Insurance Risk Models with ON/OFF Process Inputs

Athens University of Economics and Business
Department of Statistics

George Makatis

June 3, 2025

CONTENTS

ABSTRACT

In this thesis we consider extensions of the classical Cramér-Lundberg model in risk theory to incorporate delayed and distributed claim settlement and portfolio modeling issues. The models discussed are also applicable to the analysis of traffic in computer communication systems. The first chapter reviews risk theory models together with basic tools from probability theory on which our subsequent analysis will be based. In Chapter 2 we consider a single ON/OFF source with independent, identically distributed activity cycles and obtain estimates for the ruin probability based both on approximations using gaussian processes and large deviation techniques. ON/OFF sources with correlated activity cycles are also briefly considered via a process based on a bivariate exponential distribution. The main result in this section is an expression for the Laplace transform of the correlation function $\rho(s)$. The significance of this computation lies in the fact that the large deviations exponent of a gaussian process can be expressed in terms of the variance of its integral process. This variance can be in turn computed from $\rho(s)$.

In Chapter 3 we examine the superposition of a large number of markovian ON/OFF processes with identical ON and OFF parameters and obtain a diffusion approximation for the superposition and a corresponding approximation for the free reserves process. We establish limit theorems for the total claim requirements over a period of time as the number of contracts grows large and indeed we show that, after scaling, the limiting process is an Ornstein-Uhlenbeck diffusion. We also consider portfolios consisting of a large number of contracts that do not have identical statistical characteristics and obtain conditions under which a limiting gaussian processing is obtained under appropriate scaling.

In Chapter 4 we describe such portfolios by introducing the concept of a *structure function* which describes statistically the characteristics of the contracts making up the portfolio in terms of frequency of claim generation and delay in claim settlement. Using this framework we obtain again limiting theorems for the superposition of the claim processes. Under appropriate conditions the limiting processes turn out to be a gaussian process (not necessarily markovian in this case) with covariance function $c(\tau)$ which has the form $\int_0^\infty e^{-|\tau|y}\nu(dy)$ where $\nu(dy)$ is a measure on $\mathbb{R}^+$ which depends on the structure function of the portfolio. This is also extended to the case where the ON times of the sources are not necessarily exponential. This is important since while the time to the next claim may reasonably assumed to be exponential, the time to claim settlement seldom is. Lindeberg-Feller type Central Limit Theorems are given for triangular arrays of such processes and the covariance function of the corresponding gaussian limiting processes are obtained in terms of the structure functions. In particular we examine the results obtained when the structure functions are of multivariate Gamma, multivariate Pareto, Dirichlet, and inverted Dirichlet.

In Chapter 5 we examine further the limits obtained by general triangular ON/OFF processes and we establish limit theorems with limit processes that are generalizations of the classic Ornstein–Uhlenbeck process to a process with a similar correlation structure but with marginals that have an infinitely divisible distribution. This can be thought of as defining an new class of processes which we term *Generalized Covevrage Processes with Infinitely Divisible Marginals.* The classic $M/M/\infty$ and $M/GI/\infty$ processes are particular instances of such processes. Finally Chapter 6 considers a simulation application of the original Cramér-Lundberg model. A "smoothed estimator" for the ruin probability and the "time in the red" is obtained and shown to have smaller variance than that of the corresponding crude Monte Carlo estimators. In practice, however, the variance reduction obtained was rarely greater than 15%. This has been published in *Mathematical Methods of Operations Research.*

1. INTRODUCTION

Collective risk theory, as a part of insurance mathematics studies the implications of the random fluctuations of the total assets, i.e. the risk reserve, of an insurance company. Its development dates from the beginning of the twentieth century with the seminal ideas of Filip Lundberg [100]) and their subsequent development by Harald Cramér with the introduction and use of the theory of stochastic processes which was also under development at that time (see Cramér [33]), and Täcklind [133]). Later, the introduction of renewal theoretic methods and Karamata's theory of regular variation in the study of risk processes allowed the analysis of claim distributions with heavy tails, as opposed to the classical Cramér–Lundberg theory which was based on the existence of exponential moments of the claim distribution.

Diffusion approximations are based on approximating the underlying risk process with an appropriate Brownian motion. This can be justified by the Functional Central Limit Theorem and can be used to obtain approximate such functionals of risk processes as the probability of ruin, severity of ruin etc. Such approximations may be used when the assumptions that underlie the classical risk model are not valid. They depend of course, much as any approach based on the Central Limit Theorem, on the assumption that claims have a finite second moment.

In the last three decades corresponding approximations for claims with infinite second moment have also been developed under appropriate conditions, based on Functional Central Limit Theorems for Lévy stable motion of exponent a, requiring scaling factors of the form $n^{\frac{1}{a}}$ with $1 < a < 2$ (see for instance [58]).

1.1 The Cramér Lundberg Model

Consider a company which only writes ordinary insurance policies such as accident, fire, damage etc. The policyholders pay premiums regularly and at certain random times make claims to the company. A policyholder's premium, is a positive amount composed of two components. The net risk premium is the component calculated to cover the payments of claims on the average, while the security risk premium, or *safety loading*, is the component which protects the company from large deviations of claims from the average and also allows an accumulation of capital. When a claim occurs the company pays the policyholder a positive amount called the positive risk sum.

The classical collective risk model referred to in the literature as the Cramér–Lundberg model, has the following structure:

The claim sizes $\{X_k\}_{k\in\mathbb{N}}$ are a sequence of independent and identically distributed random variables, having common distribution function F and mean value $\mu < \infty$. Otherwise no insurance company would insure such a risk Note that $F(x) = 0$ for $x < 0$. The point process $N = \{N(t); t \geq 0\}$ with $N(0) = 0$, which represents the number of claims in the time interval $[0, t]$, is assumed to be a (homogeneous) Poisson process with intensity λ. Furthermore, N and $\{X_k\}_{k\in\mathbb{N}}$ are assumed to be independent. Then,

$$S(t) = \sum_{i=1}^{N(t)} X_i$$

represents the accumulated claims up to time t and has distribution function

$$G_t(x) = P(S(t) \leq x) = \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} F^{n*}(x)\,, \quad x \geq 0.$$

where $F^{n*}(x) = P\left(\sum_{i=1}^{n} X_i \leq x\right)$ is the $n-$fold convolution of F with itself. The risk process U is defined by

$$U(t) = u + ct - S(t)$$

where u is the initial capital and $c > 0$ is the constant premium rate. Assume that N has intensity λ, i.e. $E[N(t)] = \lambda t$. Then, the profit over the interval $[0, t]$ is $Q(t) = ct - S(t)$ with

$$E[Q(t)] = ct - E[N(t)]E[X_1] = t(c - \lambda\mu).$$

For an insurance company it is important that $U(t)$ stays above a certain level. This level is given by legal restrictions. By adjusting the initial capital it is no loss of generality if we assume this level to be 0. We define the ruin time

$$\tau = \inf\{t > 0 : U(t) < 0\}.$$

We will mostly be interested in the probability of ruin in a time interval $(0, t]$

$$\Psi(u, t) = P(\tau \leq t \mid U(0) = u) = P\left(\inf_{0\leq s\leq t} U(s) < 0\right)$$

and the probability of ultimate ruin

$$\Psi(u) = \lim_{t\to\infty} \Psi(u, t) = P\left(\inf_{t\geq 0} U(t) < 0\right).$$

Is is easy to see that $\psi(u, t)$ is decreasing in u and increasing in t.

The *relative safety loading* ρ is defined by $\rho = \frac{c-\lambda\mu}{\lambda\mu} = \frac{c}{\lambda\mu} - 1$. If the risk process U has positive safety loading $\rho > 0$, then $U(t)$ almost surely drifts to $+\infty$. The condition $\rho > 0$ is also restated as the so-called *net profit condition* $c > \lambda\mu$ and says that on the average we obtain a higher premium income than a claim loss. If the net profit condition is fulfilled we can conclude that $\lim_{u\to\infty} \Psi(u) = 0$.

For the Cramér–Lundberg model under the net profit condition $\rho > 0$, one can show by using renewal arguments that

$$\Phi(u) := 1 - \Psi(u) = \frac{\rho}{1+\rho} \sum_{n=0}^{\infty} \left(\frac{1}{1+\rho}\right)^n F_I^{*n}(u) \tag{1.1}$$

where

$$F_I(x) = \frac{1}{\mu}\int_0^x \overline{F}(y)dy, \qquad x \geq 0$$

denotes the integrated tail distribution, $\overline{F}(x) = 1 - F(x)$, $x \geq 0$ denotes the tail of the distribution and $\Phi(u)$ is the non-ruin probability (for a proof see (Embrechts and Klüppelberg 1993). The above formula which is also known as the Khinchine-Pollazeck formula will be the key tool for estimating ruin probabilities under the assumption of heavy tails. The following Cramér–Lundberg estimates of the ruin probability $\Psi(u)$ are fundamental in risk theory. Only for special choices of the claim size distribution F it is possible to get an explicit expression.

1.1.1 The Cramér-Lundberg Theorem

Consider the Cramér–Lundberg model including the net profit condition $\rho > 0$. Assume that there exists a $R > 0$ such that

$$\int_0^\infty e^{Rx} dF_I(x) = \frac{c}{\mu\lambda} =: 1 + \rho.$$

$R > 0$ is called the *Lundberg exponent* or the *adjustment coefficient* of the underlying process and the following relations hold:

a) For all $u \geq 0$, $\Psi(u) \leq e^{-Ru}$.

b) If moreover $\int_0^\infty xe^{Rx}\overline{F}(x)dx < \infty$, then

$$\lim_{u\to\infty} e^{Ru}\Psi(u) = C \quad \text{where} \quad C = \left[\frac{R}{\rho\mu}\int_0^\infty xe^{Rx}\bar{F}(x)dx\right]^{-1}.$$

For a proof of the theorem see Grandell [73] or Mikosch [107]. A short proof of (a) based on martingale techniques is for instance discussed in [73]. This classical risk model has been generalized in many ways. For instance, the premiums may depend on the size of the risk reserves, inflation and interest may have to be included in the model. Also, the occurrence of the claims may be described by a more general point process such as a renewal or a Cox process. (See for example Bjork and Grandell [20] for an extension to the basic martingale approach to Cox models and Embrechts, Grandell and Schmidli [52] for a discussion on finite time ruin probabilities in the Cox case. See Rolski et al. [120] for related results and techniques.

The essential hypothesis in the Cramér–Lundberg approximation is the existence of exponential moments of the claim size distribution for some $R > 0$ which means that this distribution is light-tailed. If this assumption fails, which may often happen for some types of insurance and finanancial data it is not applicable. (This is for instance the case for claims that follow a Pareto or lognormal distribution.) Techniques for analyzing risk processes in the presence of heavy tails have been developed that relate the tail behavior of $\Psi(u)$ to that of F_I. It turns out that the proper class for this purpose is the class of so called S subexponential distributions. We refer the reader to [54].

1.1.2 Risk Processes Perturbed by Diffusions

An interesting generalization of the Cramér–Lundberg model involves perturbation by an independent Brownian motion. The free reserves process becomes then

$$U_W(t) = u + ct - S(t) + \eta W(t).$$

where u, c, and $S(t)$ are as in §1.1 and $\{W(t)\}$ is a standard Brownian motion, independent of S. The process $\{W(t)\}$ can be viewed as describing small perturbations around the risk process $U(t) = u + ct - S(t)$ modeling random fluctuations of the premium income, the return earned from investment, claim sizes, administration costs etc. Gerber [63] was the first to consider the classical risk process perturbed by Brownian motion by adding to the classical integral renewal equation a derivative term and transforming the resulting equation in order to solve it by means of the Key Renewal Theorem. Dufresne and Gerber [49] analyzed further this model. They established an explicit convolution formula for the infinite-time ruin probability. Veraverbeke [141] addresses the same problem by identifying the free reserves process as a Lévy process and obtaining the Laplace transform of the ruin probability in terms of its characteristic exponent. By inverting the Laplace transform both the light and heavy tail case for the claims in the compound Poisson process part of the model is analyzed. These results can be generalized to the renewal set-up as noted by Furrer and Schmidli [59]. Schmidli [125] in particular considered a renewal claim processes perturbed by diffusion and obtained the Cramér - Lundberg approximation by means of a measure change argument.

Gaussian processes have long since been studied and their utility in stochastic modeling is well accepted. However, they do not allow for large fluctuations and may sometimes be not adequate for modeling high variability. Furrer [57] added instead of a Brownian component an $\alpha-$stable Lévy motion to the classical risk process. considered the following model

$$U_Z(t) = u + ct - S(t) + \eta Z_\alpha(t)$$

where Z_α is an α stable Lévy motion with $1 < \alpha < 2$ and η is a positive number. The convolution formula for the probability of ruin is derived and contains the Mittag-Leffler function, a generalization of the exponential function to which it reduces in the Gaussian case ($a = 2$).

1.2 More General Models of Risk Processes

In the last few decades a number of researchers have considered more general models in order to capture important aspects of specific risk processes. These aspects include delay in claim settlement (see for instance Boogaert and Haezendonck [21] and Klüppelberg and Mikosch [90] and [92]) and IBNR (incurred but not reported) claims (see Neuhaus [112], Norberg [115] and Trufin, Albrecher and Denuit [139]). Such systems have been analyzed using diffusion approximations and large deviations techniques (see for instance Klüppelberg and Mikosch [89], [90], Brémaud [26], and Macci, Stabile and Torrisi [103] among others). In the model considered by Yuen, Guo and Ng [150] main claims also induce a delayed claim, called a by-claim. Zafiropoulos and Zazanis [151] have considered a process of claims with a series of by-claims and obtained a diffusion approximation based on the Central Limit Theorem for associated random variables.

This thesis considers for the most part risk processes arising from portfolios consisting of contracts each of which may at some point generate a claim resulting in a stream of payments. In [89], [90], [26], and many other papers a Poisson shot noise process has been used to model the resulting claim process: If $S(t)$ is the cumulative amount of claims that have occurred up to time t then

$$S(t) \;=\; \sum_{n=1}^{\infty} H_n(t - T_n), \quad t \geq 0.$$

In the above expression $\{T_n\}$ are the points of a Poisson process with rate λ and $\{H_n(\cdot)\}$ are a sequence of i.i.d. random functions with the property that $H(s) = 0$ for $s < 0$ and with sample paths that are cadlag (continuous from the right with limits from the left) and non-decreasing with probability 1. $H_n(\cdot)$ is the increasing process describing the cumulative amount of payments generated by the nth claim. $H_n(\infty) = \lim_{s\to\infty} H_n(s)$ is the total amount of the nth claim.

We will use alternating renewal processes i.e. ON/OFF processes as an idealized model of such situations in which, typically, the OFF period will be much longer than the ON period. We will consider a superposition of a large number of such processes, either with the same parameters, or, in the more interesting and realistic case with different parameters. These parameters may be supposed to come from an appropriate distribution describing the structure of the portfolio. (See Bühlmann [28, ch. 3].) Under appropriate assumptions we will show that suitably normalized sequences of these processes converge to standard stochastic processes such as Brownian motion, the Ornstein–Uhlenbeck process or, more generally Gaussian processes with given covariance structure whose behavior can then be analyzed.

1.3 Weak Convergence in Metric Spaces

In discussing diffusion approximations for ON/OFF processes and for the superposition of such processes, standard techniques of weak convergence of probability measures in metric spaces will be used. We mention a few basic concepts and results without proof and refer to Billingsley [18] for a detailed account.

Suppose that S is a metric space with metric ρ. A set $A \subset S$ is *relatively compact* if its closure $\overline{A}$ is compact. This is equivalent to the condition that each sequence in A contains a convergent subsequence (whose limit does not necessarily belong to A).

Denote by $\mathcal{S}$ the Borel σ–field generated by the open (or closed) sets of S. Every probability measure on $(S, \mathcal{S})$ is *regular*, i.e. for every $A \in \mathcal{S}$ and every $\epsilon > 0$ there exists a closed set F and an open set G such that $F \subset A \subset G$ and $P(G \setminus F) < \epsilon$. A probability measure on $(S, \mathcal{S})$ is *tight* if, for each $\epsilon > 0$, there exists a compact set K such that $P(K) > 1 - \epsilon$.

A sequence of probability measures $\{P_n\}$ on $(S, \mathcal{S})$ converges weakly to the probability measure P on the same space if, for every bounded, continuous, real function on S,

$$\lim_{n\to\infty} \int_S f(x) dP_n(x) = \int_S f(x) dP.$$

Equivalently, a sequence $\{X_n\}$ of random elements of S converges weakly (equivalently, in dis-

tribution) to the random element X if, for all bounded, continuous real functions f on S,

$$E\left[f(X_n)\right] \to E\left[f(X)\right], \qquad n \to \infty. \tag{1.2}$$

A family of measures Π on $(S, \mathcal{S})$ is called *relatively compact* if every sequence $\{P_n\}$ of elements of Π contains a subsequence $\{P_{n_k}\}$ which converges weakly.

