point process by the equation

$$\mathrm{Cov}\,[\Delta N(t+\tau), \Delta N(t)] = \gamma_n(\tau)(\Delta t)^2$$

or alternatively

$$\gamma_n(\tau) = \lim_{\Delta t \to o} \frac{\mathrm{Cov}[N(t+\Delta t) - N(t), N(t+\tau+\Delta t) - N(t+\tau)]}{(\Delta t)^2} .$$

The second order properties of this process in continuous time can also be defined by the intensity function

$$I_n(\tau) = \lim_{\Delta t \to o} \frac{\Pr[\text{event in } (t+\tau, t+\tau+\Delta t) \mid \text{event at } t]}{\Delta t} .$$

Now

$$E[\Delta N(t+\tau)\, \Delta N(t)] = \Pr[\Delta N(t) = \Delta N(t+\tau) = 1]$$

$$= \lambda\Delta t \Pr[\Delta N(t+\tau) = 1 \mid \Delta N(t) = 1]$$

$$= \lambda I_n(\tau)(\Delta t)^2 .$$

Since $E[\Delta N(t)] = \lambda \Delta t$, we have that $\gamma_n(\tau) = \lambda [I_n(\tau) - \lambda]$. For all practical purposes then, we can estimate either the covariance density or the intensity function.

Now for a Poisson process

$$I_n(\tau) = \lim_{\Delta t \to o} \frac{\lambda \Delta t + o(\Delta t)}{\Delta t}$$

$$= \lambda, \text{since the no. of events}$$

in $(t+\tau, t+\tau+\Delta t)$ is independent of the number of events at t.

For $\tau < o$, we can define $\gamma_n(\tau) = \gamma_n(-\tau)$, but for $\tau = o$ it is important to note that

$$E[\Delta N(t)\ \Delta N(t)] = E[\Delta N(t)]^2 = E[\Delta N(t)].$$

It is convenient to add to $\gamma_n(\tau)$ previously defined the term $\lambda\delta(\tau)$, where $\delta(\tau)$ is the Dirac delta function. We define thus the complete covariance density

$$\gamma_n^c(\tau) = \lambda\delta(\tau) + \gamma_n(\tau).$$

5.4 Estimation of the Intensity Function

Suppose the process is observed over an interval $[0,T]$ starting from an arbitrary time. Let N(t) be the sample counting function having N(0) = 0 and jumping by one at each event. Let n be the total number events observed.

Let τ be some grouping interval. For a formal study of limiting properties, we should allow τ to tend to zero as $T \to \infty$. However, we shall regard τ as fixed and estimate the grouped intensities.

$$I_n^g(r\tau + \tfrac{1}{2}\tau) = \frac{1}{\tau} \int_{r\tau+0}^{r\tau+\tau} I_n(x)dx \qquad r = 0, 1, \ldots .$$

In practice it is usually advisable to examine a number of values of τ.

We shall base the estimate of $I_n^{g}(r\tau + \frac{1}{2}\tau)$ on the total number of pairs of events, S_r, separated by an interval between $r\tau$ and $r\tau+\tau$. In principle, we take the $\frac{1}{2}n(n-1)$ positive intervals between all possible ordered pairs of events and form them into histograms with groups of width τ. Formally, we can write

$$S_r = \int_{u=0}^{T-r} \int_{x=r\tau}^{r\tau+\tau} dN(u)\, dN(u+x)$$

$$= \int_{u=0}^{T-r\tau-\tau} \int_{x=r\tau}^{r\tau+\tau} dN(u)\, dN(u+x) + \int_{u=T-r\tau-\tau}^{T-r\tau} \int_{x=r\tau}^{T-u} dN(u)\, dN(u+x) .$$

Since $$E[dN(u)] = \lambda du$$ and

$$E[dN(u)\, dN(u+x)] = \lambda\, I_n(x)\, du\, dx ;$$

the expectation of the first double integral is

$$\lambda \int_{u=0}^{T-r\tau-\tau} du \int_{x=r\tau}^{r\tau+\tau} I_n(x)dx = \lambda\tau(T-r\tau-\tau)\ I_n^{g}(r\tau+\tfrac{1}{2}\tau).$$

When $\frac{\tau}{T-r\tau}$ is small, as is always the case in practice, the second integral is only a small correction term and can be approximated by assuming $I_n(x)$ to be constant over the relevant range and equal to $I_n^{g}(r\tau+\frac{1}{2}\tau)$. Then the expectation of the second integral is

$$\lambda \int_{u=T-r\tau-\tau}^{T-r\tau} \int_{x-r\tau}^{T-u} I_n(x)\, du\, dx$$

$$\sim \lambda \int_{u=T-r\tau-\tau}^{T-r\tau} \frac{\tau}{2}\ I_n^{g}(r\tau+\tfrac{1}{2}\tau)du$$

$$= \frac{\lambda\tau^2}{2}\ I_n^{g}(r\tau+\tfrac{1}{2}\tau).$$

Hence with this approximation

$$E(S_r) = \lambda\tau(T-r\tau-\tfrac{1}{2}\tau)\ I_n^{g}(r\tau+\tfrac{1}{2}\tau) .$$

Since $E(\frac{n}{T}) = \lambda$, we are led to define

$$\hat{m}_g(r\tau+\tfrac{1}{2}\tau) = \frac{S_r\, T}{n\tau(T-r\tau-\frac{\tau}{2})} .$$

The effective length of the series decreases with increasing r, thus providing a qualitative reason for the factor $\frac{T}{T-r\tau-\frac{\tau}{2}}$.

