Point Process Models with Applications to Safety and Reliability

W.A. Thompson, Jr
University of Missouri, Columbia

New York London
CHAPMAN AND HALL

First published in 1988 by Chapman and Hall
29 West 35th Street, New York NY 10001
Published in the UK by Chapman and Hall Ltd
11 New Fetter Lane, London EC4P 4EE

Printed in the United States of America

ISBN 0 412 01481 5

Library of Congress Cataloging in Publication Data

Thompson, William A.
Point process models with applications to safety and reliability.

Bibliography: p.
Includes index.
1. Point processes. I. Title.
QA274.42.T47 1988 519.2'32 87-23825
ISBN 0-412-01481-5

British Library Cataloguing in Publication Data

Thompson, W. A. 1928–
Point process models with applications
to safety and reliability.
1. Point processes
I. Title
519.2

ISBN 0-412-01481-5

Preface

In teaching an elementary course in stochastic processes it was noticed that many seemingly deep results in point processes are readily accessible by the device of representing them in terms of random gap lengths between points. The possibility of representing point processes in terms of sequences of random variables rather than probability measures makes them mathematically simpler than general stochastic processes. Point processes can be studied using only the tools of elementary probability, that is the joint distributions of finitely many random variables. Given the wide applicability of point process models and the difficulty of access by the measure-theoretic route, it was determined that the simpler representation is of sufficient expository importance to deserve emphasis. The present book is the result: it is specialized and short and therefore is called a monograph. In its development the material has been taught to several classes with pleasing results. Students have apparently understood theorems which by other methods appear difficult and deep. A few of the results, particularly on reliability, safety assessment and clustering, are original applied research.

An alternative title for this monograph might be 'Point processes: What they are and what they are good for.' The specialization to point processes is of course of restriction, but one cannot learn all of probability at once and point processes are more general and of more common applicability than would at first appear; the specialization is perhaps justified by the ease with which an extensive theory can be treated at an elementary level.

The monograph is written for those who have already mastered the material of a first rigorous course in probability[a] and for some reason wish to expand their knowledge of point process models. They may wish to better understand or appreciate the theory or perhaps to acquire skill in constructing probabilistic models. An elementary treatment of an extensive and applicable theory should be of special interest to those concerned with applications. The material is particularly useful for reliability, safety analysis, life distributions and clustering. The inclusion of problems makes the monograph suitable as a

[a] See Notes on the literature at the end of Appendix 1.

text for a specialized course or topics seminar for beginning graduate or advanced undergraduate students. Most of the problems merely check whether the text material was understood; but a few problems, those marked by an asterisk, extend the text and are moderately difficult. Familiarity with calculus concepts such as limit, supremum, continuity, derivative, integral, and Taylor's theorem is assumed. A few elementary differential equations are solved. The reader will need to be motivated and at home with mathematical reasoning as the arguments are elementary but not always easy.

A more ambitious task would be to treat not only models but their statistical analysis. That would at least double the size of the monograph and the time required for reading and writing it. Most of what we have to say fits comfortably under the title point process models but occasionally it has seemed desirable to comment on statistical issues. To understand this material, which is noted in the text and is not essential for the main theme, a beginning background in statistics is needed.

The intent of this monograph is to provide models for the placement of points on a time axis according to some chance mechanism. Chapter 1 is a preview of some of the theory and applications to be discussed. Section 2.1 attempts to place point processes in their probabilistic context. Appendix 1 provides general probability background which a potential reader may not have encountered. Chapter 2 sets forth the main issues and problems which arise when points are probabilistically placed on a time axis. Chapter 3 discusses homogeneous Poisson processes which are fundamental for understanding all point processes. The topic of Chapter 4 is how to quantify safety; point processes are found to be useful in this task. An extensive exposition of general Poisson processes, not available elsewhere, appears in Chapter 6. Brief discussions of renewal and superimposed processes, which are nicely exposited elsewhere, are presented for completeness in Chapters 5 and 7 respectively. Chapter 8 treats the important class of Markov point processes. Chapter 9 is a discussion, by example, of the concept of clustering of points; it is an application of Markov processes. Chapters 10 and 11 discuss topics having to do with length of life as in actuarial science. Appendix 2 treats point process concepts judged too technical to include in the text proper.