Theorem 1.1. *If a family of measures Π on $(S, \mathcal{S})$ is tight then it is relatively compact. Conversely, if Π is relatively compact and in addition S is separable and complete, then it is tight.*

Theorem 1.2 (Continuous mapping theorem)**.** *Let h be a mapping from the metric space (S_1, ρ_1) to the metric space (S_2, ρ_2) (both equipped with the corresponding Borel σ–fields, $\mathcal{S}_i$, $i = 1, 2$, generated by the open subsets of each space). Suppose that $X_n \xrightarrow{d} X$ in S_1 and denote by P_A the distribution of A. Then $h\left(X_n\right) \xrightarrow{d} h(X)$ in the space S_2, provided that the set of discontinuities of h has P_A–measure zero.*

1.3.1 The Space $\mathbb{C}[0, 1]$

$\mathbb{C}[0, 1]$ is the space of continuous, real-valued functions on $[0, 1]$ equipped with the *supremum norm:* for $x \in \mathbb{C}[0, 1]$, $\|x\| = \sup_{0 \le t \le 1} |x(t)|$. This norm generates the *uniform metric* on $\mathbb{C}[0, 1]$. We also denote by $\mathcal{C}[0, 1]$ the Borel σ-field generated by the open sets in $\mathbb{C}[0, 1]$. (For the rest of the discussion we will abbreviate the notation to $(\mathbb{C}, \mathcal{C})$.) It is a standard result that this space is a *separable* and *complete*.

Let $\pi_{t_1,\ldots,t_k} : \mathbb{C}[0, 1] \to \mathbb{R}^k$ denote the projection map defined by $\pi_{t_1,\ldots,t_k} x = (x(t_1), \ldots, x(t_k))$. Obviously this is a contintinuous function and if $\mathcal{B}(\mathbb{R}^k)$ denotes the Borel σ–field of $\mathbb{R}^k$ then the σ–field $\sigma(\mathcal{C}_f)$ generated by the family $\mathcal{C}_f := \{\pi^{-1}_{t_1,\ldots,t_k}(\mathcal{B}(\mathbb{R}^k)) : k \in \mathbb{N},\, 0 < t_1 < \cdots < t_k < 1\}$. Closed balls in $\mathbb{C}$ can be expressed as countable intersections of sets in $\mathcal{C}_f$:

$$\overline{B}_\epsilon(x) := \left\{ y \in \mathbb{C} : \sup_{0 \le t \le 1} |y(t) - x(t)| \le \epsilon \right\} = \bigcap_{r \in \mathbb{Q} \cap [0,1]} \left\{ y \in \mathbb{C} : |y(r) - x(r)| \le \epsilon \right\}.$$

Hence $\sigma(\mathcal{C}_f)$ contains the closed balls. Therefore, $\mathcal{C} = \sigma(\mathcal{C}_f)$ which, in view of the fact that $\mathcal{C}_f$ is a π–system, means that $\mathcal{C}_f$ is a *separating class.*

If $\{P_n\}$ is relatively compact and the finite-dimensional distributions of P_n converge weakly to those of P then $P_n \xrightarrow{d} P$.

The *modulus of continuity* of a real function $x : [0, 1] \to \mathbb{R}$ is defined as

$$w_x(\delta) := \sup_{|s-t|<\delta} |x(t) - x(s)|, \quad \text{for } \delta \in (0, 1].$$

The condition $\lim_{\delta \to 0} w_x(\delta) = 0$ is equivalent to the uniform continuity of x on $[0, 1]$. (We will also use the alternative notation $w(x, \delta)$ for the modulus of continuity of x.)

Relative compactness in $\mathbb{C}$ is characterized by the following

Theorem 1.3 (Arzelà–Ascoli)**.** *The set $A \subset \mathbb{C}$ is relatively compatct if*

a) $\sup_{x\in A}|x(0)|<\infty$ *and*

b) $\lim_{\delta\to\infty}\sup_{x\in A}w_x(\delta)=0$.

Theorem 1.4. *The sequence* $\{P_n\}$ *on* $(\mathbb{C},\mathcal{C})$ *is tight iff for each* $\eta>0$ *there exists* $a>0$ *and* n_0 *such that*

a) $P_n(x:|x(0)|\geq a)\leq\eta$ *for* $n\geq n_0$ *and*

b) $P_n(x:w_x(\delta)\geq\epsilon)\leq\eta$ *for* $n\geq n_0$.

Theorem 1.5. *Suppose that for all* $0\leq t_1<t_2\ldots<t_k\leq 1$ $(X^n_{t_1},\ldots,X^n_{t_k})\xrightarrow{d}(X_{t_1},\ldots,X_{t_k})$. *If in addition for every* $\epsilon>0$ $\lim_{\delta\to 0}\limsup_{n\to\infty}P(w(X^n,\delta)\geq\epsilon)=0$ *then* $X^n\xrightarrow{d}X$.

If $\{P_n\}$ is relatively compact and the finite–dimensional distributions of P_n converge weakly to those of P, then $P_n\xrightarrow{d}P$ as $n\to\infty$.

1.3.2 The space $\mathbb{D}[0,1]$

The space $\mathbb{D}[0,1]$ consists of the *cadlag* functions on $[0,1]$, i.e. all functions that continuous from the right and have left limits. In particular, $\mathbb{C}[0,1]\subset\mathbb{D}[0,1]$. As before, here we will simplify the notation to $\mathbb{D}$. For each $x\in\mathbb{D}$ and each $\epsilon>0$ there exist points $0=t_0<t_1<\ldots<t_{v-1}<t_v\leq 1$ and $w_x[t_{i-1},t_i)<\epsilon$, $i=1,2,\ldots,v$. (Here $w_x(A):=\sup_{s,t\in A}|x(s)-x(t)|$ where $A\subset[0,1]$.) If $x\in\mathbb{D}$ then x has at most countably many discontinuities and $\sup_{t\in[0,1]}|x(t)|<\infty$.

A set of points $\{t_i\}_{i=0,1,\ldots,v}$ is δ*–sparse* if $\min_{1\leq i\leq v}(t_i-t_{i-1})>\delta$. The corresponding adaptation of the modulus of continuity to cadlag functions is

$$w'_x(\delta)=\inf_{\{t_i\}}\max_{1\leq i\leq v}w_x[t_{i-1}-t_i)\quad\text{where the infimum is over all }\delta\text{–sparse sets }\{t_i\}. \tag{1.3}$$

There are difficulties that arise when the metric $\rho(x,y)=\sup_{t\in[0,1]}|y(t)-x(t)|$ based on the sup norm is used in $\mathbb{D}$. On the mathematical side, one can see that $\mathbb{D}$ is not *separable* under this metric: The family of functions $\{x_t=\mathbf{1}_{[t,1]}\}$ is uncountable and any two members of the family have distance $\rho(\mathbf{1}_s,\mathbf{1}_t)=1$. Therefore there cannot exist a countable dense set in $\mathbb{D}$. If $\{t_n\}$ is an increasing sequence in $[0,1]$ such that $t_n\uparrow\tau$ then the sequence of functions $x_n(s):=\mathbf{1}_{[t_n,1]}(s)$ does not converge to $x(s):=\mathbf{1}_{[\tau,1]}(s)$ as $n\to\infty$ with respect to the metric ρ defined above.

Skorohod [129] introduced a metric that turns $\mathbb{D}$ into a separable space as follows. Let Λ be the class of continuous, strictly increasing functions $\lambda:[0,1]\to[0,1]$ such that $\lambda(0)=0$ and $\lambda(1)=1$. These conditions imply that λ is a bicontinuous bijection from $[0,1]$ onto $[0,1]$. One may then define a distance between two elements, x, y, of $\mathbb{D}[0,1]$ via

$$d(x,y)=\inf_{\lambda\in\Lambda}\left\{\|\lambda-I\|\vee\|x-y\circ\lambda\|\right\}. \tag{1.4}$$

The norm $\|\lambda-I\|$ measures the maximum deviation of the "clock rate" λ from the unit rate I. The norm $\|x-y\circ\lambda\|=\sup_{t\in[0,1]}|x(t)-y(\lambda(t))|$ gives the maximum difference between x and the time-changed $y\circ\lambda$.

The topology generated by the metric (1.4) is called Skorohod's J_1 topology.

Theorem 1.6 (Compactness in $\mathbb{D}$)**.** *A necessary and sufficient condition for a set $A \subset \mathbb{D}$ to be relatively compact in the Skorohod topology is that*

a) $\sup_{x\in A} \|x\| < \infty$ *and*

b) $\lim_{\delta\to 0} \sup_{x\in A} w'_x(\delta) = 0$.

Theorem 1.7 (Tightness in $\mathbb{D}$)**.** *The sequence $\{P_n\}$ is tight iff the following two conditions hold:*

i) $\lim_{a\to\infty} \limsup_{n\to\infty} P_n(x : \|x\| \geq a) = 0$.

ii) For each $\epsilon > 0$ $\lim_{\delta\to 0} \limsup_{n\to\infty} P_n(x : w'_x(\delta) \geq \epsilon)$.

Condition (i) is equivalent to the condition $\lim_{a\to\infty} \limsup_{n\to\infty} P_n(x : |x(t)| \geq a) = 0$ for all $t \in T$ where T is a dense countable set that contains 1.

The following theorem in Billingsley [18] guarantees the weak convergence of a family of processes in $\mathbb{D}[0, T]$.

Theorem 1.8. *Let $\{X_n(t); t \in [0, T]\}$ be sequence of real-valued processes and a process $\{X(t); t \in [0, T\}$ such that, for all finite dimensional distributions*

$$(X_n(t_1), \ldots, X_n(t_k)) \xrightarrow{d} (X(t_1), \ldots, X(t_k)) \tag{1.5}$$

for all $k \in \mathbb{N}$ and $t_i \in [0, T]$ and further that

$$X_T - X_{T-\delta} \xrightarrow{d} 0 \quad \text{as } \delta \to 0. \tag{1.6}$$

Finally, suppose that for $t_1 \leq t_2 \leq t_3$ the following inequality holds

$$E\left[|X_n(t_2) - X_n(t_1)|^\beta \, |X_n(t_3) - X_n(t_2)|^\beta\right] \leq K(t_3 - t_1)^{1+\alpha} \tag{1.7}$$

for some $\alpha > 0$, $\beta > 1$ and $K > 0$. Then $X_n \xrightarrow{d} X$ in $D[0, T]$.

1.3.3 The Functional Central Limit Theorem

Let $\{X_n\}$ be an i.i.d. sequence with mean μ and variance $0 < \sigma^2 < \infty$. We embed the sequence of partial sums $S_n = \sum_{i=1}^n X_i$ in a process on $[0, 1]$ and consider the limit process which turns out to be Brownian motion. First consider the process $\{S_n(t); t \in [0, 1]\}$ defined by

$$S_n(t) \;=\; \frac{1}{\sigma\sqrt{n}} \left(S_{[nt]} - \mu[nt] \;+\; (nt - [nt]) \left(X_{[nt]+1} - \mu\right)\right), \quad t \in [0, 1] \tag{1.8}$$

where $[x]$ denotes the integer part of the real number x. The paths $S_n(t)$ defined by (1.8) consist of a linear interpolation between the points $(\frac{k}{n}, S_n(\frac{k}{n}))$ and thus are continuous functions. Thus both $\{S_n(t)\}$ and $\{W_t\}$ assume values in $\mathbb{C}[0, 1]$.

We also introduce another process on $[0,1]$ which coincides with $S_n(t)$ at the points $\frac{k}{n}$, $k = 0, 1, \ldots, n$ and whose sample paths are step functions.

$$\widetilde{S}_n(t) \;=\; \frac{1}{\sigma\sqrt{n}}\left(S_{[nt]} - \mu[nt]\right), \quad t \in [0,1]. \tag{1.9}$$

The sample paths of this process are piecewise constant with (possible) jumps at the points $\frac{k}{n}$. Since they are càdlàg, they belong to the space $\mathbb{D}[0,1]$.

The following result is known as the *Donsker invariance principle* or the *Functional Central Limit Theorem (FCLT).*

Theorem 1.9 (Donsker invariance principle)**.** *Suppose that* $E\left[X^2\right] < \infty$*. Then*

(a) $S_n(\cdot) \xrightarrow{d} W_\cdot$ *in* $\mathbb{C}[0,1]$ *(in the sup norm topology).*

(b) $\widetilde{S}_n(\cdot) \xrightarrow{d} W_\cdot$ *in* $\mathbb{D}[0,1]$ *(in the* J_1 *Skorokhod topology).*

In particular, if f_1 *(respectively* f_2*) is continuous except possibly on a subset* $A \subset \mathbb{C}[0,1]$ *(respectively* $A \subset \mathbb{D}[0,1]$*) for which* $P(W_\cdot \in A) = 0$*, then* $f_1(S_n(\cdot)) \xrightarrow{d} f_1(B_\cdot)$ *and* $f_2(\widetilde{S}_n(\cdot)) \xrightarrow{d} f_2(W_\cdot)$*.*

A proof of the above theorem can be found in Billingsley [18]. See also Embrechts, Klüppelberg and Mikosch [54]. The Donsker invariance principle is a powerful result. It explains why Brownian motion can be taken as a reasonable approximation to many real processes which are in some way related to sums of independent random variables. It also suggests an easy way of simulating Brownian sample paths by means of the approximating processes $S_n(\cdot)$ or $\widetilde{S}_n(\cdot)$. In this case it is best to take $\{X_n\}$ to be an i.i.d. sequence of standard normal random variables.

The importance of weak convergence in metric spaces is enhanced by the so called *continuous mapping theorem.* It means that weak convergence of a sequence of random elements in a metric space is preserved under continuous mappings

We may conclude from the Donsker invariance principle that the finite-dimensional distributions of the processes $S_n(\cdot)$ and $\widetilde{S}_n(\cdot)$ converge. Indeed, the projection mapping $\pi_{t_1,\ldots,t_m} : \mathbb{D}[0,1] \to \mathbb{R}^m$ defined by

$$\pi_{t_1,\ldots,t_m}(x) = (x_{t_1}, x_{t_2}, \ldots, x_{t_m})$$

for any given $0 \le t_1 \le \cdots \le t_m \le 1$ is continuous at elements $x \in \mathbb{C}[0,1]$. Hence weak convergence of the processes $S_n(\cdot)$ and $\widetilde{S}_n(\cdot)$ implies convergence of the finite-dimensional distributions. Moreover, the following functionals are continuous on both spaces $\mathbb{C}[0,1]$ and $\mathbb{D}[0,1]$ when endowed with the sup-norm:

$$f_1(x) = x(1), \qquad f_2(x) = \sup_{0\le t\le 1} x(t), \qquad f_3(x) = \inf_{0\le t\le 1} x(t).$$

Moreover, the multivariate function (f_1, f_2, f_3) is continuous on both spaces $\mathbb{C}[0,1]$ and $\mathbb{D}[0,1]$. From Donsker's invariance principle and the continuous mapping theorem we immediately obtain that

$$\frac{1}{\sigma\sqrt{n}}\left(S_n - n\mu, \max_{0\le k\le n}\left(S_k - k\mu\right), \min_{0\le k\le n}\left(S_k - k\mu\right)\right) \xrightarrow{d} \left(W_1, \max_{0\le t\le 1} W_t, \min_{0\le t\le 1} W_t\right). \tag{1.10}$$

However, the joint distribution of W_1, the minimum and maximum of Brownian motion on $[0,1]$ is well known (see Billingsley [18]). Thus we can take advantage of the above limit theorem in order to approximate the joint distribution of the left hand side of (1.10) which is usually hard to evaluate in practice by means of the corresponding joint distribution for Brownian motion.

1.3.4 Limit theorems for cumulative processes

The following theorem presents a small part of the results dealing with limit theorems for cumulative processes arising from regenerative processes given in Glynn and Whitt [67] that will be used in the next chapter.