Since the total number of pairs of intervals is $\frac{1}{2}n(n-1)$, we have that

$$\sum_{r=0}^{\frac{T}{\tau}-1} S_r = \tfrac{1}{2}n(n-1) .$$

To avoid unpleasant end effects, we assume that τ is so chosen so as to make $\frac{T}{\tau}$ an integer. Thus, we have

$$\sum_{r=o}^{\frac{T}{\tau}-1} \hat{m}_g(r\tau+\tfrac{1}{2}\tau)\, \frac{n\tau(T-r\tau-\frac{1}{2}\tau)}{T} = \tfrac{1}{2}n(n-1)$$

or $$\tau \sum_{r=o}^{\frac{T}{\tau}-1} w_r\, \hat{m}_g(r\tau+\tfrac{1}{2}\tau) = \frac{n-1}{2} \quad , \text{where } w_r = \frac{T-r\tau-\frac{\tau}{2}}{T}$$

and $$\sum_{r=o}^{\frac{T}{\tau}-1} w_r\, \hat{m}_g(r\tau+\tfrac{1}{2}\tau) = \frac{n-1}{2\tau}$$

and $$\sum_{r=o}^{\frac{T}{\tau}-1} \frac{w_r\, \hat{m}_g(r\tau+\tfrac{1}{2}\tau)}{\sum w_r} = \frac{n-1}{2\tau} \cdot \frac{2\tau}{T} = \frac{n-1}{T} \sim \frac{n}{T} = \hat{\lambda}$$

where $\hat{\lambda}$ is the estimator of λ. The last equation implies that if the sample function is fluctuating about an average independent of r, that average must be nearly $\hat{\lambda}$.

5.5 Properties of the Intensity Function for the Poisson Process

For the Poisson process, we can consider the n events to be independently and uniformly distributed over the interval $[0,T]$. Let $U_1, U_2, \ldots, U_n$ be independently, rectangularly distributed over $[0,T]$. Let

$$V_{ij}^{(r)} = 1 \quad \text{if} \quad r\tau \leq |U_i - U_j| < r\tau+\tau$$
$$= 0 \text{ otherwise.}$$

Then, we have that

$$S_r = \sum_{j>i} V_{ij}^{(r)} .$$

We can obtain the mean and variance of S_r from the conditional mean of $V_{ij}^{(r)}$ given U_i. If $2r\tau+\tau < T$, then we have $E\left[V_{ij}^{(r)} \mid 0 \leq U_i < r\tau\right] = \frac{\tau}{T}$, since it is easily verified that if $0 \leq U_i = u_i < r\tau$, then $r\tau + u_i \leq U_j < r\tau+\tau+ u_i$ and the result follows from independence of U_i and U_j. Similarly, we obtain

$$E\left[V_{ij}^{(r)} \mid r\tau \leq U_i = u_i < r\tau +\tau\right] = \frac{\tau}{T} + \frac{u_i - r\tau}{T}$$

$$E\left[V_{ij}^{(r)} \mid r\tau+\tau \leq U_i < T - r\tau -\tau\right] = \frac{2\tau}{T}$$

$$E\left[V_{ij}^{(r)} \mid T - r\tau -\tau \leq U_i = u_i < T - r\tau\right] = \frac{\tau}{T} + \frac{u_i-r\tau}{T}$$

$$E\left[V_{ij}^{(r)} \mid T - r\tau \leq U_i < T\right] = \frac{\tau}{T} .$$

Now, the unconditional mean of $V_{ij}^{(r)}$ can be obtained by integrating with respect to the uniform distribution of U_i, i.e.,

$$E\left[V_{ij}^{(r)}\right] = \int_{0}^{r\tau} \frac{\tau}{T} \cdot \frac{1}{T}\, du + \int_{r\tau}^{r\tau+\tau} \frac{1}{T}\left(\frac{\tau}{T} + \frac{u-r\tau}{T}\right) du$$

$$+ \int_{r\tau+\tau}^{T-r\tau-\tau} \frac{2\tau}{T} \cdot \frac{1}{T}\, du + \int_{T-r\tau-\tau}^{T-r\tau} \left(\frac{\tau}{T} + \frac{u-r\tau}{T}\right) \cdot \frac{1}{T}\, du$$

$$+ \int_{T-r\tau}^{T} \frac{\tau}{T} \cdot \frac{1}{T}\, du \quad .$$

And $$E\left[V_{ij}^{(r)}\right] = \frac{2\tau\left[T - r\tau - \frac{\tau}{2}\right]}{T^2} \quad .$$

Thus, $$E(S_r) = \sum_{j > i} E\left[V_{ij}^{(r)}\right]$$

$$= \sum_{j > i} 2\tau\left(\frac{T - r\tau - \frac{\tau}{2}}{T^2}\right)$$

$$= \frac{n(n-1)}{2}\, 2\tau\left(\frac{T - r\tau - \frac{\tau}{2}}{T^2}\right)$$

$$= \frac{n(n-1)\tau\,(T - r\tau - \frac{\tau}{2})}{T^2} \quad .$$

Finally $$E\left[\hat{m}_g(r\tau + \tfrac{1}{2}\tau)\right]$$

$$= \frac{E(S_r)T}{n\tau(T-r\tau-\frac{\tau}{2})}$$

$$= \frac{n-1}{T} \quad .$$

Since $V_{ij}^{(r)}$ is a (0,1) random variable

$$E\left[V_{ij}^{(r)2}\right] = E\left[V_{ij}^{(r)}\right] = \frac{2\tau\left[T - r\tau - \frac{\tau}{2}\right]}{T^2} \quad .$$

Conditionally on U_i, the random variables $V_{ij}^{(r)}$ and $V_{il}^{(r)}$ $(j \neq l)$ are independent, since U_j and U_l are independent.