Some comments about notation and conventions will be helpful. We prefer to use capital Latin letters to denote random variables and the corresponding lower case to denote values which the random variables may assume. EX denotes the expectation of the random variable X. Where possible, parameters and parametric functions are denoted by Greek letters. The conclusion of a proof is signified by the symbol #. Some further conventions are gathered together in Tables 1 and 2; the reader may refer back to these rather than a hunting for the appropriate place in the text.

I have received much help in writing this monograph. The insistence of Harry Ascher convinced me that something was very wrong about the way

Table 1 Notational conventions

Concept	Notation
points on the time axis	t, l_i, t_i, u_i, v_i
universal indexing set, the set of all times considered	$T, [0, \tau]$
set consisting of the single point t	$\{t\}$
number of arrivals on $(0, t]$	$N(t)$
number of arrivals on the set A	NA
hazard rate of a lifelength	$r(t)$
intensity function of a process	$\lambda(t)$
arrival rate of a process	$\mu(t)$
$EN(t)$	$M(t)$
a function which, if divided by x, approaches zero as x goes to zero	$o(x)$
a function which, if divided by x, remains bounded as x goes to zero	$O(x)$
the Poisson density function	$p(k; \alpha)$

that certain reliability problems were being formulated. Bruce McDonald believed that I could help straighten things out. It was Ram Uppuluri who introduced me to the Reactor Safety Study and the field of risk analysis. Earlier versions of this material were used in a topics course at the University of Missouri; student comments and suggestions are much appreciated. In particular, I wish to thank Hamid Fallahi and Mohamed Habibullah for their help. Several anonymous reviewers have managed to say nice things about my manuscript, while at the same time prompting me to think more carefully about certain points. D. V. Hinkley, A. E. Raftery, Mark Dozzi and Bruce Thompson have been particularly helpful and generous with their time. Tina Carmack and Judy Dooley have stuck with me through version after version of typed manuscript.

Table 2 Random variables

Name	Notation	Distribution function
arrival time	Y_i	$G_i(y)$
gap length or gap	X_i	$H_i(x) = P(X_i \leqslant x)$
forward waiting time	W_t	$K_t(w) = P[N(t, t+w] \geqslant 1]$
conditional forward waiting time	$W_t\|$	$L_t(w) = P[N(t, t+w] \geqslant 1 \| N\{t\} \geqslant 1]$
lifelength	Z_i	$F(z)$

Contents

1

Introduction

1.1 ARRIVALS IN TIME

The motivation for this monograph is to provide models for the placement of points (often called arrivals) on the time axis according to some chance mechanism. Familiar examples are the arrivals of radioactive particles at a Geiger counter, the successive collisions of a given particle with other particles in the kinetic theory of an ideal gas, the arrivals of ships to be unloaded at a dock, and the arrivals of telephone calls at a switchboard. The independent variable need not be time but may be some other nonnegative quantity such as age, length or area. An example of the latter type would be the finding of red corpuscles on the surface of a microscope slide. But, we adopt a language and notation reflecting the common situation which is that the independent variable is time.

Chapter 2 sets forth the main issues and problems which arise when points are placed according to a probabilistic mechanism on the time axis.

Chapters 3 and 6 discuss Poisson processes. A Poisson process places points on the time axis so that the number of points on any interval is Poisson distributed and events, concerning the points, defined on nonoverlapping intervals are probabilistically independent. A point process is homogeneous if the probabilistic mechanism of placing points on a portion of the time axis is the same as the placement mechanism for any translation of that portion. Homogeneous processes often serve as models of physical processes which have been in operation sufficiently long to be in 'equilibrium.'

Chapter 3 treats processes which are both homogeneous and Poisson; these processes are central to the subject of point processes for two reasons. First, the homogeneous Poisson processes serve as models for the placement of points on the time axis 'at random.' 'At random' has several meanings. Choosing a sample of three from ten people at random means that every subset of size three has the same probability, and therefore $\binom{10}{3}^{-1}$, of constituting the sample. Choosing a point X at random in the interval $(0, 1)$ means according to the uniform distribution $P(X \leqslant x) = x,\ 0 < x < 1$. This meaning does not extend directly to the interval $(0, \infty)$. However we show, in section 3.4, that in several indirect senses this concept of 'at random' does extend to the entire positive real line and in fact yields the homogeneous Poisson processes.