Suppose that $\{T_i\}$, $i = 0, 1, 2, \ldots$ is an ordinary renewal process with $T_0 = 0$ and $\tau_i := T_i - T_{i-1}$ (where of course the $\{\tau_i\}$ are i.i.d.) for which we will assume that $E\tau_1 < \infty$. Let $\{\xi_t; t \geq 0\}$ be a real valued (classical) regenerative process with respect to $\{T_i\}$ (see [7] for definitions). Let also f be a measurable real function and define the process $\{C(t); t \geq 0\}$ with $C(t) := \int_0^t f(\xi_s)ds$. Define also the sequence of random variables

$$\begin{aligned} Z_i &= \int_{T_{i-1}}^{T_i} f(\xi_s)ds, \\ V_i &= \sup_{0\leq s\leq \tau_i} \left| \int_0^s f(\xi_{T_{i-1}+u})du \right|, \quad i = 1, 2, \ldots \end{aligned}$$

The sequence of triplets $\{(\tau_i, Z_i, V_i)\}$, $i = 1, 2, \ldots,$. We further assume that $\int_0^t |f(\xi_s)|ds < \infty$ w.p. 1 for each t and that $EZ_1 < \infty$. Under the additional assumption that $EV_1 < \infty$ the following Strong Law of Large Numbers holds (see [7])

$$\lim_{t\to\infty} \frac{C(t)}{t} = \alpha := \frac{E[Z_1]}{E\tau_1} \text{ w.p. 1.} \tag{1.11}$$

Define also the quatities

$$\begin{aligned} \widetilde{Z}_1 &= Z_1 - \alpha\tau_1, \\ \widetilde{V}_1 &= \sup_{0\leq s\leq \tau_i} \left| \int_0^s \left[f(\xi_{T_{i-1}+u}) - \alpha \right] du \right|. \end{aligned}$$

Theorem 1.10 (Glynn and Whitt, 1993)**. a)** *In the above framework and under the above assumptions the following Functional SLLN also holds:*

$$\sup_{0\leq t\leq T} \left| n^{-1}C(nt) - t\alpha \right| \to 0 \quad \textit{as } n \to \infty \textit{ for all } t > 0.$$

b) *Under the additional assumptions,*

$$E\widetilde{Z}_1^2 < \infty, \tag{1.12}$$

$$\lim_{t\to\infty} t^2 P\left(\widetilde{V}_1 > t\right) = 0. \tag{1.13}$$

the following FCLT holds

$$\frac{1}{\sigma\sqrt{n}} \left[C(nt) - \alpha nt \right] \xrightarrow{d} W_t \quad \textit{in } D[0,T], \textit{ as } n \to \infty \tag{1.14}$$

where $\{W_t; t \geq 0\}$ *is Standard Brownian Motion and*

$$\sigma^2 \;=\; \frac{E(\widetilde{Z}_1)^2}{E\tau_1}. \tag{1.15}$$

1.4 Diffusion Approximation

One way of evaluating ruin probabilities is based on diffusion approximations. The objective is to obtain approximations for the distribution of $U(t)$ and $\tau = \inf\{t : U(t) < 0\}$ (i.e. time of ruin), and other functionals of $U(t)$ by applying the theory of weak convergence of probability measure on function spaces. A sequence of risk processes $\{U_n(t); t \geq 0\}$, $n = 1, 2, \ldots$ is defined and weak convergence to a Brownian motion process with drift is shown. For the limit process one is able to calculate explicitly the density of the first passage time corresponding to the random variable T. The distributions of the random variables derived from the limit process are then proposed as approximations for the process $\{X_n(t)\}$ for large n.The diffusion approximation has many disadvantages, but also some good aspects. We refer to Asmussen [8] for a discussion and some recent literature. Among the positive aspects of the diffusion approximation is that it is applicable to a wide range of risk processes which deviate from the Cramér-Lundberg model. In that case, the classical methods from renewal theory will usually break down, and the diffusion approach is then one of the few tools which work. In such more general models it is usually not possible to choose the premiums as a linear function in time; see for example Klüppelberg and Mikosch [90] for a shot noise risk model. Their main result is a FCLT with a self-similar Gaussian limit process which in the classical case, is Brownian motion. Their theorems are derived under regularity conditions on the moment and covariance functions of the shot noise process. The model is applied to delay in claim settlement in insurance portfolios. They use the asymptotic theory for studying the ruin time and ruin probability for a risk process which is based on the Poisson shot noise process. The ruin problem is finally reduced to first hitting problems of Brownian motion and of an Ornstein–Uhlenbeck process with constant boundaries.

Already in 1940 H. Hadwiger [75] was comparing a discrete risk process with a diffusion. Though theoretically not comparable with the modern approach, this can be viewed, as the first treatment of diffusion approximations in risk theory. A more modern version, based on weak convergence, is due to Iglehart [78] (see also Grandell [73, Appendix A.4] for more details and Schmidli [124]). The idea is to let the number of claims grow in a unit time interval and to make the claim sizes smaller in such a way that the risk process converges weakly to a diffusion. An interesting approximation of the risk process by an α-stable motion ($1 < \alpha < 2$) with drift when the claim distribution does not have a finite variance was given in Furrer et al.[58].

Ruin Problems

In the classical Cramer-Lundberg model the claim process $(S(t); t \geq 0)$ is compound Poisson with positive i.i.d. claims $(X_n; n = 1, 2, \ldots)$ with $EX_1 = \mu$ and $\mathsf{Var}(X_1) = \sigma^2$, independent of the homogeneous Poisson process $(N(t); t \geq 0)$ with intensity $\lambda > 0$. The corresponding risk process with initial capital u and premium rate $c = (1 + \rho)\lambda\mu > 0$ (where $\rho > 0$ is the safety

loading) is given as

$$U(t) = u + ct - S(t), \qquad t \geq 0\,.$$

Iglehart [78] discusses diffusion approximations for ruin probabilities in a mathematically rigorous fashion. Consider the ruin probability $\Psi(u,T)$ in finite time

$$\Psi(u,T) \;=\; P\left(U(t) < 0 \text{ for some } t \leq T\right) \;=\; P\left(\inf_{0\leq t\leq T}\left(ct - S(t)\right) < -u\right)$$

Then $\mathsf{Var}(S(t)) = t\widetilde{\sigma}^2$ with $\widetilde{\sigma}^2 = \lambda(\sigma^2 + \mu^2)$. Then,

$$\begin{aligned}\Psi(u,T) &= P\left(\inf_{0\leq t\leq T}\left((1+\rho)\,\lambda\mu t - S(t)\right) < -u\right) \;=\; P\left(\inf_{0\leq t\leq T/n}\left((1+\rho)\lambda\mu tn - S(tn)\right) < -u\right)\\ &= P\left(\inf_{0\leq t\leq T/n}\left(\widetilde{\sigma}^2 n\right)^{-\frac{1}{2}}\left((1+\rho)\lambda\mu tn - S(tn)\right) < -u\left(\widetilde{\sigma}^2 n\right)^{-\frac{1}{2}}\right).\end{aligned}$$

Now assume that $T_0 = \frac{1}{n}T$, $\rho_0 = \rho\lambda\mu\widetilde{\sigma}^{-1}\sqrt{n}$ and $u_0 = u\left(\widetilde{\sigma}^2 n\right)^{-\frac{1}{2}}$ are constants, i.e. we increase T and u with n, and decrease at the same time the safety loading ρ with n. This means that a small safety loading is compensated by a large initial capital. Then we obtain

$$\Psi(u,T) = P\left(\inf_{0\leq t\leq T_0}\left(\left(\widetilde{\sigma}^2 n\right)^{-\frac{1}{2}}\left(\lambda\mu tn - S(tn)\right) + \rho_0 t\right) < -u_0\right).$$

The functional $x(f) = \inf_{0\leq t\leq T_0} f(t)$ is continuous on $\mathbb{D}[0,T_0]$ and we may conclude from the Functional Central Limit Theorem and the Continuous Mapping Theorem (noting that u, ρ, and T depend on n) that

$$\Psi(u,T) \;\to\; P\left(\inf_{0\leq t\leq T_0}\left(\rho_0 t - W_t\right) < -u_0\right) \;=\; P\left(\sup_{0\leq t\leq T_0}\left(W_t - \rho_0 t\right) > u_0\right)$$

where $\{W_t; t \geq 0\}$ is standard Brownian motion. The distribution of the supremum functional of Brownian motion with linear drift is well known (see for instance Lerche [97]).

$$P\left(\sup_{0\leq t\leq T_0}\left(W_t - \rho_0 t\right) > u_0\right) = \overline{\Phi}\left(\frac{\rho_0 T_0 + u_0}{\sqrt{T_0}}\right) + e^{-2u_0\rho_0}\,\Phi\left(\frac{\rho_0 T_0 - u_0}{\sqrt{T_0}}\right).$$

(In the above Φ denotes the standard normal distribution function.) Such diffusion approximations suffer from low accuracy. On the other hand the technique has a wide range of applicability.

1.5 Large Deviations

Here we state briefly the basic results behind the logarithmic asymptotics obtained in Chapter 2 and 3. We refer the reader to Dembo and Zeitouni [41] and Ganesh, O' Connell, and Wischik [60]. Denote by $\overline{\mathbb{R}} := \mathbb{R}\cup\{-\infty,\infty\}$ the *extended real numbers*. If $\mathcal{X}\subset\mathbb{R}^d$ a function $f:\mathcal{X}\to\overline{\mathbb{R}}$ is *lower semicontinuous* if, for every $x_0\in\mathcal{X}$, $\liminf_{x\to x_0} f(x) \geq f(x_0)$. Equivalently, a function is lower semicontinuous if, for all $y\in\mathbb{R}$ the level sets $\{x\in\mathcal{X} : f(x)\leq y\}$ are closed in $\mathcal{X}$ for all $y\in\mathbb{R}$.

A function $I:\mathbb{R}^d\to\overline{\mathbb{R}}$ is called a rate function if

1. $I(x) \geq 0$ for all $\mathbb{R}^d$ and
2. I is a lower semicontinuous on $\mathbb{R}^d$ which means that all level sets of I are closed.
3. If, additionally, all level set of I are also *compact* then I is called a *good rate function.*

Definition 1.11. *A sequence of random elements of $\mathbb{R}^d$, $\{Y_n\}$, is said to satisfy a* Large Deviations Principle *with good rate function I if I is a good rate function and for any Borel set $B \subset \mathbb{R}^d$*

$$-\inf_{x\in B^o} I(x) \leq \liminf_{n\to\infty} \frac{1}{n} \log P(Y_n \in B) \leq \limsup_{n\to\infty} \frac{1}{n} \log P(Y_n \in B) \leq -\inf_{x\in \bar{B}} I(x) \tag{1.16}$$

where B^o and $\bar{B}$ are the interior and closure of B respectively.

Definition 1.12. *Let $f : \mathbb{R}^d \to \overline{\mathbb{R}}$. The convex conjugate or Fenchel–Legendre transform of f, denoted by $f^* : \mathbb{R}^d \to \overline{\mathbb{R}}$ is defined by $f^*(\theta) = \sup_{x\in\mathbb{R}^d} (\langle \theta, x\rangle - f(x))$.*

f^* is convex and lower semicontinuous. If f is a convex function, differentiable at $x \in \mathbb{R}^d$ with $\nabla f(x) = \eta$, then $f^*(\eta) = \langle \eta, x\rangle - f(x)$.

Let $\{X_n\}$ be an i.i.d. sequence of real random variables with cumulant generating function $\Lambda(\theta) = \log Ee^{\theta X_1}$. Define the sequence $\{S_n\}$ via $S_n = X_1 + \cdots + X_n$. Then $\frac{1}{n} \log Ee^{\theta S_n} = \Lambda(\theta)$. Cramér's theorem gives a large deviations principle for the mean sequence $\frac{1}{n} S_n$.

Theorem 1.13 (Cramér)**.** *If $\{X_n\}$ is an i.i.d. sequence of real random variables in the above framework and if Λ is finite in a neighborhood of 0 then the sequence of random variables $\{\frac{1}{n} S_n\}$ satisfies a Large Deviation Principle with good rate function Λ^*.*

The generalization of Cramér's theorem beyond the context of sums of i.i.d. random variables to more general sequences of random variables given in the Gärtner–Ellis Theorem greatly enhances the applicability of this result.

Definition 1.14. *The effective domain of a function $f : \mathcal{X} \to \overline{\mathbb{R}}$ is the set $\{x \in \mathcal{X} : f(x) < \infty\}$. The function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ is called* essentially smooth *if the interior of its effective domain is nonempty, f is differentiable there, and f is* steep *meaning that, for any sequence $\{\theta_n\}$ which converges to a boundary point at the effective domain of f, $\lim_{n\to\infty} |\nabla f(\theta_n)| = +\infty$.*

Let Λ be a convex, real-valued function such that $\Lambda(0) = 0$ and let Λ^* be its convex conjugate. Then Λ^* is non-negative, convex, and lower semi-continuity. If x is differentiable at the origin with $\mu = \Lambda'(0)$ then, in addition, $\Lambda^*(\mu)$ and Λ^* is decreasing on $(-\infty, \mu]$ and increasing on $[\mu, \infty)$.

Theorem 1.15. *(Gärtner–Ellis) Let $\{S_n\}$ be a sequence of random variables with values in $\mathbb{R}^d$ for which the limit $\Lambda(\theta) := \lim_{n\to\infty} \frac{1}{n} \log Ee^{\theta S_n}$ exitst for each $\theta \in \mathbb{R}^d$. Suppose that Λ is finite in a neighborhood of $\theta = 0$ and that Λ is essentially smooth and lower semi-continuous. Then the sequence $\{\frac{1}{n} S_n\}$ satisfies a Large Deviation Principle in $\mathbb{R}^d$ with good convex rate function Λ^*, the convex conjugate of Λ.*

The following theorem, in the spirit of the above results, is given in [68] and is used to obtain logarithmic asymptotics in a number of models examined in Chapter 2.

Theorem 1.16 (Glynn and Whitt, 1994)**.** *Let $\{X_n : n \geq 1\}$ be a sequence of real valued r.v.'s (not necessarily stationary or mutually independent). Let, for each $n \in \mathbb{N}$, $S_n := X_1 + X_2 + \cdots + X_n$ and $M_n := \max(X_1, X_2, \ldots, X_n)$. If there exists a function ψ and positive constants θ^* and ϵ^* such that the following hold*

GW1 $\frac{1}{n} \log E \exp(\theta S_n) \to \psi(\theta)$ *as* $n \to \infty$ *for* $|\theta - \theta^*| < \epsilon^*$, ψ *is finite in a neighborhood of* θ^* *and differentiable of* θ^* *and differentiable at* θ^* *with* $\psi(\theta^*) = 0$,

GW2 $\psi'(\theta^*) > 0$,

GW3 $E \exp(\theta^* S_n) < \infty$ *for* $n \geq 1$

GW4 $\limsup_{n\to\infty} E \exp(\theta X_n) < \infty$ *for* $|\theta| < \epsilon^*$

then $M_n \to M$ *w.p. 1 as* $n \to \infty$ *and*

$$x^{-1} \log P(M > x) \to -\theta^* \quad \text{as } x \to \infty. \tag{1.17}$$

A generalization was obtained by Duffield and O'Connell [48]. They showed that if there exist increasing scaling functions (a_t, υ_t) such that the pair $\left(\frac{W_t}{a_t}, \upsilon_t\right)$ satisfies a large deviation principle with rate function I. It is assumed that there exists a scaling function (h_t) such that the limit

$$g(c) \;=\; \lim_{t\to\infty} \frac{\upsilon\left(a^{-1}(c/t)\right)}{h_t}$$

exists for each $c > 0$.

Let $(Z_t, t \in \mathbb{R}_+)$ be a zero-mean Gaussian process with stationary increments and covariance function $\Gamma(s,t) = E[Z_s Z_t]$ and set $W_t := Z_t - \mu t$. This is quite a general model for the free reserves process (or the workload process in queueing applications), which includes fractional Brownian motion; (See (Aldous [3, Ch. 3]) and the references therein.) Setting $\sigma_t^2 := \Gamma(t,t)$ we have that

$$\lambda(\theta) := \lim_{t\to\infty} \frac{\sigma_t^2}{t^2} \log E\left[e^{\frac{\theta t W_t}{\sigma_t^2}}\right] = \frac{1}{2}\theta^2 - \theta\mu$$

The scaling functions that satisfy the Hypotheses in [48] are $a_t = t$ and and $\upsilon_t = \frac{t^2}{\sigma_t^2}$, provided the limit

$$g(c) := \lim_{t\to\infty} \frac{\sigma_t^2}{c^2 \sigma_{t/c}^2}$$

exists for each $c > 0$. In particular, if the variance σ_t^2 is asymptotically linear in t, i.e. $\frac{\sigma_t^2}{t} \to \sigma^2 > 0$ say, as $t \to \infty$, then it is shown in [48] that

$$\lim_{b\to\infty} \frac{1}{b} \log P(\sup_t W_t > b) = -\frac{2\mu}{\sigma^2}. \tag{1.18}$$

1.6 The Lindeberg-Feller Central Limit Theorem

We begin this section with the following two important theorems of Helly on convergene in distribution.