Therefore,
$$E\left[V_{ij}^{(r)}\ V_{il}^{(r)} \mid U_i = u_i\right]$$
$$= E\left[V_{ij}^{(r)} \mid U_i = u_i\right]\ E\left[V_{il}^{(r)} \mid U_i = u_i\right].$$

These can be computed directly for the various values of U_i from the formulae above. By integrating out U_i with respect to its uniform distribution, we obtain

$$E\left[V_{ij}^{(r)}\ V_{il}^{(r)}\right] = \frac{4\tau^2}{T^2} - \frac{6r\tau^3}{T^3} - \frac{10\tau^3}{3T^3}, j \neq 1 .$$

Now, if i, j, l, k are different, then $V_{ij}^{(r)}$ and $V_{lk}^{(r)}$ are independent. Hence,
$$E(S_r^{\,2}) = E\left[\sum_{j>i} V_{ij}^{(r)}\right]^2$$
$$= \tfrac{1}{2}n(n-1)\ E\left[(V_{ij}^{(r)})^2\right]$$
$$+ n(n-1)(n-2)\ E\left[V_{ij}^{(r)}\ V_{il}^{(r)}\right]$$
$$+ \tfrac{1}{4}n(n-1)(n-2)(n-3)\left[E(V_{ij}^{(r)})\right]^2 .$$

Upon substitution from above, we obtain, after a great deal of tedious algebra,

$$\mathrm{Var}(S_r) = E(S_r^{\,2}) - \left[E(S_r)\right]^2$$
$$= \frac{n(n-1)\tau(T - r\tau - \frac{5}{2}\tau)}{T^2} + n(n-1)(2rn + \tfrac{2}{3}n + \tfrac{2}{3})\frac{\tau^3}{T^3}$$
$$- n(n-1)(4n-6)(r + \tfrac{1}{2})^2\frac{\tau^4}{T^4} .$$

Finally, we obtain

$$\mathrm{Var}\left[\hat{m}_g(r\tau + \tfrac{1}{2}\tau)\right] = \frac{T^2\ \mathrm{Var}(S_r)}{n^2\tau^2(T - r\tau - \frac{1}{2}\tau)^2} .$$

For many purposes, an adequate approximation to the above is

$$\frac{n-1}{n\tau(T - r\tau + \frac{\tau}{2})} .$$

Thus, we can take as a test of the hypothesis that the series is Poisson, the variable

$$\sum_{r=0}^{s-1} \frac{\left[\hat{m}_g(r\tau + \tfrac{1}{2}\tau) - \frac{n-1}{T}\right]^2}{\mathrm{Var}\left[\hat{m}_g(r\tau + \tfrac{1}{2}\tau)\right]}$$

which would be, to a first approximation, distributed as χ^2 with s degrees of freedom and perform the usual test. We mention that as $n \to \infty$, the S_r are asymptotically normal. This follows from a result of Hoeffding (16).

CHAPTER VI

Spectral Analysis of Point Process

6.1 Introduction

Previously we have discussed correlation analysis and the estimation of the intensity function for point processes. Generally speaking, correlation analysis is useful in determining departures from the Poisson or purely random series and in the study of superimposed processes. The correlation function is a natural one for considering the evolution of processes in time. When the practical effect of the process is more conveniently measured separately for the different harmonic components, it is then more natural to use spectral analysis.

For the sake of completeness, we shall give a brief intuitive discussion of spectral analysis for stationary processes. The specification to point processes will then be made.

6.2 Stationary Processes

In the context of harmonic analysis, it is convenient and gives some extra generality, to allow complex valued processes. This enables us to deal directly, for example, with A.C. signals. Let $X_{(1)}$, $X_{(2)}$, $Y_{(1)}$, $Y_{(2)}$ be real valued random variables.

Let $$X = X_{(1)} + i\, X_{(2)}$$
$$Y = Y_{(1)} + i\, Y_{(2)}.$$

We then define

$$E(X) = E[X_{(1)}] + i\, E[X_{(2)}]$$
$$E(Y) = E[Y_{(1)}] + i\, E[Y_{(2)}].$$

If we suppose $E(X) = E(Y) = 0$, we can then define the covariance of X and Y by $Cov(X,Y) = E(X,\overline{Y})$ where $\overline{Y}$ denotes the complex conjugate of Y. Then $Var(X) = E(X,\overline{X}) = E[X_{(1)}^2 + X_{(2)}^2]$ is real and positive. The physical justification for these definitions is that, if $X_{(1)} + i\, X_{(2)}$ represents an A.C. signal, then Var(X) is proportional to the mean power generated.

Now, if $X = X_{(1)} + i\, X_{(2)}$ and $Y = Y_{(1)} + i\, Y_{(2)}$, then

$Cov(X,Y) = E[X_{(1)}\, Y_{(1)} + X_{(2)}\, Y_{(2)}] + i\, E[X_{(2)}\, Y_{(1)} - X_{(1)}\, Y_{(2)}]$,

so that $Cov(X,Y) = 0$ does not imply that the components are separately uncorrelated. We shall say that if $Cov(X,Y) = 0$, then X and Y are orthogonal random variables.

