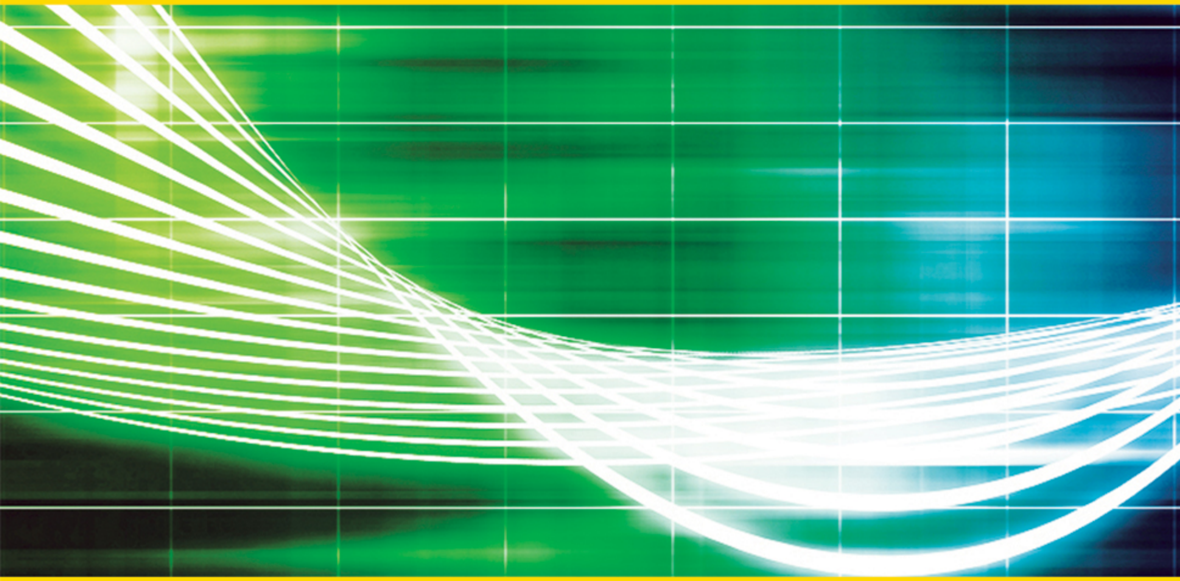
MATHEMATICS AND STATISTICS SERIES

MATHEMATICAL MODELS AND METHODS IN RELIABILITY SET



Volume 2

Recurrent Event Modeling Based on the Yule Process

Application to Water Network Asset Management

Yves Le Gat

ISTE

WILEY

Recurrent Event Modeling Based on the Yule Process

Mathematical Models and Methods in Reliability Set

coordinated by
Nikolaos Limnios and Bo Henry Lindqvist

Volume 2

Recurrent Event Modeling Based on the Yule Process

Application to Water Network Asset Management

Yves Le Gat

iSTE

WILEY

First published 2016 in Great Britain and the United States by ISTE Ltd and John Wiley & Sons, Inc.

ISTE Ltd
27-37 St George's Road
London SW19 4EU
UK

www.iste.co.uk

John Wiley & Sons, Inc.
111 River Street
Hoboken, NJ 07030
USA

www.wiley.com

Library of Congress Control Number: 2015955809

British Library Cataloguing-in-Publication Data
A CIP record for this book is available from the British Library
ISBN 978-1-84821-891-8

Contents

Preface

The research work presented in this book arises from the involvement of the author in engineering studies of the reliability of drinking water pipes. This type of infrastructure is organized as a network of pipelines, and failures, namely leakage or breakage, tend to occur in an aggregative manner on the same network segments. Building relevant strategies of infrastructure asset management requires, therefore, accurate modeling tools of the repeated failures that can affect some pipes, due to the heavy socioeconomic and environmental consequences of leakage and breakage.

Yves LE GAT
October 2015

1

Introduction

Examples of recurrent failures abound in the literature devoted to the reliability of technical objects, and in many cases, the occurrence rates tend to increase not only with the ageing of the object, but also with the number of past failures. The effect of ageing can be relevantly modeled using the now classical non-homogeneous Poisson process (NHPP), a comprehensive presentation of which can be found in [LAW 87], and a good example of application to drinking water pipe failures in [RØS 00]. In this same context of pipe failures, the PhD work of [EIS 94] emphasizes the critical importance of past failures. The consideration of the dependency of the failure process on its past is not a trivial question, and motivates a theoretical effort which the present book attempts to contribute to.

The basic concept of a *stochastic process* underlies all developments of the present work. A stochastic process must be understood as a function $X()$ of time t, each $X(t)$ being considered as a random variable (r.v.).

The stochastic process theory is the *natural* mathematical framework for studying the repetition of random events of the same kind. As presented by [COO 02], this question can be addressed from two alternative perspectives, which are equivalent and respectively consist of modeling:

– either the distribution of successive inter-arrival times;

– or the distribution of the number of events that occur in a given time interval.