1.6.1 Helly's Theorems and Convergence in Distribution

Theorem 1.17 (Helly's first Theorem). *Any sequence of uniformly bounded, non-decreasing functions $F_1, F_2, \ldots$ on $\mathbb{R}$ contains a subsequence $F_{n_1}, F_{n_2}, \ldots$, that converges weakly a non–decreasing function F.*

Theorem 1.18 (Helly's second Theorem). *Let f be a continuous function on $\mathbb{R}$ and suppose that the sequence of non-decreasing uniformly bounded functions $F_1, F_2, \ldots$ converge weakly to F on some finite interval $[a, b]$, where a, b are continuity points of the F; then*

$$\lim_{n\to\infty} \int_a^b f(x)dF_n(x) = \int_a^b f(x)dF(x).$$

Theorem 1.19 (Generalized second Theorem of Helly). *If the function f is continuous and bounded over the entire line $-\infty < x < +\infty$, the sequence of non–decreasing uniformly bounded functions $F_1, F_2, \ldots$, converge weakly to F and $\lim_{n\to\infty} F_n(-\infty) = F(-\infty)$, $\lim_{n\to\infty} F_n(+\infty) = F(+\infty)$, it follows that*

$$\lim_{n\to\infty} \int f(x)dF_n(x) = \int f(x)dF(x).$$

Next, we state two limit theorems for characteristic functions namely the Direct Limit Theorem and the Converse Limit Theorem which are proven with the help of the Generalized Second Theorem and Helly's First Theorem respectively.

The following two theorems allow one to establish weak convergence of a sequence of distributions by establishing the convergence of the corresponding sequence of characteristic functions.

Theorem 1.20 (Direct Limit Theorem). *If a sequence of distribution functions $F_1, F_2, \ldots$ converges weakly to the distribution function F, then the sequence of characteristic functions $\phi_1, \phi_2, \ldots$ converges to the characteristic function f. This convergence is uniform in each finite interval.*

Theorem 1.21 (Convergence Limit Theorem). *If a sequence of characteristic functions $\phi_1, \phi_2, \ldots$, converges to the continuous function ϕ, then the sequence of distribution functions $F_1, F_2, \ldots$ converges weakly to some distribution F.*

1.6.2 The Lindeberg–Feller conditions and the Central Limit Theorem

Consider a double array of random variables

$$\begin{array}{l} \xi_{11},\, \xi_{12},\, \ldots, \xi_{1k_1} \\ \xi_{21},\, \xi_{22},\, \ldots, \xi_{2k_2} \\ \quad \ldots\ldots\ldots\ldots\ldots \\ \xi_{n1},\, \xi_{n2},\, \ldots, \xi_{nk_n} \\ \quad \ldots\ldots\ldots\ldots\ldots \end{array} \tag{1.19}$$

We suppose that $k_n \to \infty$ as $n \to \infty$ and each line of the array $\{\xi_{nj}, 1 \le j \le k_n\}$ consists of independent random variables with *finite second moments*. The question then is to find necessary and sufficient conditions under which, if

$$S_n := \sum_{j=1}^{k_n} \xi_{nj}$$

then for appropriate sequences $\{a_n\}$, $\{b_n\}$,

$$\frac{S_n - a_n}{b_n} \xrightarrow{d} Z$$

where Z is a standard normal random variable.

Let $\alpha_{nj} := E\xi_{nj}$, $\sigma_{nj}^2 := \mathsf{Var}(\xi_{nj})$, and

$$B_n^2 := \sum_{j=1}^{k_n} \sigma_{nj}^2.$$

The double array (1.19) is called *elementary* if the sequence $\{B_n^2\}$ is *bounded* and, furthermore,

$$\lim_{n\to\infty} \frac{1}{B_n^2} \max_{1\le j\le k_n} \sigma_{nj}^2 \;=\; 0. \tag{1.20}$$

Intuitively, in order to have convergence to a normal limit it is clear that the contribution of each individual term to the sum S_n must be small. This is made precise in the following theorem.

Define the centered random variables

$$\bar{\xi}_{nj} := \xi_{nj} - \alpha_{nj}.$$

Theorem 1.22 (Lindeberg–Feller)**.** *With the above notation, the elementary double array of centered random variables $\{\bar{\xi}_{nj}\}$ satisfies*

$$\frac{1}{B_n} \sum_{j=1}^{k_n} \bar{\xi}_{nj} \;\xrightarrow{d}\; Z \quad \text{as } n \to \infty,$$

where Z is a standard normal random variable if and only if

$$\lim_{n\to\infty} \frac{1}{B_n^2} \sum_{j=1}^{k_n} E[\bar{\xi}_{nj}^2 \mathbf{1}(|\bar{\xi}_{nj}| > \epsilon B_n)] \;=\; 0 \quad \text{for any } \epsilon > 0. \tag{1.21}$$

(1.21) are known as the *Lindeberg condition.*

Note that the Lindeberg condition (1.21) implies (1.20).

1.7 Infinitely Divisible Distributions

Definition 1.23. *A probability law μ is called* infinitely divisible *if for any n there exists a probability law μ_n (depending on n) such that $\mu = \mu_n^{*n}$, the convolution of μ_n n times with itself.*

Equivalently, μ is infinitely divisible if, assuming the random variable X to follow μ, for each $n \in \mathbb{N}$ there exist independent, identically distributed random variables $\xi_1^{(n)}, \ldots, \xi_n^{(n)}$ (whose distribution depends on n) such that

$$X \;=\; \xi_1^{(n)} + \cdots + \xi_n^{(n)}.$$

The fundamental result on infinitely divisible distributions is the following

Theorem 1.24 (Lévy–Khinchine Representation)**.** *The probability law μ in $\mathbb{R}$ is infinitely divisible if and only if its characteristic function has the form*

$$\phi(t) \;:=\; \int_{\mathbb{R}} e^{itx}\mu(dx) \;=\; e^{\psi(t)}$$

where

$$\psi(t) \;=\; ibt - \frac{1}{2}\sigma^2 t^2 + \int_{\mathbb{R}} \left(e^{itx} - 1 - itx\mathbf{1}(|x| < 1)\right) \Pi(dx). \tag{1.22}$$

In the above representation $b \in \mathbb{R}$, $\sigma > 0$, and Π, the Lévy measure, is concentrated on $\mathbb{R} \setminus \{0\}$ and satisfies $\int_{\mathbb{R}}(1 \wedge x^2)\Pi(dx) < \infty$. Each triplet (b, σ, Π) uniquely specifies an infinite divisible law and vice versa.

We refer the reader to Gnedenko [69], Sato [123], and Applebaum [5] for background. When the infinitely divisible law μ has finite second moment then the logarithm of its characteristic function can be written in the form

$$\psi(t) \;=\; i\gamma t + \int_{-\infty}^{\infty} \frac{e^{itx} - 1 - itx}{x^2}\, dG(x) \tag{1.23}$$

where γ is real and G a nondecreasing bounded function on $\mathbb{R}$.

Sums of independent infinitely divisible random variables are also infinitely divisible. If a sequence of infinitely divisible random variables converges in distribution to a finite limit then this limit is also infinitely divisible. The theorems that follow give sufficient conditions for a given sequence of infinitely divisible distribution functions to converge to (an infinite divisible) distribution.

Theorem 1.25 (Limit Theorem)**.** *Suppose that F_n is a sequence of infinitely divisible distributions with finite second moment. In order for the sequence $\{F_n\}$ to converge in distribution as $n \to \infty$, to some distribution function F and for their means and variances to converge to the mean and variance of the limit law, it is necessary and sufficient that there exist a constant γ and the function $G(x)$, refering to representation (1.23), for which, as $n \to \infty$,*

(i) G_n *converges in distribution to* G,

(ii) $G_n(\infty) - G_n(-\infty) \to G(\infty) - G(-\infty)$,

(iii) $\gamma_n \to \gamma$, *where* γ_n *and* G_n *are defined by formula (1.23), for the law* F_n, *and the constant* γ *and the function* G *define, by the same formula, the limit law* F.

Also, the following holds (see Gnedenko [69])

Theorem 1.26. *Every law that is a limit law for the sums in an elementary system is infinitely divisible with finite variance and, conversely, every infinitely divisible law with finite variance is a limit law for the distributions functions of the sums of some elementary system.*

The following theorem connects the convergence in distribution of an elementary array to that of a related elementary array of *infinitely divisible* random variables. Using the notation of section 1.6.2 consider an *elementary array* $\{\xi_{nj} : j = 1, \dots, k_n\}_{n=1,2,\dots}$ and denote by F_{nj} the distribution of ξ_{nj}. Denote also by $\overline{F}_{nj}$ the distribution of the corresponding *centered* random variables: $\overline{F}_{nj}(x) = F(x + \alpha_{nj})$, $x \in \mathbb{R}$.

Theorem 1.27. *The sequence* $\{\xi_n\}$ *of random variables defined in terms of the elementary array (1.19) via the sums*

$$\xi_n = \xi_{n_1} + \xi_{n_2} + \cdots + \xi_{nk_n}$$

converges in distribution as $n \to \infty$ *to a limit random variable if and only if the corresponding sequence of infinitely divisible laws whose characteristic function have logarithms given by the formula*

$$\psi_n(t) = \sum_{j=1}^{k_n} \left\{ itE[\xi_{nj}] + \int \left(e^{itx} - 1\right) d\overline{F}_{nj}(x) \right\}$$

to converge to a limit law. Furthermore, the limit laws for both sequences coincide.

Finally we state a theorem that guarantees convergence to the Poisson Law for elementary arrays or random variables whose means and variancs satisfy an appropriate limiting condition.

Theorem 1.28. *Let* $\{\xi_{nj} : j = 1, \dots, k_j\}_{n=1,2,\dots}$ *elementary system that obeys the conditions*

$$\sum_{j=1}^{k_n} E\xi_{nj} \to \lambda \quad \textit{and} \quad \sum_{k=1}^{k_n} \mathsf{Var}(\xi_{nj}) \to \lambda$$

for some $\lambda > 0$. *The laws of the sums* $\xi_n = \xi_{n1} + \xi_{n2} + \cdots + \xi_{nk_n}$ *converge to the law*

$$P(x) = \begin{cases} 0, & \text{for } x < 0 \\ \sum_{k=0}^{[x]} e^{-\lambda} \frac{\lambda^k}{k!}, & \textit{for } x \geq 0 \end{cases}$$

if and only if for any $\tau > 0$

$$\sum_{j=1}^{k_n} \int_{|x-1|>\tau} x^2 dF_{nj}(x + E[\xi_{nj}]) \to 0$$

as $n \to \infty$.

2. ON/OFF SOURCES AND RUIN PROBABILITIES

In this chapter we consider ON/OFF sources of the following type. ON and OFF periods alternate and are assumed to be i.i.d. positive random variables with finite mean and second moment. We shall designate by $\{X_i\}$ the sequence of OFF periods and by $\{Y_i\}$ the corresponding ON periods. We also define a process $\{\xi_t\}$ taking the values 0 or 1 according to whether t belongs to an ON or an OFF period. We will construct a stationary version of the process $\{\xi_t\}$ and focus on the covariance function $\rho(t) = \mathsf{Cov}(\xi_s, \xi_{s+t})$ and on the behavior of $\int_0^t \rho(u)du$ as $t \to \infty$.

In the second part of the chapter renewal reward processes and generalizations are considered. If $\{T_n\}$ is a renewal process with interevent sequence $\{X_i\}$ and corresponding reward at the end of each such interval Y_i, $N(t) = \sum_{i=1}^{\infty} \mathbf{1}(T_i \leq t)$ denotes the number of renewals up to time t and $\sum_{i=1}^{N(t)} Y_i$ the total reward up to this time. The Laplace transform $M(\theta, t) := E \exp\left(-\theta \sum_{i=1}^{N(t)} Y_i\right)$ of this reward and the corresponding Laplace transform $\widehat{M}(\theta, s) = \int_0^{\infty} e^{-st} M(\theta, t)dt$ is studied for these and other similar models.

2.1 ON/OFF Sources Modeled as Alternating Renewal Processes

In this section we analyze ON/OFF sources for which, during the ON period, the rate of the source is constant and equal to 1. ON and OFF cycles are assumed to be independent, identically distributed random variables. We examine various such models of increasing generality in terms of the structure of these cycles. In the simplest case, which we include for completeness, ON and OFF periods are independent, exponential random variables. In the more general, the ith cycle consists of a pair of non-negative random variables, (X_i, Y_i) corresponding to the OFF and ON periods respectively. The sequence of pairs $\{(X_i, Y_i)\}$ is i.i.d. but the components of a pair may be dependent random variables.

2.1.1 Exponentially distributed ON and OFF periods: Markovian Source

Assume that the ON periods are i.i.d. exponentially distributed random variables with rate μ and OFF periods i.i.d. independent of the ON periods and exponentially distributed with rate λ. Here it is easier to first consider the indicator process $\{\xi_t\}$ which is a two state Markov process, taking the value 1 in the ON state and the value 0 in the OFF state, with generator

$$Q = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}$$

and transition probability matrix

$$\begin{bmatrix} P_{00}(t) & P_{01}(t) \\ P_{10}(t) & P_{11}(t) \end{bmatrix} = \begin{bmatrix} \frac{\mu}{\lambda+\mu} & \frac{\lambda}{\lambda+\mu} \\ \frac{\mu}{\lambda+\mu} & \frac{\lambda}{\lambda+\mu} \end{bmatrix} + e^{-(\lambda+\mu)t} \begin{bmatrix} \frac{\lambda}{\lambda+\mu} & -\frac{\lambda}{\lambda+\mu} \\ -\frac{\mu}{\lambda+\mu} & \frac{\mu}{\lambda+\mu} \end{bmatrix}. \tag{2.1}$$

Suppose that the indicator process, $\{\xi_t\}$, is stationary. We begin by computing the (stationary) covariance

$$\rho(t) := \mathsf{Cov}(\xi_{s+t}, \xi_s) \;=\; E[\xi_{s+t}\xi_s] - (E\xi_{s+t})\,(E\xi_s)$$

Since $\{\xi_t\}$ is stationary and takes values in $\{0, 1\}$,

$$\begin{aligned} \rho(t) &= P\left(\xi_{t+s} = 1, \xi_s = 1\right) - \left(P(\xi_s = 1)\right)^2 \;=\; P(\xi_s = 1)\left(P_{11}(t) - P(\xi_s = 1)\right) \\ &= \frac{\lambda\mu}{(\lambda+\mu)^2}\, e^{-(\lambda+\mu)t}. \end{aligned} \tag{2.2}$$

From the above expressions, using the Fubini theorem, we can easily compute the expected cumulative input due to an ON/OFF source as

$$E\left[\int_0^t \xi_u du\right] \;=\; \frac{\lambda}{\lambda+\mu} t.$$

The second moment can also be easily computed as follows

$$\begin{aligned} \mathsf{Var}\left(\int_0^t \xi_s ds\right) &= E\left[\int_0^t \xi_u du\right]^2 - \left(\frac{\lambda t}{\lambda+\mu}\right)^2 \;=\; \int_0^t \int_0^t E\left[\xi_u\xi_w - \left(\frac{\lambda}{\lambda+\mu}\right)^2\right] du dw \\ &= \frac{2\lambda}{\lambda+\mu} \int_{w=0}^t \int_{u=0}^w \left(P_{11}(w-u) - \frac{\lambda}{\lambda+\mu}\right) du dw, \\ &= \frac{2\lambda\mu}{(\lambda+\mu)^2} \int_{w=0}^t \int_{u=0}^w e^{-(w-u)(\lambda+\mu)} du dw \\ &= \frac{2\lambda\mu}{(\lambda+\mu)^3}\, t \;-\; \frac{2\lambda\mu}{(\lambda+\mu)^4}\left(1 - e^{-(\lambda+\mu)t}\right). \end{aligned} \tag{2.3}$$

2.1.2 FCLT-Weak Convergence

In order to be able to obtain approximations for expectations of functionals of the sample path such as the ruin probability we need to establish functional versions of the Central Limit Theorem for the free process

$$S(t) = \int_0^t \xi_s ds - ct \tag{2.4}$$

under appropriate scaling, then examine

$$\Psi\left(u\right) = P\left(\sup_{t\geq 0} S(t) > u\right).$$

Define the scaled free process as

$$S_n(t) := \frac{\int_0^{nt} \xi_s ds - c_n nt}{\sqrt{n}} \tag{2.5}$$

where $\{c_n\}$ a real sequence to be chosen appropriately, and $\sqrt{n}$ the standard CLT scaling. The conditions $\{c_n\}$ must satisfy become apparent if we write the right hand side of (2.5) as

$$\frac{1}{\sqrt{n}}\left(\int_0^{nt}\xi_s ds - nt\frac{\lambda}{\lambda+\mu}\right) + \sqrt{n}t\left(\frac{\lambda}{\lambda+\mu} - c_n\right)$$

where in the above expression the first term will be shown to converge to $\{\sigma W(t)\}$ with

$$\sigma = \sqrt{\frac{2\lambda\mu}{(\lambda+\mu)^3}}, \tag{2.6}$$

i.e. Brownian motion with variance constant σ. The requirement that the second term converge to $-t$ as $n\to\infty$ determines the asymptotic behavior of $\{c_n\}$. Using renewal theoretic arguments one can show that

$$E\left[\int_0^{nt}\xi_s ds - \int_0^{nt}E\xi_s ds\right]^2 = nt\frac{2\lambda\mu}{(\lambda+\mu)^3} + h(t), \qquad \text{where } h(t) = O(1) \text{ as } t\to\infty.$$

(See also Gnedenko and Korolyev [71].) Hence

$$\mathsf{Var}\left(\int_0^{nt}\xi_s ds - nt\frac{\lambda}{\lambda+\mu}\right) \simeq nt\frac{2\lambda\mu}{(\lambda+\mu)^3},$$

which justifies (2.6). In order to determine the asymptotic behavior of $\frac{1}{\sqrt{n}}\left(\int_0^{nt}\xi_s ds - nt\frac{\lambda}{\lambda+\mu}\right)$ we rewrite it as

$$\frac{1}{\sqrt{n}}\left(\int_0^{nt}\xi_s ds - \int_0^{nt}E\xi_s ds + \int_0^{nt}E\xi_s ds - nt\frac{\lambda}{\lambda+\mu}\right)$$

or

$$\frac{1}{\sqrt{n}}\left(\int_0^{nt}\xi_s ds - \int_0^{nt}E\xi_s ds\right) + \frac{1}{\sqrt{n}}\left(\int_0^{nt}E\xi_s ds - nt\frac{\lambda}{\lambda+\mu}\right). \tag{2.7}$$

The first term above converges to $\{\sigma W(t)\}$ by the FCLT for Markov Processes (Ethier and Kurz [56]). The second term can easily be evaluated using (2.1) as follows

$$\frac{1}{\sqrt{n}}\left(\int_0^{nt}E\xi_s ds - nt\frac{\lambda}{\lambda+\mu}\right) = -\frac{1}{\sqrt{n}}\frac{\lambda}{(\lambda+\mu)^2}\left[1 - e^{-(\lambda+\mu)nt}\right].$$

The right hand side of the above equation clearly goes to zero as $n\to\infty$. Thus we can state the following

Proposition 2.1. *Consider the scaled free process $S_n(t)$ derived from the two–state Markov process described above. Suppose that the scaling sequence $\{c_n\}$satisfies the following conditions*

C1. $c_n > \frac{\lambda}{\lambda+\mu}\ \forall n.$

C2. $\sqrt{n}\left(\frac{\lambda}{\lambda+\mu} - c_n\right) \to -1.$

C3. c_n *decreases monotonically to* $\frac{\lambda}{\lambda+\mu}$.