Generally speaking,

$$Var(X + Y) = Var(X) + Var(Y) + Cov(X,Y) + Cov(Y,X)$$

so that if $Cov(X,Y) = Cov(Y,X) = 0$

then $Var(X + Y) = Var(X) + Var(Y)$.

We now consider, for a fixed ω, processes of the type $Re^{i\omega t}$ $(-\infty < t < \infty)$ in continuous time and $Re^{i\omega n}$ $(n = \ldots, -1, 0, 1, \ldots)$ in discrete time, where R is a complex-valued random variable. In continuous time ω can be any real number, but in the discrete case, there is no loss of generality in taking $-\pi \leq \omega < \pi$, since for all n and integral k, $e^{i(\omega + 2k\pi)n} = e^{i\omega n}$.

We now examine the conditions under which $Re^{i\omega n}$ represents a second-order (weakly) stationary process. Obviously $E(Re^{i\omega n})$ is

independent of n if and only if $E(R) = 0$ and we shall assume this.

$$\therefore \qquad E\left[Re^{i\omega(n+k)}\ \overline{Re}^{-i\omega n}\right] = e^{i\omega k}\ E(R\ \overline{R}).$$

The above is independent of n and a stationary correlation function exists provided $E(R\ \overline{R}) < \infty$.

It can be easily shown that $\sum_{i=1}^{k} R_i e^{j\omega_i n}$ represents a stationary process, if and only if, $E(R_i) = 0$ and $Cov(R_i, R_j) = 0$ for $j \neq i$. The correlation function of the sum is then

$$\sum_{i=1}^{k} E(R_i\ \overline{R}_i).$$

Thus the variance of R_i can be regarded as determining the contribution of the component at frequency ω_i to the total variance of the process.

Suppose we formally let $k \to \infty$ in the above sum. In discrete time, we take a set of ω_j's covering the interval $[-\pi, \pi)$. In continuous time we cover, in the limit, the whole real axis. The limiting form of the sum is obtained by considering

$$X_n = \int_{-\pi-o}^{\pi} e^{i\omega n}\, dS(\omega), \text{ in discrete time } (n = \ldots, -1, 0, 1, \ldots)$$

or

$$X(t) = \int_{-\infty}^{\infty} e^{i\omega t}\, dS(\omega) \qquad (-\infty < t < \infty)$$

where $S(\omega)$ is now a stochastic process defined over $[-\pi, \pi)$ in discrete time and over $(-\infty, \infty)$ in the continuous case. At points of discontinuity, we take (by convention) $S(\omega)$ to be continuous on the right.

For example, if ω_j's are to be equally spaced over $[-\pi, \pi)$ with $\omega_j = j\Delta\omega$, we can consider $X_n = \int_{-\pi-o}^{\pi} e^{i\omega n}\, dS(\omega)$ to be the limit as $\Delta\omega \to o$ of

$$\sum_{j=-\frac{\pi}{\Delta\omega}}^{\pi/\Delta\omega} e^{i\omega_j n} R_j \qquad \text{where}$$

$$R_j = \int_{\omega_j - \frac{1}{2}\Delta\omega}^{\omega_j + \frac{1}{2}\Delta\omega} dS(\omega) \ .$$

Now the requirements for R_j so that the sum above defines a stationary process is that $E(R_j) = o$ and $Cov(R_j, \bar{R}_1) = o$, $j \neq 1$. Thus,

$$E\left[\int_{\omega_1}^{\omega_2} dS(\omega)\right] = o \text{ and } E\left[\int_{\omega_1}^{\omega_2} dS(\omega) \int_{\omega_3}^{\omega_4} d\bar{S}(\omega)\right] = o$$

$$(\omega_1 < \omega_2 < \omega_3 < \omega_4).$$

We call $S(\omega)$ a process with orthogonal increments. For $\omega_1 < \omega_2$, we have

$$\begin{aligned} Var[S(\omega_2)] &= Var\left[\int_{-\pi}^{\omega_1} dS(\omega) + \int_{\omega_1}^{\omega_2} dS(\omega)\right] \\ &= Var[S(\omega_1)] + Var\left[\int_{\omega_1}^{\omega_2} dS(\omega)\right] . \end{aligned}$$

Thus, $Var[S(\omega)]$ is a non-decreasing real valued function of ω, say $G(\omega)$. $G(\omega)$ will be continuous on the right. This function completely specifies the second order properties of $S(\omega)$ and $Var(X_n) = Var\left[\int_{-\pi}^{\pi} dS(\omega)\right] = G(\pi)$ in discrete time, and $Var[X(t)] = G(\infty)$ in continuous time, where we have taken $G(-\pi - o) = o$ and $G(-\infty) = o$. If we define $F(\omega) = \frac{G(\omega)}{\sigma_x^2}$, it has the mathematical properties of a distribution function defined over $[-\pi, \pi)$ or $(-\infty, \infty)$. We call $F(\omega)$ the spectral distribution function. Now $F(\omega_o)$ is the proportion of the total variance of the process contributed by the harmonic components