A second reason for the central position of the homogeneous Poisson processes is that they may be characterized in terms of their properties in several ways, and the other main classes of point processes are obtained by omitting or altering one or another of these properties. Thus the homogeneous Poisson processes are special cases of the other classes. For instance, if the homogeneity property is dropped, the general class of Poisson processes is obtained. Chapter 6 contains the discussion of general Poisson processes.

Chapter 5 treats renewal processes for which, in a sense, each time a point occurs it is as though we were starting over from the beginning. Chapter 7, entitled Superimposed processes, treats properties of a summary process obtained from several point processes operating simultaneously.

Chapter 8 treats Markov point processes for which the probabilistic mechanism which produces future points depends on the number of prior points but is otherwise unaffected by the past. Chapter 9 is a discussion, by example, of the concept of clustering; it is an application of the previously discussed Markov processes. Clustering has to do with whether points in space or time are more or less regularly spaced than points distributed according to the homogeneous Poisson processes.

Chapters 10 and 11 treat topics having to do with length of life as in actuarial science. Chapter 10, The order statistics process, is an example of a Markov process as discussed in Chapter 8.

We strive for an account which is mathematically general but, where the generalization is clear, we like to use the suggestive and intuitive language of one application or another. Favorite applications will be to reliability and to the related area of safety assessment. Applications to reliability appear throughout the monograph but our discussion of safety assessment is concentrated primarily in Chapter 4 but then again in sections 3 and 4 of Chapter 11.

1.2 RELIABILITY

Loosely, reliability is the study of whether and when things will work. The statistical theory of reliability owes its initial success to its origin in attempting to make rockets work sufficiently well to carry out projects of the space program. A common reliability approach is to synthesize a complex engineering system into more elementary systems (called components) and to attempt to take advantage of the structure of the system in order to build mathematical models to predict whether a system will perform its required function or perhaps explain why it did or did not perform its function.

Much reliability theory (e.g. Mann, Schafer and Singpurwalla, 1974; Barlow and Proschan, 1975) deals strictly with nonrepairable systems and is essentially the study of lifetime distributions particularly in terms of the hazard rate or force of mortality.

Enlarging on this, let Z denote lifelength, that is, the length of time until a particular functioning object fails to function properly. Once the object fails, it stays in that state; we are considering it to be nonrepairable. Except for the intuitive background, we may think of lifelength as meaning simply a nonnegative random variable. Practical examples are the lifelength of a lightbulb, or of a person, or the storage life of a shelved drug. For continuous lifelength we write

$$P(Z \leqslant t) = F(t) = \int_0^t f(x)\,\mathrm{d}x$$

The exponential distribution

$$F(t) = \begin{cases} 0, & t < 0 \\ 1 - \mathrm{e}^{-\lambda t}, & t \geqslant 0 \end{cases}$$

has proved useful as a model for lifelength (see Epstein and Sobel (1954)), but it has a 'no aging property' which is peculiar in this context. If the random variable Z is exponential, then

$$P[Z > t + h | Z > t] = P[Z > h]$$

for all $t \geqslant 0$, $h > 0$. That is, in a probability sense, residual life is independent of age.

Obviously, many objects age, i.e., become more prone to failure, as they become older. Some actually strengthen as they get older, e.g. some electronic circuits and mechanical devices during early life.

The concept of hazard rate is useful to describe variation over time of the tendency of an object to fail. The **hazard rate**, force of mortality or failure rate of the lifelength Z, or the distribution F, is defined to be

$$\begin{aligned} r(t) &= \lim_{\Delta \to 0} \frac{P(t \leqslant Z < t + \Delta | t \leqslant Z)}{\Delta} \\ &= \lim_{\Delta \to 0} P(t \leqslant Z < t + \Delta)/[\Delta P(t \leqslant Z)] \\ &= \lim_{\Delta \to 0} \int_t^{t+\Delta} f(x)\,\mathrm{d}x/[\Delta \bar{F}(t)] = f(t)/\bar{F}(t) \end{aligned}$$

where $\bar{F}(t) = P(Z \geqslant t)$. Hazard rate is useful and has a meaningful interpretation: $r(t)\Delta$ represents approximately the probability that an object of age t will fail in the interval $[t, t + \Delta)$. If $r(t)$ does not decrease, then F or Z is said to be **increasing hazard rate**. A decreasing hazard rate is similarly defined.