Then, the scaled process converges in distribution to Brownian motion with drift -1, *i.e.*

$$S_n(t) \xrightarrow{d} \sigma W(t) - t. \tag{2.8}$$

2.2 Asymptotic variance

The following proposition is standard and will be useful in the sequel.

Proposition 2.2. *Suppose that $\{\xi_t\}$, $t \geq 0$ is a stationary process with values in $\{0,1\}$ and let $\rho(t) := \mathsf{Cov}(\xi_0, \xi_t)$. Let $S^0(t) = \int_0^t \xi_s ds$ and $\sigma_t^2 := \mathsf{Var}\left(S^0(t)\right)$. Let also $\widehat{\rho}(s) := \int_0^\infty e^{-st}\rho(t)dt$ define the Laplace transform of the covariance function. Then, if $\widehat{\rho}(s)$ is analytic for $\Re(s) > 0$ the limit $\sigma^2 := \lim_{t\to\infty} \frac{1}{t}\sigma_t^2$ exists and is equal to $2\int_0^\infty \rho(u)du$. This last integral converges under the assumptions of the theorem and $\lim_{s\to 0, \Re(s)>0} \widehat{\rho}(s) = \int_0^\infty \rho(u)du$.*

Proof. We begin with the following expression of the variance.

$$\begin{aligned}\sigma_t^2 &= E\left(\int_0^t \xi_s ds\right)^2 - \left(\int_0^t E\xi_s ds\right)^2 = 2\left[\int_{s=0}^t \left(\int_{u=0}^s \left(E[\xi_u\xi_s] - E[\xi_u]E[\xi_s]\right) du\right) ds\right] \\ &= 2\int_0^t \left[\int_0^s \rho(s-u)du\right] ds = 2\int_0^t \left[\int_0^s \rho(u)du\right] ds = 2\int_0^t (t-u)\rho(u)du,\end{aligned}$$

the third equation following by the stationarity of the underlying process, $\{\xi_t\}$. Thus

$$\frac{1}{t}\sigma_t^2 = 2\int_0^\infty \mathbf{1}(u \leq t)\left(1 - \frac{u}{t}\right)\rho(u)du$$

and, therefore,

$$\sigma^2 := \lim_{t\to\infty} \frac{1}{t}\sigma_t^2 = 2\int_0^\infty \rho(u)du$$

provided that the integral converges. This last proviso is guaranteed by a Tauberian theorem (see Korevaar [94]) in view of the assumed behavior of $\widetilde{\rho}(s)$ at zero. Settting $C := \int_0^\infty \rho(u)du < \infty$

$$\sigma^2 = 2\lim_{s\to 0}\widehat{\rho}(s) = 2\int_0^\infty \rho(t)dt = 2C. \tag{2.9}$$

□

2.3 Alternating Renewal Processes. Exponentially distributed OFF periods, generally distributed ON periods

We consider an alternating renewal process consisting of ON and OFF periods. In this section we will assume that the OFF periods are exponentially distributed with rate λ while the ON (or active) periods have an arbitrary distribution G with support on the nonnegative reals and *finite mean* $m := \int_0^\infty y dG(y)$.

We consider again the indicator process $\{\xi_t\}$ which is equal to 1 during the ON and 0 during the OFF periods. We suppose again that under the probability measure P this process is stationary. From a standard regenerative argument it follows that $P(\xi_s = 0) = \frac{1/\lambda}{(1/\lambda)+m} = \frac{1}{1+a}$ and $P(\xi_s = 1) = \frac{a}{1+a}$.

The covariance function is given by

$$\rho(t) := E[\xi_{t+s}\xi_s] - (E[\xi_t])^2, \quad 0 \le s \le t.$$

In order to obtain an expression for it we note that

$$\begin{aligned} E\left[(1-\xi_s)(1-\xi_{t+s})\right] &= P(\xi_s = 0, \xi_{t+s} = 0) \;=\; P(\xi_{t+s} = 0 \mid \xi_s = 0)P(\xi_s = 0) \\ &= \frac{1}{1+a}\left[1 - P_0(\xi_t = 1)\right]. \end{aligned} \tag{2.10}$$

We have thus expressed $E\left[(1-\xi_s)(1-\xi_{t+s})\right]$ in terms of $P_0(\xi_t = 1) := P(\xi_t = 1|\xi_0 = 0)$. If we also use the fact that

$$E\left[(1-\xi_s)(1-\xi_{t+s})\right] = 1 - E[\xi_s] - E[\xi_{t+s}] + E[\xi_s\xi_{t+s}]$$

together with the stationarity of the process which implies that $E\left[\xi_s\right] = E\left[\xi_{t+s}\right] = \frac{1}{1+a}$, we conclude that

$$E\left[(1-\xi_s)(1-\xi_{t+s})\right] = 1 - \frac{2a}{1+a} + \rho(t) + \left[E\xi_t\right]^2. \tag{2.11}$$

Combining (2.10) and (2.11) and using the notation $P_0(t) := P_0(\xi_t = 1)$ we obtain

$$\rho(t) = \frac{a}{1+a} - \frac{1}{1+a}P_0(t) - \left(\frac{a}{1+a}\right)^2. \tag{2.12}$$

Consider a version of the alternating renewal process beginning at time 0 with an exponential (λ) random variable X_1 (the first OFF period), followed by an ON period Y_1, independent of X_1, with distribution G. This first cycle is followed by the next cycle, (X_2, Y_2), independent of the first and so forth. Let $H(x) := P(X_1 + Y_1 \le x)$ denote the distribution of the length of each cycle. Then $P_0(t)$ satisfies the renewal equation

$$P_0(t) \;=\; P(X_1 < t \le X_1 + Y_1) + \int_0^t P_0(t-u)dH(u). \tag{2.13}$$

Denote by $\widehat{p}_0(s) := \int_0^\infty e^{-st}P_0(t)dt$ the Laplace transform of $P_0(t)$. Also $\widehat{H}(s) = \int_0^\infty e^{-su}dH(u) = \frac{\lambda}{\lambda+s}\widehat{G}(s)$ denotes the Laplace transform of the distribution H. Taking Laplace transforms in (2.13) we obtain

$$\begin{aligned} \widehat{p}_0(s) &= \int_0^\infty P(X_1 < t \le X_1 + Y_1)e^{-st}dt + \widehat{p}_0(s)\widehat{H}(s) \\ &= E\int_{X_1}^{X_1+Y_1} e^{-st}dt \;+\; \widehat{p}_0(s)\widehat{G}(s)\frac{\lambda}{\lambda+s} \;=\; \frac{E[e^{-sX_1}] - E[e^{-s(X_1+Y_1)}]}{s} + \widehat{p}_0(s)\widehat{G}(s)\frac{\lambda}{\lambda+s} \\ &= \frac{\lambda}{\lambda+s}\frac{1-\widehat{G}(s)}{s} + \widehat{p}_0(s)\widehat{G}(s)\frac{\lambda}{\lambda+s}. \end{aligned}$$

From the above we obtain the expression

$$\widehat{p}_0(s) \;=\; \frac{1}{s}\,\frac{a\widehat{G}_I(s)}{1+a\widehat{G}_I(s)}, \tag{2.14}$$

In (2.14) $a = \lambda m$ where $m = \int_0^\infty x dG(x)$ is the mean of G and, as usual, $G_I(x) := \frac{1}{m}\int_0^x \overline{G}(y)dy$ is the *integrated tail distribution* that corresponds to G with Laplace transform given by $\widehat{G}_I(s) = \frac{1-\widehat{G}(s)}{ms}$.

The Laplace transform of the covariance function is then given by

$$\widehat{\rho}(s) := \int_0^\infty e^{-st}\rho(t)dt = \frac{1}{s}\frac{a}{1+a}\left(1 - \frac{\widehat{G}_I(s)}{1+a\widehat{G}_I(s)}\right) - \frac{1}{s}\left(\frac{a}{1+a}\right)^2$$

or equivalently

$$\widehat{\rho}(s) = \frac{a}{(1+a)^2}\frac{1}{s}\frac{1-\widehat{G}_I(s)}{1+a\widehat{G}_I(s)}. \tag{2.15}$$

Since for all $s > 0$ we have that $\widehat{G}_I(s) < 1$ we can use the following series expansion for the Laplace transform of the covariance function

$$\widehat{\rho}(s) = \frac{a}{(1+a)^2}\frac{1-\widehat{G}_I(s)}{s}\sum_{k=0}^{\infty}(-a)^k\widehat{G}_I^k(s).$$

This, if we assume the existence of the second moment of G and we define the *doubly integrated tail distribution* G_{II} with Laplace transform $\widehat{G}_{II}(s) = \frac{1-\widehat{G}_I(s)}{sm_I}$, where $m_I = \frac{E[Y^2]}{2EY}$, can be written as

$$\widehat{\rho}(s) = \frac{am_I}{(1+a)^2}\sum_{k=0}^{\infty}(-a)^k\widehat{G}_{II}(s)\widehat{G}_I^k(s).$$

In view of 2.15 the value of the limit C can be obtained by appealing to (2.9) as follows

$$C = \lim_{s\to 0}\widehat{\rho}(s) = \frac{a}{(1+a)^2}\frac{\lim_{s\to 0}\frac{1-\widehat{G_I}(s)}{S}}{1+a\lim_{s\to 0}\widehat{G_I}(s)} = \frac{a}{(1+a)^2}\frac{EY^2}{2EY}.$$

The last equation above was based on the fact that $\lim_{s\to 0}\frac{1-\widehat{G}_I(s)}{s} = \frac{EY^2}{2EY}$, the mean of the integrated tail distribution. Finally, if we have OFF periods which also follow some arbitrary distribution, say F, (the ON periods following the distribution G), the Laplace transform of the covariance function is given by the formula

$$\widehat{\rho}(s) = \frac{as}{(1+a)^2}\frac{\widehat{F}_I(s)\widehat{G}_{II}(s)m_{G_I} + a\widehat{G}_I(s)\left\{m_{F_I}\widehat{F}_{II}(s) - \widehat{F}_I(s)m_F\right\}}{\widehat{F}_I(s) + a\widehat{F}(s)G(s)}.$$

The limit is once more computed by using (2.9)

$$C = \lim_{s\to 0}\widehat{\rho}(s) = \frac{a}{(1+a)^2}\frac{m_{G_I} + a\left(m_{F_I} - m_F\right)}{1+a} = \frac{a}{(1+a)^3}\frac{EY}{2}\left[\frac{EY^2}{[EY]^2} + \frac{EX^2}{[EX]^2} - 2\right].$$

2.4 Alternating Renewal ON/OFF Sources With General Distributions

In this section we assume that the OFF periods $\{X_i; i = 1, 2, \ldots\}$ are i.i.d. random variables with distribution F and the ON periods $\{Y_i; i = 1, 2, \ldots\}$ are i.i.d. with distribution G. We will further assume that the two sequences of random variables are independent of each other and

we will denote by $H := F * G$ the convolution of the distributions of the OFF and ON periods which, under our assumptions, is the distribution of the duration of each OFF/ON cycle.

As before, $\{\xi_t\}$ is stationary under the probability measure P and, form a standard regenerative argument we have $P(\xi_t = 1) = \frac{a}{1+a}$ where $a = \frac{m_G}{m_F}$. We will also define the conditional probability measure P_{1e} defined via $P_{1e}(A) = P(A|\xi_0 = 1)$ for all $A \in \mathcal{F}$. In other words, P_{1e} is the stationary probability measure conditioned on being in an ON interval. Intuitively it is clear that the remaining life of the ON interval at time 0 follows the equilibrium distribution $G_I(x) := \frac{1-G(x)}{m_G}$. This can of course be made rigorous by means of a standard regenerative argument. We will denote this first "remaining life" ON interval by Y_I.

The covariance function is given by

$$\begin{aligned}\rho(t) &= P(\xi_t = 1, \xi_{t+s} = 1) - P(\xi_t = 1)P(\xi_{t+s} = 1) \\ &= P(\xi_{t+s} = 1|\xi_t = 1)P(\xi_t = 1) - P(\xi_t = 1)P(\xi_{t+s} = 1) \\ &= \frac{a}{1+a}\left(P_{1e}\left(\xi_s = 1\right) - \frac{a}{1+a}\right).\end{aligned}$$

Suppose also that, under the probability measure P_0 the process is an alternating renewal process starting with an OFF period, X_1, The probability that the source in the ON state at time t, $P_0(t) := P_0(\xi_t = 1)$, can be obtained via a standard renewal theoretic argument as the solution of the following renewal equation

$$P_0(t) = \overline{G \star F}(t) - \overline{F}(t) + \int_0^t P_0(t-u)dH(u) \tag{2.16}$$

(see e.g.[7] for background on renewal theory). Upon taking Laplace transforms in (2.16) we obtain $\widehat{p}_0(s) = \frac{1}{s}(\widehat{F}(s) - \widehat{F}(s)\widehat{G}(s)) + \widehat{p}_0(s)\widehat{H}(s)$ and, because of the independence of the ON and OFF periods which implies that $\widehat{H}(s) = \widehat{F}(s)\widehat{G}(s)$, we have that

$$\widehat{p}_0(s) = \frac{\widehat{F}(s)}{s}\frac{1-\widehat{G}(s)}{1-\widehat{F}(s)\widehat{G}(s)}. \tag{2.17}$$

Consider now process under P_{1e}. It becomes a *modified alternating renewal process* that begins with an ON period, Y_I, which is distributed according to the integrated tail distribution G_I. Then $P_{1e}(t) := P_{1e}(\xi_t = 1)$ satisfies the equation

$$P_{1e}(t) \;=\; \overline{G}_I(t) + \int_0^t P_0(t-u)dG_I(u).$$

Taking Laplace transforms in the above equation (with $\widehat{p}_{1e}(s) = \int_0^\infty P_{1e}(t)e^{-st}dt$) we obtain

$$\widehat{p}_{1e}(s) \;=\; \frac{1-\widehat{G}_I(s)}{s} + \widehat{p}_0(s)\widehat{G}_I(s). \tag{2.18}$$