$\omega \leq \omega_o$. In other words, the spectral distribution function specifies how the total variance of $[X_n]$ or $[X(t)]$ is sub-divided among the orthogonal components making up the representations of these processes. Thus, we may consider processes having the structure

$$X_n = \int_{-\pi}^{\pi} e^{i\omega n}\, dS(\omega) \text{ or } X(t) = \int_{-\infty}^{\infty} e^{i\omega t}\, dS(\omega)$$

with $S(\omega)$ having orthogonal increments. The second-order properties of $S(\omega)$ are specified by a function $G(\omega)$, which is the total variance σ_x^2 of the process and by the spectral distribution function

$$F(\omega) = \frac{G(\omega)}{\sigma_x^2}.$$

We summarize the properties of processes defined by the above equations.

$$(1)\ E[dS(\omega)] = o$$

$$(2)\ \mathrm{Cov}[dS(\omega_1), dS(\omega_2)] = E[dS(\omega_1)\ d\bar{S}(\omega_2)] = o,\ \omega_1 \neq \omega_2$$

$$(3)\ \mathrm{Var}[dS(\omega)] = E[|dS(\omega)|^2] = dG(\omega).$$

In discrete time, $\mathrm{Var}[X_n] = G(\pi)$, i.e.,

$$\mathrm{Var}(X_n) = E[X_n \bar{X}_n] = E\left[\int_{-\pi}^{\pi} e^{i\omega_1 n}\, dS(\omega_1) \int_{-\pi}^{\pi} e^{-i\omega_2 n}\, d\bar{S}(\omega_2)\right]$$

$$= \int_{-\pi}^{\pi} dG(\omega)$$

since pairs $\omega_1 \neq \omega_2$ contribute zero to the expectation.

We can obtain the autocorrelation function

$$R(h) = E(X_{n+h} \bar{X}_n) = E\left[\int_{-\pi}^{\pi} e^{i\omega_1 (n+h)}\, dS(\omega_1) \int_{-\pi}^{\pi} e^{-i\omega_2 n}\, d\bar{S}(\omega_2)\right]$$

$$= \int_{-\pi}^{\pi} e^{i\omega h}\, dG(\omega)$$

since only contributions where $\omega_1 = \omega_2$ need be considered. Thus the function $G(\omega)$ or $F(\omega)$, the spectral distribution function, has two interpretations. It gives the probabilistic properties of the components in a Fourier analysis of the process itself. It also gives directly the components in a Fourier analysis of the autocorrelation function.

Suppose $F(\omega)$ is a discrete distribution. Then, there is a finite or denumerable set of values $\omega_1, \omega_2, \ldots$ at which $F(\omega)$ has positive increments $f_1, f_2, \ldots$ with $\sum f_i = 1$. Then we can write $X_n = \sum e^{i\omega_j n} R_j$ where

$$E(R_j) = 0; \quad \mathrm{Var}(R_j) = f_j \sigma_x^2 \text{ and } \mathrm{Cov}(R_j, R_l) = 0 \ , \ l \neq j.$$

Such a process has a discrete spectrum. We have also

$$\sigma_x^2 R(h) = \sum_j e^{i\omega_j h} f_j \quad \text{and}$$

$$\frac{f_j}{\sigma_x^2} = \lim_{a \to \infty} \frac{1}{2a} \sum_{h=-a}^{a} e^{-i\omega_j h} R(h) \ ,$$

the limit being zero when ω_j is replaced by a value of ω not in the set $\omega_1, \omega_2, \ldots$. Note that the f_j's can not be determined from a single realization no matter how long. All we can determine, from analysis of a single realization, are the values $R_1, R_2, \ldots$.

Suppose $F(\omega)$ is absolutely continuous with a spectral density function $f(\omega)$. Then we have

$$\sigma_x^2 R(h) = \int_{-\pi}^{\pi} e^{i\omega h} f(\omega) d\omega \ .$$

Therefore

$$\frac{f(\omega)}{\sigma_x^2} = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-i\omega h} R(h) \qquad \text{or}$$

$$\sigma_x^2 R(h) = \int_{-\infty}^{\infty} e^{i\omega h} f(\omega) d\omega \qquad \text{and}$$

$$\frac{f(\omega)}{\sigma_x^2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega h} R(h)dh$$

i.e., the correlation function and spectral density function are Fourier transform pairs.

6.3 Spectral Analysis of Point Processes

For point processes, there are several ways to proceed if one wants to analyze the process by means of the spectrum. Generally speaking, we have available both N(T), the number of events in an interval T, the sequence of intervals $X_1, X_2, \ldots, X_n$ between events.

First of all, we shall consider the sequence of intervals $X_1, X_2, \ldots, X_n$. These can be thought of as a stochastic process in discrete time and can be handled by standard methods, of which we give a summary. First, we estimate the correlation function R(h) by means of the sample correlation function

$$C_h = \frac{1}{n-|h|} \sum_{t=1}^{n-|h|} X_t X_{t+|h|} .$$

The C_h are computed after removing the trend in the sample due to the mean, i.e., we assume $E(X_t) = 0$.

We then estimate the spectral density by means of the periodogram

$$\mathcal{I}_n(\omega_i) = 2 \sum_{h=-(n-1)}^{n-1} (1 - \frac{|h|}{n}) C_h \cos \omega_i h .$$

It is well known (Jenkins (17)) that the periodogram $\mathcal{I}_n(\omega)$ provides an estimate of the spectral density $f(\omega)$, but since

$$\mathrm{Var}[\mathcal{I}_n(\omega)] \simeq [2\sigma^2 f(\omega)]^2$$ it follows

that the estimate is not even consistent.