Recurrent Event Modeling Based on the Yule Process: *Application to Water Network Asset Management,* First Edition. Yves Le Gat.

The method chosen by [EIS 94] arises from the first approach. The "classical" presentation of [ROS 83] arises from the second approach. The linear extension of the Yule process (called LEYP throughout the rest of the book) aims at building a failure occurrence model that cumulates the advantages of both NHPP and [EIS 94]'s approaches. This involves a theoretical setup, focused on the *counting process* concept, which is to be developed throughout the next two chapters.

A counting process is a particular stochastic process, simply designed to count repeated events, as presented in section 1.2.1.

As this presentation is to have a general scope, the entity subjected to repeated failures will be called a *technical object* or more simply an *object*; this term will be replaced by "water main" or "water pipe" when the context refers more specifically to failures that affect a water network.

1.1. Notation

The following mathematical notations will be used throughout this book:

– $\mathbb{N}$ and $\mathbb{N}^*$ respectively denote the sets of natural integers $\{0, 1, 2, \ldots, \infty\}$ and the set of strictly positive natural integers $\{1, 2, \ldots, \infty\}$;

– $\mathbb{R}$, $\mathbb{R}_+$ and $\mathbb{R}_+^*$ are the real sets $]-\infty, +\infty[$, $[0, +\infty[$ and $]0, +\infty[$;

– $\mathrm{P}(A)$ and $\mathrm{P}(A \mid B)$ respectively denote the probability of the event A, and the conditional probability of A given that the other event B occurs;

– $\mathrm{P}(A \cap B)$ and $\mathrm{P}(A, B)$ equivalently denote the joint probability of events A and B ; $\mathrm{P}\left(\bigcap_j A_j\right)$ more generally stands for the joint probability of events A_j;

– $t \in \mathbb{R}_+$ is a positive time variable that stands for the age of a technical object;

– $N(t) \in \mathbb{N}$ is an integer-valued step function that counts the failures;

– $\mathrm{d}N(t) = N(t + \mathrm{d}t) - N(t)$ is the differential of $N(t)$, i.e. $\mathrm{d}N(t) = 1$ whenever a failure occurs within $[t, t + \mathrm{d}t[$, $\mathrm{d}N(t) = 0$ otherwise;

– $\Delta N(t) = N(t) - N(t-)$ stands for the increment of $N(t)$ at t;

– $\mathscr{N}_{[a,t[}$ stands for the auto-exciting σ-algebra generated by the process $N(t)$ within $[a,t[$;

– $\mathscr{N}_{t-}$ stands for the auto-exciting σ-algebra $\mathscr{N}_{[0,t[}$;

– $\boldsymbol{Z}$ is a vector of failure factor values specific to a given technical object, also called "covariates";

– $\mathscr{F}_{[a,t[} = \mathscr{N}_{[a,t[} \vee \sigma(\boldsymbol{Z})$ denotes the information on the process $\mathscr{N}_{[a,t[}$ increased by the knowledge of the covariates $\boldsymbol{Z}$, or more technically the smallest σ-algebra that contains all events composed with events of σ-algebras $\mathscr{N}_{[a,t[}$ and $\sigma(\boldsymbol{Z})$;

– $\lambda(t)$ is a real positive function bounded on any compact interval, and its integral is $\Lambda(t) = \int_0^t \lambda(u)\mathrm{d}u$;

– $\mathrm{E}X$ and $\mathrm{E}(X \mid A)$ respectively denotes the expectation of the random variable (r. v.) X and its conditional expectation given A;

– $\mathrm{Var}(X)$ denotes the variance of the r. v. X;

– $\mathcal{U}_E$ stands for the uniform distribution on the set E;

– $\mathcal{U}_{[0,1]}$ denotes in particular the uniform distribution on interval $[0,1]$;

– $\mathcal{N}(\mu,\sigma^2)$ stands for the Gaussian distribution with expectation μ and variance σ^2;

– $\mathcal{P}o(\mu)$ is the Poisson distribution with expectation $\mu \in \mathbb{R}_+$;

– $\mathcal{NB}(\theta,p)$ is the negative binomial distribution with two parameters $\theta \in \mathbb{R}_+^*$ and $p \in [0,1]$;

– $\mathcal{NM}(\theta,(p_j)_{j=1,\dots,n})$ is the negative multinomial distribution with $n+1$ parameters $\theta \in \mathbb{R}_+^*$ and $p_j \in [0,1]$;

– $\mathcal{M}(k,(p_j)_{j=1,\dots,n})$ is the multinomial distribution with $n+1$ parameters $k \in \mathbb{N}^*$ and $p_j \in [0,1]$, where $\sum_{j=1}^n p_j = 1$;

– $\chi^2(k)$ is the Chi-squared distribution with $k \in \mathbb{N}^*$ degrees of freedom;

– $L(\boldsymbol{\theta})$ stands for the likelihood of a theoretical process with parameter $\boldsymbol{\theta}$ given a sequence of observed events;

– $\prod$ stands for the product integral operator, which plays the same role for products as the integral operator $\int$ plays for sums;