Taking into account (2.17) this gives

$$\widehat{p}_{1e}(s) = \frac{1-\widehat{G}_I(s)}{s} + \widehat{F}(s)\widehat{G}_I(s)\frac{1}{s}\frac{1-\widehat{G}(s)}{1-\widehat{F}(s)\widehat{G}(s)} \tag{2.19}$$

which can also be written as $\widehat{p}_{1e}(s) = \frac{1-\widehat{G}_I(s)}{s} + \widehat{F}(s)\widehat{G}_I(s)\frac{a\widehat{G}_I(s)}{\widehat{F}_I(s)+a\widehat{F}(s)\widehat{G}_I(s)}$. Hence the Laplace

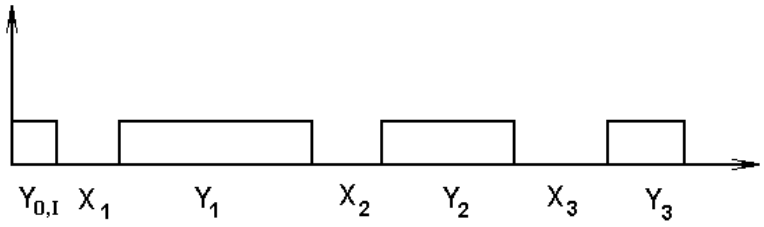


Fig. 2.1: A stationary ON/OFF process conditioned on being ON at time zero

transform of $\rho(t)$ is

$$\widehat{\rho}(s) = \frac{a}{1+a}\left(\widehat{p}_{1e}(s) - \frac{1}{s}\,\frac{a}{1+a}\right)$$

which, expressed in terms of the Laplace transforms of the OFF and ON periods becomes

$$\widehat{\rho}(s) = \frac{a}{1+a}\left(\frac{1-\widehat{G}_I(s)}{s} + \frac{1}{s}\,\frac{1-\widehat{G}(s)}{1-\widehat{F}(s)\widehat{G}(s)}\widehat{F}(s)\widehat{G}_I(s) - \frac{1}{s}\,\frac{a}{1+a}\right).$$

If we assume that the distributions F and G have finite second moments, the integrated tail distributions can be defined for F_I as $F_{II}(x) := \int_0^x \frac{1-F_I(x)}{m_F}dx$ with mean $m_{F_I} := \int_0^\infty x\frac{1-F_I(x)}{m_F}dx = \frac{E[X^2]}{2EX}$ and Laplace transform $\widehat{F}_{II}(s) = \frac{1-\widehat{F}_I(s)}{sm_{F_I}}$, with similar definitions for G.

$$\widehat{\rho}(s) = \frac{a}{(1+a)^2}\,\frac{\widehat{F}_I(s)\widehat{G}_{II}(s)m_{G_I} + a\widehat{G}_I(s)\left\{m_{F_I}\widehat{F}_{II}(s) - \widehat{F}_I(s)m_F\right\}}{\widehat{F}_I(s) + a\widehat{F}(s)G(s)}.$$

2.5 Independent Cycles with Dependent OFF and ON periods

Consider a sequence of i.i.d. pairs of nonnegative random variables $\{(X_i, Y_i)\}_{i\in\mathbb{Z}}$. The joint distribution of the elements of the pair is $F(x,y) := P(X_1 \le x, Y_1 \le y)$, $x, y \ge 0$. As before, the X_i's correspond to the source's OFF periods while the Y_i's to the ON periods. Let $\Phi(s,u) := \int_0^\infty \int_0^\infty e^{-sx-uy}dF(x,y)$ denote the joint Laplace transform of the OFF and ON period. We will also use the notation $\Phi^X(s) := \Phi(s,0) = Ee^{-sX_1}$ for the Laplace transform of the OFF period and similarly $\Phi^Y(u)$ for that of the ON period.

Let $\{T_n\}$ denote the point process of beginnings of OFF periods, defined via $T_0 = 0$, and $T_{n+1} = T_n + X_n + Y_n$, $n = 0, 1, 2, \ldots$. We may also extend the point process for negative indices. We may set

$$T_n = \begin{cases} \sum_{k=0}^{n-1}(X_k + Y_k), & n = 1, 2, \ldots \\ 0, & n = 0 \\ -\sum_{k=1}^{n} X_{-k} + Y{-k}, & n = -1, -2, \ldots \end{cases}$$

The process $\{\xi_t\}$, $t \in \mathbb{R}$ can then be defined as

$$\xi_t \;=\; \sum_{n\in\mathbb{Z}} \mathbf{1}(T_n + X_n < t \le T_{n+1}). \tag{2.20}$$

We will suppose that $\{T_n\}$ is *stationary* under the probability measure P and will denote by P^0 the *Palm transformation* of P under $\{T_n\}$ (see [12]). Correspondingly we denote by E^0 expectation with respect to P^0. Intuitively, the Palm probability of an event is its conditional probability given that the time origin coincides with a point of the process $\{T_n\}$. (In other words under P^0 we have the "ordinary version" of the renewal process.) In particular $F(x,y) = P^0(X_0 \le x, Y_0 \le y)$. Consider also the process

$$\chi(t,u,s) := \sum_{n\in\mathbb{Z}} \mathbf{1}(T_n + X_n < t \le T_{n+1}) e^{-u(T_{n+1}-t)} + \mathbf{1}(T_n < t < T_n + X_n) e^{-s(T_n+X_n-t)}.$$

The stationary expectation of such a process in terms of the Palm probability is given by the Palm inversion formula

$$E\chi(0,u,s) \;=\; \frac{1}{E^0[T_1 - T_0]} E^0\left[\int_{T_0}^{T_1} \chi(t,u,s)dt\right].$$

Under P^0, $T_0 = 0$, $T_1 = X_0 + Y_0$, and

$$\begin{aligned}\int_{T_0}^{T_1} \chi(t,u,s)dt &= \int_0^{X_0} e^{-(X_0-t)s}dt + \int_{X_0}^{X_0+Y_0} e^{-(X_0+Y_0-t)u}dt \\ &= \frac{1}{s}\left[1 - e^{-sX_0}\right] + \frac{1}{s}\left[e^{-uX_0} - e^{-u(X_0+Y_0)}\right]\end{aligned}$$

and therefore

$$E\chi(0,u,s) \;=\; \frac{1}{E^0[X_0+Y_0]}\left(\frac{1}{s}E^0\left[1 - e^{-sX_0}\right] + \frac{1}{s}E^0\left[e^{-uX_0} - e^{-u(X_0+Y_0)}\right]\right).$$

If we denote by $P_0(t) := P^0(\xi_t = 1)$ the Palm probability that at time t the process is in the ON state then the Laplace transform of this function is given by

$$\widehat{p}_0(s) = \int_0^\infty e^{-st} P_0(t)dt = E^0 \int_0^\infty \mathbf{1}(\xi_t = 1)dt$$

the second equation following by an application of the Fubini theorem. Thus,

$$\widehat{P}_0(s) = E\left[\int_0^{X+Y} e^{-st}\mathbf{1}(\xi_t = 1)dt\right] + E\left[\int_{X+Y}^\infty e^{-st}\mathbf{1}(\xi_t = 1)dt\right]$$

or

$$\widehat{P}_0(s) = E\left[e^{-sX_0}\right] - E\left[e^{-s(X_0+Y_0)}\right] + E\left[e^{-s(X_0+Y_0)}\right]\widehat{P}_0(s),$$

whence we obtain that

$$\widehat{P}_0(s) = \frac{1}{s}\,\frac{\Phi^X(s) - \Phi(s,s)}{1 - \Phi(s,s)}. \tag{2.21}$$

Finally, suppose that the process starts with an ON period in equilibrium, i.e. that the first on period is distributed according to the corresponding integrated tail distribution. Denoting by $\widehat{p}_{1e}(s)$ the Laplace transform of α_t in that case we obtain

$$\widehat{P}_{1e}(s) = \frac{1 - \Phi_I^Y(s)}{s} + \Phi_I^Y(s)\widehat{p}_0(s)$$

and taking into account (2.21)

$$\widehat{P}_{1e}(s) = \frac{1 - \Phi_I^Y(s)}{s} + \frac{\Phi_I^Y(s)}{s}\,\frac{\Phi^X(s) - \Phi(s,s)}{1 - \Phi(s,s)}. \tag{2.22}$$

In this last case the Laplace transform of the covariance function is

$$\begin{aligned}
\widehat{\rho}(s) &= \frac{a}{1+a}\left(\widehat{P}_{1e}(s) - \frac{1}{s}\frac{a}{1+a}\right) \\
&= \frac{a}{1+a}\left(\frac{1-\Phi_I^Y(s)}{s} + \frac{\Phi_I^Y(s)}{s}\frac{\Phi^X(s)-\Phi(s,s)}{1-\Phi(s,s)} - \frac{1}{s}\frac{a}{1+a}\right) \\
&= \frac{a}{1+a}\left(\frac{1}{1+a}\frac{1}{s} - \frac{\Phi_I^Y(s)\left[1-\Phi^X(s)\right]}{s(1-\Phi(s,s))}\right) \\
&= \frac{a}{(1+a)^2}\,\frac{1-\Phi(s,s)-\Phi_I^Y(s)\left[1-\Phi^X(s)\right](1+a)}{s(1-\Phi(s,s))}.
\end{aligned}$$

But

$$\begin{aligned}
1-\Phi(s,s) &= 1-E\left[e^{-s(X+Y)}\right] = E\left[1-e^{-s(X+Y)}\right] = E\left[1-e^{-sY}+e^{-sY}-e^{-s(X+Y)}\right] \\
&= E\left[(1-e^{-sY})+e^{-sY}(1-e^{-sX})\right]
\end{aligned}$$

and $1+a = \frac{m_Y+m_X}{m_X}$ since $a = \frac{m_Y}{m_X}$. Thus the Laplace transform of the covariance function becomes

$$\begin{aligned}
\widehat{\rho}(s) &= \frac{a}{(1+a)^2}\,\frac{E\left(\frac{1-e^{-sY}}{s}\right)+E\left(e^{-sY}\frac{1-e^{-sX}}{s}\right)-\Phi_I^Y(s)\Phi_I^X(s)\left[m_Y+m_X\right]}{E\left(\frac{1-e^{-sY}}{s}\right)+E\left(e^{-sY}\frac{1-e^{-sX}}{s}\right)} \\
&= \frac{a}{(1+a)^2}\,\frac{\Phi_I^Y(s)m_Y E\left(1-\frac{1-e^{-sX}}{sm_X}\right)+m_X E\left[\frac{1-e^{-sX}}{sm_X}\left(\frac{e^{-sY}-\Phi_I^Y(s)}{s}\right)s\right]}{\Phi_I^Y(s)m_Y + E\left(e^{-sY}\frac{1-e^{-sX}}{sm_X}\right)m_X}.
\end{aligned}$$

So,

$$\begin{aligned}
\widehat{\rho}(s) &= \frac{a}{(1+a)^2}\,\frac{\Phi_I^Y(s)m_Y E\left(\frac{1-\frac{1-e^{-sX}}{sm_X}}{s}\right)+m_X E\left[\frac{1-e^{-sX}}{sm_X}\left(\frac{e^{-sY}-1+1-\Phi_I^Y(s)}{s}\right)\right]}{\Phi_I^Y(s)m_Y + E\left(e^{-sY}\frac{1-e^{-sX}}{sm_X}\right)m_X} \\
&= \frac{a}{(1+a)^2}\,\frac{\Phi_I^Y(s)m_Y\Phi_{II}^X(s)\frac{EX^2}{2m_X}+EX\frac{EY^2}{2m_Y}\Phi_{II}^Y(s)\Phi_I^X(s)-m_Xm_YE\left[\frac{1-e^{-sX}}{sm_X}\left(\frac{1-e^{-sY}}{sm_Y}\right)\right]}{\Phi_I^Y(s)m_Y+E\left(e^{-sY}\frac{1-e^{-sX}}{sm_X}\right)m_X}.
\end{aligned}$$

It only remains to take the limit as $s\to 0$ in the above expression for $\widehat{\rho}(s)$. We then have

$$\begin{aligned}
\lim_{s\to 0}\widehat{\rho}(s) &= \frac{a}{(1+a)^2}\left[\frac{EY\frac{EX^2}{2EX}+EX\frac{EY^2}{2EY}-EXEY\frac{E[XY]}{EXEY}}{EY+EX}\right] \\
&= \frac{a}{(1+a)^2}\frac{EXEY}{2\left[EY+EX\right]}\left[\frac{EX^2}{[EX]^2}+\frac{EY^2}{[EY]^2}-2\frac{E[XY]}{EXEY}\right] \\
&= \frac{1}{2}\frac{(EXEY)^2}{(EX+EY)^3}E\left[\frac{X}{EX}-\frac{Y}{EY}\right]^2.
\end{aligned}$$

We have thus obtained the interesting formula

$$C := \int_0^\infty \rho(u)du = \frac{1}{2}\frac{(EXEY)^2}{(EX+EY)^3}E\left[\frac{X}{EX}-\frac{Y}{EY}\right]^2 \tag{2.23}$$

for any ON/OFF process with independent cycles but possible dependence between the ON and OFF period in the same cycle.

2.6 Ruin probabilities for risk processes with ON/OFF sources

Here we use the above estimates in order to obtain asymptotic expressions for the ruin probability in a risk process where the input is an ON/OFF process of the type described in the previous section. The free reserves process described by

$$S(t) = \int_0^t \xi_s ds - ct$$

with initial capital u. The probability of ruin is defined as usual as $\Psi(u) = P\left\{\sup_{t\geq 0} S(t) > u\right\}$ and our goal is to find suitable asymptotic estimates for this ruin probability as $u \to \infty$ in the spirit of the classical diffusion approximation to the Cramér-Lundberg model (see for instance Iglehart [78], Grandell [73], and Asmussen and Steffensen [11]). For this purpose, assuming as always the stationarity of $\{\xi_s\}$ and using the fact that $E\xi_s = \frac{a}{1+a}$, we examine the family of processes

$$\widetilde{S}_n(t) := \frac{1}{\sqrt{n}}\left(\int_0^{nt} \xi_s ds - \frac{a}{1+a} nt\right). \tag{2.24}$$

Functional Central Limit Theorems for ON/OFF processes have been examined in the literature. We mention here Whitt [145] and [146] and Pang, Yang, and Zhou [117].

Suppose that $\{\xi_s\}$ is the ON/OFF process of section 2.5. We may then apply the results of Glynn and Whitt [67] in Theorem 1.10 it suffices to check the conditions (1.12) and (1.13). In terms of the notation of Theorem 1.10, ξ_t is given by (2.20),

$$\tau_1 = X_1 + Y_1, \quad Z_1 = \int_0^{X_1+Y_1} \mathbf{1}(X_1 < s) ds = Y_1, \quad \alpha = \frac{EY}{EX + EY}.$$

Thus

$$\begin{aligned} \widetilde{Z}_1 &= Y_1 - \alpha(X_1 + Y_1) = (1-\alpha)Y_1 - \alpha X_1 \leq X_1 + Y_1 \\ \widetilde{V}_1 &= \sup_{0\leq s\leq X_1+Y_1} \left|\int_0^s (\mathbf{1}(s > X_1) - \alpha) ds\right| = \max\{\alpha X_1, |(1-\alpha)Y_1 - \alpha X_1|\} \\ &\leq X_1 + Y_1. \end{aligned}$$

Since $EX_1^2 < \infty$, $EY_1^2 < \infty$, this means that

$$E\widetilde{Z}_1^2 \leq E(X_1 + Y_2)^2 \leq 2EX_1^2 + 2EX_2^2 < \infty. \tag{2.25}$$

For the same reason $E\widetilde{V}_1^2 < \infty$ and hence

$$t^2 P(\widetilde{V}_1 > t) = E[t^2 \mathbf{1}(\widetilde{V}_1 > t)] = E[\widetilde{V}_1^2 \mathbf{1}(\widetilde{V}_1 > t)] < \infty.$$

We may then use Dominated Convergence (with dominating r.v. $\widetilde{V}_1^2$) to conclude that

$$\lim_{t\to\infty} t^2 P(\widetilde{V}_1 > t) = E[\lim_{t\to\infty} \widetilde{V}_1^2 \mathbf{1}(\widetilde{V}_1 > t)] = 0. \tag{2.26}$$

From (2.25), (2.26), we see that conditions (1.12) and (1.13) of Theorem 1.10) are satisfied. Therefore this establishes the FCLT $\widetilde{S}_n(t) \xrightarrow{d} \sigma W_t$ where $\sigma^2 = 2C$ (with C given in (2.23)) is the asymptotic variance of the ON/OFF process.

Suppose now that in the definition of the family of processes $\{S_n(t)\}$ the sequence of premium rates $\{c_n\}$ satisfies the conditions

1. $c_n > \frac{a}{1+a}$,
2. $\sqrt{n}\left(\frac{a}{1+a} - c_n\right) \to -1$ as $n \to \infty$,
3. c_n decreases monotonically to $\frac{a}{1+a}$.