A common situation is that of 'bathtub-shaped' hazard rate (Fig. 1.1); initially the hazard rate decreases from a relatively high value due to

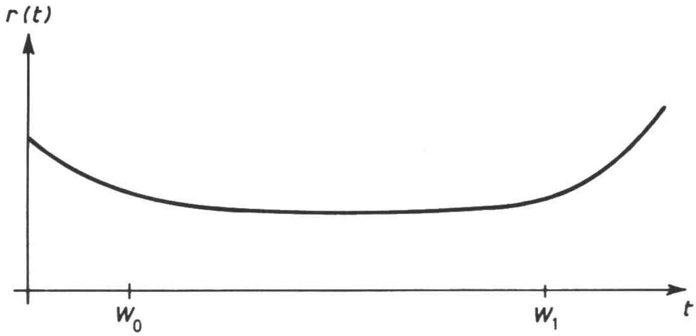

Fig. 1.1 Bathtub-shaped hazard rate.

manufacturing defects or infant mortality to a relatively stable middle life value and then slowly increases with the onset of old age or wearout. This behavior can be observed in any human actuarial life table.

A recurring theme of this monograph is the uses and abuses of the hazard rate concept. Many of the abuses are due to the unfortunate name 'failure rate.' The difficulty is that copies of a piece of equipment being simultaneously tested do not fail at rate $r(t)$. Stated in more detail, if n independent lifelengths with common distribution F and density f are sampled and if $Y_1, \ldots, Y_n$ denote the ordered failure times (the order statistics of section A1.4) then they constitute a process of failures evolving in time, the so-called order statistics process to be discussed in more detail in Chapter 10. There it is shown that the rate of expected failure of the order statistics process is **not** $r(t)$ nor $nr(t)$ but $nf(t)$, a fact which has brought many a practicing reliability analyst to grief. The name 'hazard rate' is not as easily misinterpreted. It is also shown in Chapter 10 that the conditional intensity of the Markov process of deaths in an actuarial cohort is $r(t)$ times the random number of individuals at risk at time t. This is perhaps the most meaningful sense in which $r(t)$ may be called a failure or death rate.

We call $\bar{F}(t) = P(Z \geqslant t)$ the **survival function**. For hardware, $\bar{F}(t)$ is often called the **reliability** since it is the probability that the equipment will perform its function adequately for a mission time t. For continuous distributions $\bar{F}(t) = 1 - F(t)$.

Note that

$$r(t) = \frac{\mathrm{d}}{\mathrm{d}t}[-\ln \bar{F}(t)] \tag{1.1}$$

Solving this differential equation we find that the survival function is

$$\bar{F}(t) = \exp\left[-\int_0^t r(x)\,\mathrm{d}x\right] \tag{1.2}$$

Differentiating, the density is

$$f(t) = r(t)\exp\left[-\int_0^t r(x)\,dx\right] \tag{1.3}$$

Some distributions which have been important in life studies are: (i) the exponential with constant hazard rate; (ii) the Weibull, with $r(t) = \rho\alpha t^{\alpha-1}$, $\rho, \alpha > 0$; and (iii) the Gompertz with $r(t) = B\exp(Ct)$, B, $C > 0$. Makeham's formula, $r(t) = A + B\exp(Ct)$, A, B, $C > 0$, has been important in the theory of life insurance (see Jordan (1975)). But reliability is not just actuarial science recast in an engineering setting; in some cases force of mortality may not even be to the point.

For our purpose it is convenient to draw a distinction between repairable and nonrepairable system reliability. We have been discussing nonrepairable reliability. Familiar repairable systems are automobiles and computers. The distinction between repairable and nonrepairable is not as clear as it first appears. The distinction often amounts to a **decision** whether to treat a product as throwaway or repairable goods. Many products, from bottles to automobiles, may be reprocessed or repaired and put back into service; although it may be uneconomical, unsafe or generally unwise to do so.