Bartlett (1 , 2) suggested estimates of the following form

$$f_s(\omega) = \pi^{-1}\left[\, 1 + 2\sum_{s=1}^{m-1}(1 - \tfrac{s}{m})r_s\cos\omega s \,\right]$$

where $r_s = \frac{C_s}{C_0}$ are the sample estimates of the normalized autocorrelations. This estimate, which uses the first $m < n$ autocorrelations, is known as the truncated and smoothed periodogram, and is the most commonly used today.

Grenander and Rosenblatt, p.58(7) suggested a general class of estimates of the form

$$f_G(\omega) = \pi^{-1}\left[\, 1 + 2\sum_{s=1}^{n-1}\lambda_s(\omega)r_s\cos\omega s \,\right]$$

or equivalently

$$f_G^*(\omega) = \pi^{-1}\int_0^{\pi} \mathcal{J}_n(x)\,\mu_\omega(x)dx$$

where the weighting functions $\lambda_s(\omega)$ and $\mu_\omega(x)$ are chosen so as to make the estimates consistent.

In actual practice, these estimates will be used in cases where the true spectrum is initially at most roughly known, and the departure from this supposedly true spectrum is the problem of interest. The choice of suitable weight functions is not a critical step and the main problem is to choose a suitable value for the number of correlation coefficients to be estimates, say ℓ. Now, providing the true spectrum is relatively smooth over the bandwidth d for which we want to resolve points of the spectrum, then we should chose ℓ so that $\ell > \frac{2\pi}{d}$. The various weighting functions commonly used and some of their properties

are summarized in the table below.

Spectral Weight Functions

	λ_s	μ_ω	$\frac{\text{Variance}}{f^2(\omega)}$
(1) Daniell	$\frac{\sin s\varepsilon}{\lvert s\varepsilon\rvert}$, $(\lvert s\rvert \lesssim n)$	$\frac{1}{2\varepsilon}$, $(-\varepsilon \lesssim \omega \lesssim \varepsilon)$ 0, otherwise	$\frac{\pi}{2\varepsilon}$
(2) Bartlett	$1 - \frac{\lvert s\rvert}{\mathfrak{k}}$, $(\lvert s\rvert \lesssim \mathfrak{k})$ 0, $(\lvert s\rvert > \mathfrak{k})$	$\frac{1}{2\pi\mathfrak{k}}\left[\frac{\sin\frac{1}{2}\mathfrak{k}\omega}{\sin\frac{1}{2}\omega}\right]^2$	$\frac{2\mathfrak{k}}{3n}$
(3) Tukey	$\frac{1}{2}\left[1 + \frac{\cos\pi s}{\mathfrak{k}}\right]$, $\lvert s\rvert \lesssim \mathfrak{k}$ 0, $\lvert s\rvert > \mathfrak{k}$	$\frac{1}{4\pi}\left\{\frac{\sin(\mathfrak{k}+\frac{1}{2})\omega}{\sin\frac{1}{2}\omega} + \frac{1}{2}\left[\frac{\sin[(\mathfrak{k}+\frac{1}{2})\omega+\frac{\pi}{\mathfrak{k}}]}{\sin\frac{1}{2}(\omega+\frac{\pi}{\mathfrak{k}})} + \frac{\sin[(\mathfrak{k}+\frac{1}{2})(\omega-\frac{\pi}{\mathfrak{k}})]}{\sin\frac{1}{2}(\omega-\frac{\pi}{\mathfrak{k}})}\right]\right\}$	$\frac{3\mathfrak{k}}{4n}$
(4) Parzen	$1 - \frac{6s^2}{\mathfrak{k}^2(1-\frac{\lvert s\rvert}{\mathfrak{k}})}$, $\lvert s\rvert \lesssim \frac{\mathfrak{k}}{2}$ $2(1-\frac{\lvert s\rvert}{\mathfrak{k}})^3$, $\frac{\mathfrak{k}}{2} < \lvert s\rvert \lesssim \mathfrak{k}$ 0, $\lvert s\rvert > \mathfrak{k}$	$\frac{3}{8\pi\mathfrak{k}^3}\left[\frac{\sin\frac{1}{4}\mathfrak{k}\omega}{\sin\frac{1}{4}\omega}\right]^4$	$\frac{\mathfrak{k}}{2n}$

n = total no. of observations

$\mathfrak{k}$ = effective lag, i.e., no. of correlations computed.

It is a well-known fact in spectral analysis that there is a sort of indeterminacy between resolvability and statistical reliability as measured by the variance of the estimator. From the above calculated values of the variances of the estimators (approximate), we see that increasing the value of λ increases the variance but also enables us to resolve more clearly between power at different frequences. On the other hand, decreasing the value of λ decreases the variance but decreases the resolvability. The attempt at choosing the various weight functions then is an attempt to suitably balance resolvability and statistical reliability for a given class of theoretical estimates.