Then $S_n(t) \xrightarrow{d} \sigma W_t - t$. This leads to the diffusion approximation for the infinite horizon ruin probability

$$P\left(\sup_{t\geq 0} S^0(t) - ct > u\right) \approx \exp\left(-\frac{u}{C}\left(c - \frac{a}{1+a}\right)\right)$$

where $S^0(t) = \int_0^t \xi_s ds$ and $C = \int_0^\infty \rho(u)du$.

2.7 ON/OFF Sources with Independent Non Identical Cycles

Let (X_i, Y_i) be independent vectors *not necessarily identically distributed* with $X_i \geq 0$ w.p.1. Let f_i be the density of X_i. Then, the random variable $X_1 + \cdots + X_n$ has density $f_1 * \ldots * f_n$ where $*$ denotes convolution. In particular, $\widehat{F}_n(s) := E[e^{-sX_n}] = \int_0^\infty e^{-sx} f_n(x)dx$. Set $T_n = T_{n-1} + X_n$, $n = 1, 2, \ldots$, $T_0 = 0$. If $N(t) = \sum_{n=1}^\infty \mathbf{1}(T_n \leq t)$ then we are interested in the distribution of

$$V(t) := \sum_{n=1}^{N(t)} Y_n.$$

We first establish the following

Theorem 2.3. *With the above notation,*

$$\widehat{M}(\theta, s) := \int_0^\infty e^{-st} E\left[e^{-\theta \sum_{n=1}^{N(t)} Y_n}\right] dt = \sum_{n=0}^\infty \frac{1}{s}\left(1 - \widehat{F}_{n+1}(s)\right) \prod_{i=1}^n E[e^{-\theta Y_i - sX_i}]. \tag{2.27}$$

In particular, when the random vectors are identically distributed with $E[e^{-\theta Y_i - sX_i}] =: \psi(\theta, s)$,

$$\widehat{M}(\theta, s) = \frac{1 - \psi(0, s)}{s(1 - \psi(\theta, s))}. \tag{2.28}$$

Proof. Set

$$M(\theta, t) := Ee^{-\theta \sum_{n=1}^{N(t)} Y_n} = \sum_{n=0}^\infty E\left[e^{-\theta \sum_{i=1}^{N(t)} Y_i}; N(t) = n\right]$$

where

$$\{N(t) = n\} = \{X_1 + X_2 + \cdots + X_n \leq t < X_1 + X_2 + \cdots + X_n + X_{n+1}\}.$$

If $S_n = \{\vec{x} : x_1 + \cdots + x_n = x, x_i \geq 0,\ i = 1, 2, \ldots, n\}$ then, $E\left[e^{-\theta \sum_{i=1}^{N(t)} Y_i}; N(t) = n\right]$ can be written as

$$\begin{aligned}
&\int_0^t E\left[e^{-\theta \sum_{i=1}^{n} Y_i}; X_1 + \cdots + X_n = x\right] \overline{F}_{n+1}(t-x)dx \\
&= \int_0^t \left(\int_{S_n} E\left[e^{-\theta Y_1} \cdots e^{-\theta Y_n}; X_1 = x_1, \ldots, X_n = x_n\right] dx_1 \cdots dx_n\right) \overline{F}_{n+1}(t-x)dx \\
&= \int_0^t \left(\int_{S_n} E\left[e^{-\theta Y_1} \middle| X_1 = x_1\right] \cdots E\left[e^{-\theta Y_n} \middle| X_n = x_n\right] f_1(x_1) \cdots f_n(x_n) dx_1 \cdots dx_n\right) \overline{F}_{n+1}(t-x)dx \\
&= \int_0^t \left(\int_{S_n} \prod_{i=1}^{n} h_i(x_i) f_i(x_i) dx_i\right) \overline{F}_{n+1}(t-x)dx \qquad (2.29)
\end{aligned}$$

where $h_i(x_i) := E\left[e^{-\theta Y_i} | X_i = x_i\right]$. We may evaluate this integral by defining a new density function, $\widetilde{f}$ via the change of measure

$$\widetilde{f}_i(x_i) := \frac{h_i(x_i) f_i(x_i)}{\rho_i(\theta)}, \qquad x_i \geq 0 \qquad (2.30)$$

where $\rho_i(\theta) := E\left[e^{-\theta Y_i}\right]$. Note that $\int_0^\infty h_i(x_i) f_i(x_i) dx_i = \int_0^\infty E\left[e^{-\theta Y_i} | X_i = x_i\right] f_i(x_i) dx_i = \rho_i(\theta)$ and hence that (2.30) defines indeed a new probability measure (depending on θ). So, the expression in (2.29) becomes

$$\begin{aligned}
&m \int_0^t \left(\int_{S_n} \prod_{i=1}^{n} \rho_i(\theta) h_i(x_i) f_i(x_i) dx_1 \cdots dx_n\right) \frac{1}{m} \overline{F}_{n+1}(t-x)dx \\
&\qquad = m \left(\prod_{i=1}^{n} \rho_i(\theta)\right) \int_0^t f_{n+1}^I(x_0) \widetilde{f}_i(x_i) dx_0 \cdots dx_n \\
&\qquad = m \left(\prod_{i=1}^{n} \rho_i(\theta)\right) \left(\widetilde{f}_1 * \cdots * \widetilde{f}_n * f_{n+1}^I\right)(t)
\end{aligned}$$

Putting all of the above together we obtain the result

$$M(\theta, t) := E\left[e^{-\theta \sum_{n=1}^{N(t)} Y_n}\right] = \sum_{n=0}^{\infty} m \left(\prod_{i=1}^{n} \rho_i(\theta)\right) \widetilde{f}_1 * \cdots * \widetilde{f}_n * f_{n+1}^I(t).$$

If

$$\widehat{\widetilde{f}}_i(s) = \int_0^\infty e^{-st} \widetilde{f}_i(t) dt$$

then, the Laplace transform of $M(\theta, t)$ is

$$\begin{aligned}
\widehat{M}(\theta, s) &= \int_0^\infty e^{-st} M(\theta, t) dt = m \sum_{n=0}^{\infty} \int_0^\infty e^{-st} \left(\prod_{i=1}^{n} \rho_i(\theta)\right) \widetilde{f}_1 * \cdots * \widetilde{f}_n * f_{n+1}^I(t) dt \\
&= m \sum_{n=0}^{\infty} \left(\prod_{i=1}^{n} \rho_i(\theta) \widehat{\widetilde{f}}_i(s)\right) \widehat{f}_{n+1}^I(s)
\end{aligned}$$

But,

$$\begin{aligned}
\widehat{f^I}_{n+1}(s) &= \frac{1}{m} \int_0^\infty e^{-st} \overline{F}_{n+1}(t) dt = -\frac{1}{sm} \left[e^{-st} \overline{F}_{n+1}(t)\Big|_0^\infty + \int_0^\infty e^{-st} dF_{n+1}(t)\right] \\
&= \frac{1}{sm} \left(1 - \widehat{F}_{n+1}(s)\right)
\end{aligned}$$

and

$$\begin{aligned}\widehat{\widetilde{f}}(s) &= \int_0^\infty e^{-st}\widetilde{f}_i(t)dt = \int_0^\infty e^{-st}\frac{1}{\rho_i(\theta)}E[e^{-\theta Y_i}|X_i=t]f(t)dt \\ &= \frac{1}{\rho_i(\theta)}\int_0^\infty E[e^{-\theta Y - sX}|X_i=t]f(t)dt = \frac{1}{\rho_i(\theta)}E[e^{-\theta Y_i - sX_i}] = \frac{E[e^{-\theta Y_i - sX_i}]}{E[e^{-\theta Y_i}]}.\end{aligned}$$

Taking into account the above Laplace transforms, the expression in (2.31) becomes

$$\begin{aligned}\widehat{M}(\theta,s) &= \sum_{n=0}^{\infty}\prod_{i=1}^{n}\rho_i(\theta)\frac{1}{s}\left(1-\widehat{F}_{n+1}(s)\right)\prod_{i=1}^{n}\frac{E[e^{-\theta Y_i - sX_i}]}{\rho_i(\theta)} \\ &= \sum_{n=0}^{\infty}\frac{1}{s}\left(1-\widehat{F}_{n+1}(s)\right)\prod_{i=1}^{n}E[e^{-\theta Y_i - sX_i}].\end{aligned}$$

□

We can use (2.27) in order to find the Laplace transform of $M(\theta,t)$ in the case where $E[e^{-\theta Y_i - sX_i}] = \frac{\mu_i}{\mu_i+\theta}\frac{\lambda_i}{\lambda_i+s}$.

2.7.1 Correlated Exponential Sources

In this section we depart from the model of independent cycles. A number of authors have used various approaches in considering risk models with correlated claim processes. See for instance [1] and [111]. We start with the bivariate exponential distribution (see Downton [42]) with Laplace transform given by

$$\psi(s,\theta) = \frac{\lambda\mu}{(\lambda+s)(\mu+\theta)-\rho s\theta}. \tag{2.31}$$

This correspond to a distribution, absolutely continuous with respect to the Lebesgue measure in the positive quadrant, with exponential marginals having rates λ and μ and correlation between the two variables given by $\rho \in [0,1)$. Assuming (2.31) gives the joint Laplace transform of the length of each segment, X_i, and the reward from this segment, Y_i, and using (2.28) we obtain

$$\widehat{M}(\theta,s) = \int_0^\infty e^{-st}M(\theta,t)dt = \frac{1}{s}\,\frac{1-\frac{\lambda}{\lambda+s}}{1-\frac{\lambda\mu}{(\lambda+s)(\mu+\theta)-\rho s\theta}}$$

which simplifies to give

$$\widehat{M}(\theta,s) = \frac{1}{\lambda+s} + \frac{\mu\lambda}{(\mu+s)\left[\theta(\mu+s(1-\rho))+\mu s\right]}. \tag{2.32}$$

Analyzing the above expression in partial fractions we obtain

$$\widehat{M}(\theta,s) = \frac{1}{\lambda+s}\left(\frac{\rho\theta}{\rho\theta-\mu}\right) + \mu\frac{\mu+\theta(1-\rho)}{\mu-\rho\theta}\frac{1}{\theta\lambda+s\left[\mu+\theta(1-\rho)\right]}.$$

Inverting the above with respect to s, and replacing θ by $-\theta$ we obtain the following expression for the moment generating function $E[e^{\theta\sum_{i=1}^{N(t)}Y_i}]$ the partial inverse

$$M(-\theta,t) = E[e^{\theta\sum_{i=1}^{N(t)}Y_i}] = e^{-\lambda t}\frac{\rho\theta}{\rho\theta+\mu} + \frac{\mu}{\mu+\rho\theta}e^{-\frac{\lambda\theta}{\mu+\theta(1-\rho)}t}. \tag{2.33}$$

This can be rewritten as

$$M(\theta,t) \;=\; e^{-\lambda t} - e^{-\lambda t}\frac{\frac{\mu}{\rho}}{\frac{\mu}{\rho}-\theta} + \frac{\frac{\mu}{\rho}}{\frac{\mu}{\rho}-\theta}e^{-\frac{\lambda}{1-\rho}t\left(1-\frac{\frac{\mu}{1-\rho}}{\frac{\mu}{1-\rho}+\theta}\right)}. \tag{2.34}$$

To invert (2.34) note that $\frac{\frac{\mu}{\rho}}{\frac{\mu}{\rho}-\theta}$ is the Laplace transform of a probability distribution on $(-\infty,0]$ with density $\frac{\mu}{\rho}e^{-\frac{\mu}{\rho}y}1(y\le 0)$, while

$$e^{-\frac{\lambda}{1-\rho}t\left(1-\frac{\frac{\mu}{1-\rho}}{\frac{\mu}{1-\rho}+\theta}\right)}$$

is the Laplace transform of a compound Poisson random variable $\sum_{i=1}^{K}\xi_i$ with K a Poisson random variable with mean $\frac{\lambda}{1-\rho}t$ and the i.i.d. random variables ξ_i are exponential with rate $\frac{\mu}{1-\rho}$. The mean and variance of $\sum_{n=1}^{N(t)}Y_n$ can be obtained from (2.32) by differentiating w.r.t. θ and inverting the resulting Laplace transform w.r.t. s.

$$-\frac{d}{d\theta}\,\widehat{M}(\theta,s)\Big|_{\theta=0} = \frac{\lambda}{\lambda+s}\left(\frac{\lambda}{\mu^2}\frac{1}{s^2}+\frac{1-\rho}{\mu^2}\frac{1}{s}\right)$$

hence, inverting the Laplace transform,

$$E\sum_{n=1}^{N(t)}Y_n = \frac{\lambda}{\mu}t - \frac{\rho}{\mu}\left(1-e^{-\lambda t}\right).$$

Also,

$$\frac{d^2}{d\theta^2}\,\widehat{M}(\theta,s)\Big|_{\theta=0} = 2\frac{\lambda}{\lambda+s}\frac{\mu\left(\lambda+s(1-\rho)\right)^2}{(\mu s)^3} = 2\frac{\lambda}{\lambda+s}\left(\frac{\lambda^2}{\mu^2}\frac{1}{s^3}+\frac{2\lambda(1-\rho)}{\mu^2}\frac{1}{s^2}+\frac{(1-\rho)^2}{\mu^2}\frac{1}{s}\right)$$

and, therefore,

$$E\left(\sum_{n=1}^{N(t)}Y_n\right)^2 = \frac{\lambda^2}{\mu^2}t^2 + \frac{2\lambda(1-2\rho)}{\mu^2}t + \frac{2\rho^2}{\mu^2}\left(1-e^{-\lambda t}\right).$$

Hence, from the above we have

$$\mathsf{Var}\left(\sum_{n=1}^{N(t)}Y_n\right) = \frac{2\lambda t}{\mu^2}(1-\rho-\rho e^{-\lambda t}) + \frac{\rho^2}{\mu^2}\left(1-e^{-2\lambda t}\right).$$

We can use (2.34) in conjunction with Theorem 1.16 in order to obtain the following Large Deviations result.

Theorem 2.4. *In the context of the above model, assuming that*

$$\frac{\lambda}{\mu} < c, \tag{2.35}$$

it holds that

$$\lim_{x\to\infty} x^{-1}\log P\left(\sup_{t>0}\sum_{i=1}^{N(t)}Y_i - ct > x\right) \;=\; -\frac{\mu c-\lambda}{c(1-\rho)}. \tag{2.36}$$

Proof.

$$\psi(\theta) := \lim_{t\to\infty}\frac{1}{t}\log Ee^{\theta\left(\sum_{n=1}^{N(t)}Y_n-ct\right)} = \lim_{t\to\infty}\frac{1}{t}\log\left(e^{-\lambda t}\frac{\rho\theta}{\rho\theta+\mu}+\frac{\mu}{\mu+\rho\theta}e^{\frac{\lambda\theta}{\mu-\theta(1-\rho)}t}\right) - \theta c$$
$$= \frac{\lambda\theta}{\mu-\theta(1-\rho)} - c\theta. \tag{2.37}$$

Referring to Theorem 1.16 we will check whether the conditions of the theorem are satisfied. Note that the above limit holds in the interval $-c_0 < \theta < \frac{\mu}{1-\rho}$ (for some $c_0 > 0$) and the limiting function ψ given in (2.37) is clearly differentiable there. The equation $\psi(\theta) = 0$ has the unique solution

$$\theta^* = \frac{\lambda-\mu c}{c(1-\rho)}. \tag{2.38}$$

Also

$$\psi'(\theta^*) = \frac{\lambda}{\mu-\theta^*(1-\rho)}\left[1+\frac{\theta^*}{\mu-\theta^*(1-\rho)}\right]-c = c\left[1+\frac{\mu c-\lambda}{1-\rho}\right]-c > 0,$$

the last inequality following from (2.35). Conditions GW3 and GW4 are also satisfied. □

2.7.2 A Generalization of the Correlated Model

We propose the following model as a generalization of the bivariate exponential model discussed above. Suppose that X is a strictly positive random variable with density f and corresponding Laplace transform $\varphi(s) = Ee^{-sX}$. Let ζ_0 be a positive random variable with Laplace transform $\beta_0(s)$, and $\zeta_i, i = 1, 2, \ldots$, i.i.d. random variables with Laplace transform $\beta_1(s)$. Finally, let $\{K(t); t \geq 0\}$ denote the counting process of a Poisson process with unit rate. All the above random elements are independent. Then, if $r > 0$ is a positive parameter, let

$$Y = \zeta_0 + \sum_{i=1}^{K(rX)}\zeta_i \tag{2.39}$$

and from the above it follows that

$$E\left[e^{-\theta Y}|X=x\right] = \beta_0(\theta)e^{-rx(1-\beta_1(\theta))}. \tag{2.40}$$

Hence the joint Laplace transform of the random vector (X, Y) is

$$\psi(s,\theta) := E[e^{-sX-\theta Y}] = E\left[e^{-sX}E\left[e^{-\theta Y}|X\right]\right] = E\left[e^{-sX}\beta_0(\theta)e^{-rX(1-\beta_1(\theta))}\right]$$
$$= \beta_0(\theta)\varphi(s+r(1-\beta_1(\theta))). \tag{2.41}$$

To understand the sense in which the above model generalizes (2.31) rewrite the latter as

$$\psi(s,\theta) = \frac{\mu}{\mu+\theta}\frac{\lambda}{\lambda+s\left(1-\rho\frac{\theta}{\mu+\theta}\right)}$$

and invert with respect to s to obtain the function

$$\frac{\mu}{\mu+\theta}\frac{\lambda}{1-\rho\frac{\theta}{\mu+\theta}}e^{-x\lambda\frac{\mu+\theta}{\mu+\theta(1-\rho)}} = \frac{\frac{\mu}{1-\rho}}{\frac{\mu}{1-\rho}+\theta}\lambda e^{-x\lambda\frac{\mu+\theta}{\mu+\theta(1-\rho)}}.$$

Dividing this with the density of X, namely $\lambda e^{-\lambda x}$, we obtain

$$E\left[e^{-\theta Y}|X=x\right] = \frac{\frac{\mu}{1-\rho}}{\frac{\mu}{1-\rho}+\theta} e^{-x\lambda\left(\frac{\mu+\theta}{\mu+\theta(1-\rho)}-1\right)} = \frac{\frac{\mu}{1-\rho}}{\frac{\mu}{1-\rho}+\theta} e^{-x\lambda\frac{\rho}{1-\rho}\left(1-\frac{\frac{\mu}{1-\rho}}{\frac{\mu}{1-\rho}+\theta}\right)}.$$

This conditional Laplace transform is clearly of the form (2.40) with $\beta_0(\theta) = \beta_1(\theta) = \frac{\frac{\mu}{1-\rho}}{\frac{\mu}{1-\rho}+\theta}$, $\varphi(s) = \frac{\lambda}{\lambda+s}$, and $r = \lambda\frac{\rho}{1-\rho}$.