Most of us have lived in a society of throwaway consumer goods where items are simply manufactured, used, and discarded when they fail. But for several reasons the throwaway strategy will no longer suffice. First, things (such as space vehicles) are becoming so complex that the throwaway strategy might never produce a single working item. Second, as in the example of a nuclear power plant or jet airliner, the first failure may be too much. Finally, it is being realized that design for repair may be an economical policy after all.

When we move from throwaway to repairable industrial products, the statistical reliability considerations become more complex. For the throwaway product, reliability can be defined as the probability that the product will adequately perform its desired function for a prescribed length of time. In this way we are primarily interested in the probability distribution of time to first (and only) failure. But for repairable items we become interested in the rate at which failures are expected to occur. Actually, the entire probabilistic process by which failures occur will be of interest. Often the successive failures of a repairable system can be modeled as a point process.

1.3 SAFETY ASSESSMENT

A second favorite application is to safety assessment. With increased population density, environmental awareness and technological intensity, the desire and the need to quantify safety is great. There is a special need to assess the impact of technological innovation on societal safety through alteration of the environment. There are several approaches to the quantification of safety, some of them quite old and now well established. A few suggestive references are Jordan (1975), Freudenthal (1975), Gumbel (1958) and Rowe (1977). The most extensive safety study to date is the Nuclear Regulatory Commission Reactor Safety Study (1975). For this reason much of our discussion will be in terms of the safety of nuclear power plants. But this is just the best developed and understood example of a general problem.

A theory of safety assessment would outline general methods and procedures to quantify the safety of an activity. The activity might be existing or contemplated; it could be industrial, governmental, or military. There will of course be many problems of detail (which only a subject matter specialist could solve) in fitting the theory to specific applications; some activities will certainly fall outside the contemplated theory, perhaps motivating competing theories. But many activities have common features and their safety could be assessed by variations of a common theory. Safety assessment has a substantial probabilistic and statistical component.

1.4 RANDOM STRESS AND STRENGTH[a]

Structural problems in engineering are often formulated in terms of stress and strength. Systems are designed so that component strength exceeds anticipated stress. Freudenthal (1975) has suggested that probability should play a greater role in this design task. As an illustration of the possibilities, consider the following safety problem.

If the propellant of a shoulder-fired rocket is still burning when the rocket leaves the tube, then the operator will be burned. Given the firing data of Table 1.1, is the rocket safe to use?

One formulation is as follows. Let B and T denote burning and tube time and $D = T - B$. The probability of an accident is $p = P(D < 0)$. We might structure questions concerning the safety of the rocket in terms of p. We might, for example, require $p \leqslant .005$. If B and T are bivariate normal then D will be normal with unknown mean μ and variance σ^2. Now

$$p = P[(D - \mu)/\sigma \leqslant -\mu/\sigma] = \Phi(-\mu/\sigma)$$

where Φ is the standard normal distribution function. The requirement $p \leqslant .005$ is equivalent to $\mu/\sigma \geqslant 2.58$.

With just 20 isolated pieces of data we will not be very confident about any particular distributional assumptions. Timing variations in similar engineering studies are often normally distributed and the present data do not deviate in any excessive way from that distribution. In order to proceed, we ask what if

[a] This section is a digression and can be skipped: it does not involve point processes and it employs elementary statistical ideas outside of the assumed background of this monograph, but it does illustrate the context of safety assessment.

Table 1.1 Rocket firing data

Firing	Burning time (Coded)	Tube time (Same code)
1	58.671	69.524
2	61.284	69.542
3	60.619	71.256
4	60.699	69.462
5	60.101	70.404
6	58.619	70.602
7	59.426	72.732
8	60.096	70.420
9	61.389	69.528
10	61.249	71.412

the data were normal? The present analysis depends strongly on that hypothesis.