6.4 Estimating the Spectral Density for a Counting Process

The complete covariance density for a counting process, $\Delta N(t)$, was defined in section 5.3 as

$$\gamma_n^c(\tau) = \lambda\delta\tau + \gamma_n(\tau) \text{ where}$$

$$\gamma_n(\tau) = \lim_{\Delta t \to 0} \frac{\mathrm{Cov}\left[N(t+\Delta t)-N(t),\ N(t+\tau+\Delta t)-N(t+\tau)\right]}{(\Delta t)^2}.$$

We can then define a complete spectral density function as the Fourier transform of

$$\gamma_n^c(\tau), \text{ i.e., } g(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \lambda\delta(\tau)e^{-i\omega\tau}d\tau + \frac{1}{2\pi}\int_{-\infty}^{\infty} \gamma_n(\tau)e^{-i\omega\tau}d\tau.$$

Since $\gamma_n(\tau) = \lambda\left[I_n(\tau) - \lambda\right]$, then we have

$$g(\omega) = \frac{\lambda}{2\pi} + \frac{\lambda}{2\pi}\int_{-\infty}^{\infty}\left[I_n(\tau) - \lambda\right]e^{-i\omega\tau}d\tau.$$

It is convenient to define a spectral density for non-negative ω by

$$g_+(\omega) = 2g(\omega) = \frac{\lambda}{\pi} + \frac{\lambda}{\pi} \int_{-\infty}^{\infty} \left[I_n(\tau) - \lambda \right] e^{-i\omega\tau} d\tau .$$

Since we are considering only real stationary processes, $I_n(\tau) = I_n(-\tau)$ and $I_n(\tau) \to \lambda$ as $\tau \to \infty$. A very useful alternative form for $g_+(\omega)$ is then

$$g_+(\omega) = \frac{\lambda}{\pi} \left[1 + I_n^*(i\omega) + I_n^*(-i\omega) \right]$$

where $I_n^*(s)$ is the ordinary (one-sided) Laplace transform of $I_n(\tau)$. In turn, $\gamma_n(\tau)$ is obtained from $g_+(\omega)$ by the usual inversion formula for Fourier transforms.

We can thus write

$$g_+(\omega) = \frac{\lambda}{\pi} \left[1 + \int_0^{\infty} e^{it\omega} I_n(t)dt + \int_0^{\infty} e^{-it\omega} I_n(t)dt \right]$$

which is the most useful form for our present purposes.

We now define

$$H_{t_o}(\omega) = \frac{1}{\sqrt{\pi t_o}} \sum_{t=o}^{t_o} e^{it\omega} dN(t)$$

$$= \frac{1}{\sqrt{\pi t_o}} \sum_{s=1}^{n} e^{iT_s\omega}$$

where $T_1, T_2, \ldots, T_n$ are the times in t_o at which an event occurs.

$$\therefore H_{t_o}(\omega) = \frac{1}{\sqrt{\pi t_o}} \left[\sum_{s=1}^{n} \cos(T_s\omega) + i \sum_{s=1}^{n} \sin(T_s\omega) \right]$$

We can thus define for the process $\{\Delta N(t)\}$ the periodogram

$$\mathcal{I}_{t_o}(\omega) = H_{t_o}(\omega)\, \overline{H}_{t_o}(\omega) \text{ where}$$

the bar denotes complex conjugation.

Thus

$$\mathcal{J}_{t_o}(\omega) = \frac{1}{\pi t_o} \sum_{s=1}^{n} \sum_{k=1}^{n} e^{i\omega(T_s - T_k)}$$

$$= \frac{1}{\pi}\left[\frac{n}{t_o} + \frac{1}{t_o} \sum_{s=1}^{n-1} \sum_{j=1}^{n-s} e^{i\omega(T_{s+j} - T_j)} + \frac{1}{t_o} \sum_{s=1}^{n-1} \sum_{j=1}^{n-s} e^{-i\omega(T_{s+j} - T_j)}\right].$$

If we define

$$\tilde{I}_n(t) = \frac{1}{n} \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \delta(t_{i+j} - t_i - t)$$

as an estimator of $I_n(t)$ where $\delta(t)$ is the Dirac delta function, and recalling that the Fourier transform of a Dirac delta function is

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \delta(a-t) e^{i\omega t} dt = \frac{1}{2\pi} e^{i\omega a},$$

we have that

$$\mathcal{J}_{t_o}(\omega) = \frac{n}{t_o \pi}\left[1 + \int_o^{\infty} e^{it\omega} \tilde{I}_n(t)dt + \int_o^{\infty} e^{-it\omega} \tilde{I}_n(t)dt\right].$$

Since $\frac{n}{t_o}$ is an estimator of λ, we see that $\mathcal{J}_{t_o}(\omega)$ is indeed an estimator of $g_+(\omega)$.

$\mathcal{J}_{t_o}(\omega) = \tilde{g}_+(\omega)$ has sampling properties similar to the periodogram for intervals between events. There are a number of differences, however, that make estimation of $g_+(\omega)$ more difficult than they were for the spectrum of intervals.

(1) The spectral density $g_+(\omega)$ is not a periodic function and hence neither is its estimator $\tilde{g}_+(\omega)$. The problem arises then over what values of ω to estimate $g_+(\omega)$.

(2) The lack of periodicity arises because the covariance density, in the time domain, is a function in continuous time. Thus,

if we require to smooth the estimates by some weight function, we can no longer perform the integrations as simply as before.

The function $g_+(\omega)$ is actually a "power spectrum" analogous to $\sigma^2 F(\omega)$ in the case of intervals. It is convenient to estimate the normalized spectrum $\frac{g_+(\omega)}{\gamma}$, $(\gamma = \frac{1}{\lambda})$ rather than the power spectrum. This can be done quite simply by normalizing the time scale by multiplying by the estimate $\frac{n}{t_o}$ of γ. We then obtain

$$\tilde{g}_+(\omega') = \frac{1}{\pi n} \sum_{s=1}^{n} \sum_{k=1}^{n} \exp\left[i \left\{ \frac{\omega t_o}{n} (T_s - T_k) \frac{n}{t_o} \right\} \right]$$

$$= \frac{1}{\pi n} \sum_{s=1}^{n} \sum_{k=1}^{n} \exp\left[i \left\{ \omega'(T_s' - T_k') \right\} \right].$$

The spectrum is computed at the points $\omega_p' = \frac{2\pi p}{n}$.

REFERENCES

1. Bartlett, M. S. (1950) Periodogram Analysis and Continuous Spectra, Biometrika 37, 1-16.
2. Bartlett, M. S. (1963) The Spectral Analysis of Point Processes, JRSSB, 264-295.
3. Bartlett, M. S. (1955) An Introduction to Stochastic Processes with Special Reference to Methods and Applications, Cambridge.
4. Bartlett, M. S. (1963) Statistical Estimation of Density Functions, Sankhya 25, 245-254.
5. Bartlett, M. S. (1937) Properties of Sufficiency and Statistical Tests, Proc. Roy. Soc. A 160, 268-282.
6. Birnbaum, Z. W. (1952) Numerical Tabulation of Kolmogorov's Statistic for Finite Sample Size, JASS 47, 425-441.
7. Blackman, R. B. and Tukey, J. W. (1958) The Measurement of Power Spectra, Dover.
8. Cox, D. R. (1954) On the Superposition of Several Strictly Periodic Sequences of Events, Biometrika 40, 354-360.
9. Cox, D. R. (1955) Some Statistical Methods Connected with Series of Events, JRSSB 17, 129-164.
10. Cox, D. R. (1965) On the Estimation of the Intensity Function of a Stationary Point Process, JRSSB 27, 332-337.
11. Cox, D. R. (1953) Some Simple Approximate Tests for Poisson Variates, Biometrika 40, 354-360.

12. Cox, D. R. and Smith, W. L. (1953) The Superposition of Several Strictly Periodic Sequences of Events, Biometrika 40, 1-11.

13. Cox, D. R. and Smith, W. L. (1954) On the Superposition of Renewal Processes, Biometrika 41, 91-99.

14. Daniel, S. H. E. (1962) The Estimation of Spectral Densities, JRSSB 185-198.

15. Hartley, H. O. (1949) Tests of Significance in Harmonic Analysis, Biometrika 36, 194-201.

16. Hoeffding, W. (1950) A Class of Statistics with Asymptotically Normal Distributions, AMS 19, 293-325.

17. Jenkins, G. M. (1961) General Considerations in the Estimation of Spectra, Technometrics 3, 133-166.

18. Jenkins, G. M. and Priestley, M. B. (1957) The Spectral Analysis of Time Series, JRSSB 19, 1-12.

19. Jones, R. H. (1965) A Reappraisal of the Periodogram in Spectral Analysis, Technometrics 7, 531-542.

20. Kampe de Feriat, J. and Frenkiel, F. N. (1962) Correlations and Spectra for Non-Stationary Random Function, Math. Comp., 1-19.

21. Lewis, T. (1961) The intervals between regular events displaced in time by independent random deviations of large dispersion, JRSSB 476-483.

22. Lomnicki, Z. A. and Zaremba, Z. K. (1957) On Estimating the Spectral Density Function of Stochastic Process, JRSSB 19, 13-37.

23. McFadden, J. A. (1962) On the lengths of intervals in stationary Point Processes, JRSSB 23, 364-382.

24. McFadden, J. A. and Weissblum, W., Higher Order Properties of Stationary Point Processes, JRSSB 25, 413-431.

25. Mood, A. M. (1963) Introduction to the Theory of Statistics, McGraw-Hill, New York.

26. Parzen, E. (1957) On choosing an estimate of the spectral density function of a stationary time series, AMS 28, 329-348.

27. Parzen, E. (1958) On Asymptotically Efficient, Consistent Estimates of the Spectral Density Function of a Stationary Time Series, JRSSB 20, 303-310.

28. Parzen, E. (1961) Mathematical Considerations in the Estimation of Spectra, Technometrics 3, 167-190.

29. Parzen, E. (1960) Modern Probability Theory and Its Applications, John Wiley, New York.

30. Priestley, M. B. (1962) The Analysis of Stationary Processes with Mixed Spectra I, JRSSB 215-233.

31. Priestley, M. B. (1962) Analysis of Stationary Processes with Mixed Spectra II, JRSSB 311-529.

32. Priestley, M. B. (1965) The Role of Bandwidth in Spectral Analysis Applied Stat. 14, 33-47.

33. Priestley, M. B. (1966) Design Relations for Non-Stationary Processes, JRSSB 228-240.

34. Priestley, M. B. (1962) Basic Considerations in Estimation of Spectra, Technometrics 4, 551-559.

35. Slutsky, E. (1938) Sur les fonctions aleatoires presques periodogram et sur la decomposition des fonctions aleatoires stationnaires en composantes, Actualites industre, 30-35.

36. Thompson, C. M. and Merrington, M. (1948) Tables for Testing the Homogenity of a Set of Estimated Variances, Biometrika 33, 296-304.

37. Wold, H. (1948) On Stationary Point Processes and Markov Chains, Skand. Aktuar. 31, 229-240.