From Table 1.1, the unbiased estimates of μ and σ^2 are $\bar{d} = \sum_{i=1}^{10} d_i/10 = 10.273$ and $s_D^2 = \sum_{i=1}^{10} (d_i - \bar{d})^2/9 = (1.616)^2$, yielding 6.36 as an estimate of μ/σ. This corresponds to 0.00000 as an estimate for p, the probability of an accident.

We might wish to test the hypothesis that the rocket is unsafe, $\mu/\sigma \leqslant 2.58$. $T = \sqrt{10}\bar{d}/s_D$ has the noncentral t distribution with 9 degrees of freedom and noncentrality parameter $\delta = \sqrt{10}\mu/\sigma$. The hypothesis, unsafe according to the above criterion, corresponds to $\delta \leqslant 8.16$; large values of T are unfavorable to the hypothesis. The observed value of T is 20.1. From tables of the noncentral t distribution (e.g. Resnikoff and Liebermann, 1957), the observed significance level is

$$\sup_{\delta} P(T > 20.1 \mid \delta < 8.16) = 0.4\%$$

If the rocket were unsafe, $p > .005$, and we were to repeat the experiment indicated in Table 1.1 a large number of times then, under our assumptions, we would obtain sample evidence of safety as strong or stronger than that actually obtained in no more than 0.4% of the repetitions.

NOTES ON THE LITERATURE

The Weibull distribution was promoted for reliability purposes by Weibull (1951) as the simplest suitable function of the form (1.4), of Problem 4. The material on the reliability application in this monograph and indeed much of the theory is taken from Thompson (1981a).

PROBLEMS

1. In describing the rate of decay of radioactive substances, 'half life' is frequently employed. A particle is as likely to decay before its **half life** as after. If the lifelength of a particle follows the exponential distribution, give a formula relating expected life to half life.
2. In section 1.2 it was given that the hazard rate of the Weibull distribution is $r(t) = \rho\alpha t^{\alpha-1}$, $\rho, \alpha > 0$. What then are the distribution and density functions of the Weibull? What is the half life?
3. In the definition of the Weibull distribution, why are the parameters ρ and α restricted to be nonnegative?
4. If a chain is only as strong as its weakest link and if the probability of failure of a link at load x is

$$1 - \exp[\phi(x)] \tag{1.4}$$

   then what is the probability of failure of a chain of n links at load x?
5. If, in Problem 4, X denotes the strength of a randomly chosen link, then how is $\phi(x)$ related to the hazard rate of X?

2

Point processes

2.1 THE PROBABILISTIC CONTEXT

Point processes are of course a special topic in the theory of probability and we assume familiarity with the material of a first rigorous course in probability.[a] However, we need to establish notation and terminology and to place our subject in context.

The theory of probability hypothesizes the primitive concept of a probabilistic experiment. Examples are: (i) noting whether a nonrepairable system does or does not perform its intended function on demand; (ii) observing the wanderings of a particle in solution; (iii) recording the number of failures of a repairable system as a function of time. These experiments are probabilistic, as opposed to deterministic, because when they are performed we cannot predict how they will come out. The theory further hypothesizes a collection of fundamental or indivisible possible outcomes for the experiment. The individual possible outcomes (denoted by s) are called sample points and the collection of all of them is called the sample space, $S = \{s\}$. For example (i), the sample space is easy to describe: s_1 consists of the system performing its intended function while s_2 consists of the system not performing its intended function and $S = \{s_1, s_2\}$. For example (ii), presumably any three-dimensional continuous path would be a possible outcome. Example (iii) is a point process, the object of this writing, and we will discuss its sample space in more detail later in this chapter.

Probability is a set function, $P(\cdot)$, defined for a class of subsets (called events) of the sample space. The class of subsets for which probability is defined is closed under countable set operations (union, intersection and complementation) and probability must satisfy the Kolmogorov axioms:

1. The probabilities of all events are nonnegative.
2. The probability of the sample space is one.
3. If $A_1, A_2, \ldots$ are pairwise disjoint events (finite or denumerable in number) then $P(A_1 \cup A_2 \cup \ldots) = P(A_1) + P(A_2) + \cdots$

[a] Some of the probability results which we use may exceed this background. Appendix 1 is provided to alleviate this difficulty. See Notes on the literature at the end of Appendix 1.