Proposition 2.5. *In the context of the correlated model with $\varphi(s) := \frac{\lambda}{\lambda+s}$ and therefore with*

$$\psi(s,\theta) = \frac{\lambda\beta_0(\theta)}{\lambda s + r(1-\beta_1(\theta))}$$

the joint statistics of $\left(\sum_{i=1}^{N(t)} X_i, \sum_{i=1}^{N(t)} Y_i\right)$ are given by

$$\begin{aligned}\Phi(\theta,u) &= Ee^{-u\sum_{i=1}^{N(t)} X_i - \theta\sum_{i=1}^{N(t)} Y_i} \\ &= e^{-\lambda t}\,\frac{u + r[1-\beta_1(\theta)] - \lambda\beta_0(\theta)e^{-t(u+r[1-\beta_1(\theta)]-\lambda\beta_0(\theta))}}{u + r[1-\beta_1(\theta)] - \lambda\beta_0(\theta)}\,. \end{aligned} \tag{2.42}$$

Proof. We first begin by evaluating

$$\begin{aligned} &E\left[e^{-u\sum_{i=1}^n X_i - \theta\sum_{i=1}^n Y_i} 1(X_1+\cdots+X_n \in dx)\right] \\ &\quad = dx\int_{S_n}\left(\prod_{i=1}^n E\left[e^{-\theta Y_i}\middle| X_i = x_i\right] e^{-ux_i}\lambda e^{-\lambda x_i}\right) dx_1\cdots dx_n \\ &\quad = dx\int_{S_n}\left(\prod_{i=1}^n \beta_0(\theta)e^{-(r[1-\beta_1(\theta)]+u)x_i}\lambda e^{-\lambda x_i}\right) dx_1\cdots dx_n \\ &\quad = dx\,(\beta_0(\theta))^n\int_{S_n}\lambda^n e^{-B\sum_{i=1}^n x_i}dx_1\cdots dx_n \;=\; (\lambda\beta_0(\theta))^n\,\frac{x^{n-1}}{(n-1)!}e^{-Bx}dx \end{aligned}$$

with $B := r[1-\beta_1(\theta)] + \lambda + u$. So, according to (2.29) we obtain that, for $n \geq 1$,

$$\begin{aligned} &E\left[e^{-u\sum_{i=1}^{N(t)} X_i - \theta\sum_{i=1}^{N(t)} Y_i}; N(t) = n\right] \\ &\quad = \int_0^t e^{-\lambda(t-x)}\,E\left[e^{-u\sum_{i=1}^n X_i - \theta\sum_{i=1}^n Y_i}1(X_1+\cdots+X_n\in dx)\right] \\ &\quad = \int_0^t (\lambda\beta_0(\theta))^n\,\frac{x^{n-1}}{(n-1)!}e^{-Bx}e^{-\lambda(t-x)}dx. \end{aligned} \tag{2.43}$$

Adding over all n,

$$\begin{aligned}\Phi(u,\theta) &= e^{-\lambda t} + \sum_{n=1}^\infty\int_0^t (\lambda\beta_0(\theta))^n\,\frac{x^{n-1}}{(n-1)!}e^{-Bx}e^{-\lambda(t-x)}dx \\ &= e^{-\lambda t} + e^{-\lambda t}\lambda\beta_0(\theta)\int_0^t e^{-(B-\lambda)x+\lambda\beta_0(\theta)x}dx \\ &= e^{-\lambda t} + e^{-\lambda t}\lambda\beta_0(\theta)\frac{1-e^{-t(r[1-\beta_1(\theta)]-\lambda\beta_0(\theta)+u)}}{r[1-\beta_1(\theta)]-\lambda\beta_0(\theta)+u}. \end{aligned}$$

From this (2.42) follows. □

Remark: When $\theta = 0$ (2.42) reduces

$$\Phi(u,0) = E\left[e^{-u\sum_{i=1}^{N(t)} X_i}\right] = \frac{ue^{-\lambda t} - \lambda e^{-ut}}{u-\lambda}. \tag{2.44}$$

An example of the correlated model with $\beta_0(\theta) = \beta_1(\theta) = \frac{\mu}{\mu+\theta}$.

We consider the model (2.40) with $\beta_0(\theta) = \beta_1(\theta) = \frac{\mu}{\mu+\theta}$ and $r = \lambda\rho$. Taking also $u = 0$, (2.42) gives

$$\Phi(0,\theta) \;=\; e^{-\lambda t}\,\frac{\rho\theta - \mu e^{-\frac{\lambda}{\mu+\theta}(\rho\theta-\mu)t}}{\rho\theta-\mu}.$$

or equivalently

$$\Phi(0,\theta) \;=\; e^{-\lambda t} - \frac{\mu}{\mu-\rho\theta}e^{-\lambda t} + \frac{\mu}{\mu-\rho\theta}\,e^{-\lambda t(1+\rho)\left(1-\frac{\mu}{\mu+\theta}\right)} \tag{2.45}$$

Inverting (2.45) with respect to θ we obtain the density function of $\sum_{i=1}^{N(t)} Y_i$,

$$\varphi(x) \;=\; e^{-\lambda t}\delta_0(x) - e^{-\lambda t}\frac{\mu}{\rho}e^{\frac{\mu}{\rho}x}\mathbf{1}(x<0) + \int_0^\infty f_{cp}(x+y)\frac{\mu}{\rho}e^{-\frac{\mu}{\rho}y}dy \tag{2.46}$$

where f_{cp} is the density of the compound Poisson distribution with Laplace transform

$$e^{-\lambda t(1+\rho)\left(1-\frac{\mu}{\mu+\theta}\right)}.$$

Thus the density is equal to

$$f_{cp}(x) \;=\; e^{-\lambda t(1+\rho)}\left(\delta_0(x) + \mu\lambda t(1+\rho)\sum_{n=1}^{\infty}\frac{(\lambda t(1+\rho)\mu x)^{n-1}}{n!(n-1)!}e^{-\mu x}\right)$$

In the special case in which $\rho = 0$ (2.45) gives

$$\Phi(0,\theta) \;=\; e^{-\lambda t\left(1-\frac{\mu}{\mu+\theta}\right)}$$

which corresponds to the classical compound Poisson process.

Using (2.44) and (2.45) we can find the mean and variance of $\sum_{i=1}^{N(t)} X_i$ and $\sum_{i=1}^{N(t)} Y_i$ by differentiating (2.44) and (2.45) with respect to u and θ respectively.

$$E\left[\sum_{i=1}^{N(t)} X_i\right] = \frac{\lambda t + e^{-\lambda t} - 1}{\lambda}, \quad \mathsf{Var}\left[\sum_{i=1}^{N(t)} X_i\right] = \frac{1 - e^{-\lambda t} - 2\lambda t e^{-\lambda t}}{\lambda^2},$$

and

$$\begin{aligned} E\left[\sum_{i=1}^{N(t)} Y_i\right] &= \frac{\lambda}{\mu}(1+\rho)t - \frac{\rho}{\mu}(1-e^{-\lambda t}), \\ \mathsf{Var}\left[\sum_{i=1}^{N(t)} Y_i\right] &= 2\frac{\lambda}{\mu^2}(1+\rho)t + \frac{\rho^2}{\mu^2} - 2\frac{\lambda}{\mu^2}(1+\rho)\rho t e^{-\lambda t} - \frac{\rho^2}{\mu^2}e^{-2\lambda t}. \end{aligned}$$

In the insurance risk context we are interested in

$$\begin{aligned}\Psi_t(\theta) &= \Phi_t(c\theta,-\theta) = E\left[e^{\theta\left(\sum_{i=1}^{N(t)}Y_i - c\sum_{i=1}^{N(t)}X_i\right)}\right] \\ &= e^{-\lambda t} + \lambda\mu e^{-\lambda t}\frac{1-e^{-\left[(c\theta+\rho\lambda)-\frac{\lambda\mu}{\mu-\theta}(1+\rho)\right]t}}{(c\theta+\rho\lambda)(\mu-\theta)-\lambda\mu(1+\rho)} \end{aligned} \tag{2.47}$$

In order to apply the results of Theorem 1.16 we consider

$$\psi(\theta) := \lim_{t\to\infty}\frac{1}{t}\log\Psi_t(\theta) = -\lambda-\theta c-\lambda\rho+\frac{\lambda\mu(1+\rho)}{\mu-\theta} \qquad (\theta<\mu).$$

Then

$$\theta^* = \mu - \frac{\lambda(1+\rho)}{c} \tag{2.48}$$

is the unique positive solution of $\phi(\theta)=0$. Further,

$$\psi'(\theta^*) = c\left(\frac{\mu c}{\lambda(1+\rho)}-1\right)$$

The quantity inside the parenthesis is positive provided that

$$c > \frac{\lambda(1+\rho)}{\mu}, \tag{2.49}$$

a condition that states that the premium rate c must exceed the claim rate. Then

$$\lim_{x\to\infty} x^{-1}\log P\left(\sup_{t>0}\sum_{i=1}^{N(t)}Y_i - ct > x\right) = -\mu+\frac{\lambda(1+\rho)}{c}. \tag{2.50}$$

2.7.3 A Correlated Model With Erlang Intervals

In the framework of section (2.7.2) suppose that X is an Erlang(k) distributed random variable i.e. $Ee^{-sX} = \left(\frac{\lambda}{\lambda+s}\right)^k$ where $k\in\mathbb{N}$ and Y is given by 2.39 where, again, ζ_0 is a random variable with positive values and Laplace transform $\beta_0(s)$, and ζ_i, $i=0,1,2,\ldots$ are i.i.d. with Laplace transform $\beta_1(s)$. Alos, K is a unit rate Poisson process, also independent from the other random variables, and $r>0$ a positive constant. $E[e^{-\theta Y}|X]$ is again given by (2.39) so repeating the arguments leading to (2.43) under the assumption that the random variables $\{X_i\}$ are Erlang(k) we have

$$\begin{aligned}\Phi_t(u,\theta) &:= E\left[e^{-u\sum_{i=1}^{N(t)}X_i-\theta\sum_{i=1}^{N(t)}Y_i}\right] \\ &= \overline{F}(t)+\sum_{n=1}^{\infty}\beta_0(\theta)^n\int_0^t e^{-x(r[1-\beta_1(\theta)]+u)}f^{*n}(x)\overline{F}(t-x)dx. \end{aligned} \tag{2.51}$$

In the above expression $\overline{F}(x)$ is the probability that an Erlang(k) random variable exceeds x and therefore $\overline{F}(t-x)=\sum_{j=0}^{k-1}\frac{[\lambda x)]^j}{j!}e^{-\lambda(t-x)}$. Further, the density of the sum of n independent such random variables is $P(X_1+\cdots+X_n\in dx)=f^{*n}(x)dx=\lambda\frac{(\lambda x)^{nk-1}}{(nk-1)!}e^{-\lambda x}dx$. If we set

$$z := (\beta_0(\theta))^{\frac{1}{k}} \quad \text{and} \quad A := u+r[1-\beta_1(\theta)],$$

(2.51) becomes

$$\begin{aligned}\Phi_t(u,\theta) &= \sum_{j=0}^{k-1}\frac{(\lambda t)^j}{j!}e^{-\lambda t}+z\int_0^t\sum_{n=1}^{\infty}z^{nk-1}e^{-Ax}\lambda\frac{(\lambda x)^{nk-1}}{(nk-1)!}e^{-\lambda x}\overline{F}(t-x)dx\\ &= \sum_{j=0}^{r-1}\frac{(\lambda t)^j}{j!}e^{-\lambda t}+\lambda z\int_0^t e^{-Ax}\sum_{n=1}^{\infty}[\lambda xz]^{nk-1}\frac{e^{-\lambda x}}{(nk-1)!}\sum_{j=0}^{k-1}\frac{[\lambda(t-x)]^j}{j!}e^{-\lambda(t-x)}dx\\ &= \sum_{j=0}^{k-1}\frac{(\lambda t)^j}{j!}e^{-\lambda t}+\lambda z e^{-\lambda t}\int_0^t e^{-Ax}\sum_{n=1}^{\infty}\frac{(\lambda xz)^{nk-1}}{(nk-1)!}\sum_{j=0}^{k-1}\frac{[\lambda(t-x)]^j}{j!}dx.\end{aligned}$$

We now make use of the identity

$$\sum_{n=1}^{\infty}\frac{z^{nk-1}}{(nk-1)!} = \frac{1}{k}\left(e^z+we^{wz}+w^2e^{w^2z}+\cdots+w^{k-1}e^{w^{k-1}z}\right) =: \sum_{i=0}^{k-1}c_ie^{w_iz}$$

where $1, w, w^2, \ldots, w^{k-1}$ (or $w_0, w_1, w_2, \ldots w_{k-1}$) are the k roots of unity $\left(w_i^k=1, w_{k-1}=w^{k-1}\right)$. For the proof of the above we can use the definition of e^{w_ix}, multiply it by w^{i+1} for $i = 0, 1, \ldots, k-1$ and add the sums. Thus we have

$$\begin{aligned}\Phi_t(u,\theta) &= \sum_{j=0}^{k-1}\frac{(\lambda t)^j}{j!}e^{-\lambda t}+\lambda z e^{-\lambda t}\int_0^t e^{-Ax}\sum_{i=0}^{k-1}c_ie^{w_i\lambda xz}\sum_{j=0}^{k-1}\frac{[\lambda(t-x)]^j}{j!}dx\\ &= \sum_{j=0}^{k-1}\frac{(\lambda t)^j}{j!}e^{-\lambda t}+\lambda z e^{-\lambda t}\sum_{i=0}^{k-1}c_i\int_0^t e^{[w_i\lambda z-A]x}\sum_{j=0}^{k-1}\frac{[\lambda(t-x)]^j}{j!}dx.\end{aligned}$$

Turning to asymptotics we note that

$$\Phi_t(c\theta,-\theta) = q(t)e^{-\lambda t}+\sum_{i=0}^{k-1}p_i(t)e^{t\,[w_i\lambda z-A-\lambda]}$$

where $q(t)$ and $p_i(t)$, $i = 0, \ldots, k-1$ are polynomials in t. The largest of the exponents corresponds to the root $w_0 = 1$. Hence

$$\psi(\theta) := \lim_{t\to\infty}\frac{1}{t}\log\Phi_t(c\theta,-\theta) = w_0\lambda z-A-\lambda = \lambda\left(M_0(\theta)\right)^{\frac{1}{k}}+\lambda\rho\left(M_1(\theta)-1\right)-c\theta-\lambda.$$

In the above we have set $M_i(\theta) := \beta_i(-\theta)$, $i = 0, 1$ and $r = \rho\lambda$. Thus we may write

$$\psi(\theta) = \lambda(1+\rho)\left[\frac{1}{1+\rho}\left(M_0(\theta)\right)^{\frac{1}{k}}+\frac{\rho}{1+\rho}\left(M_1(\theta)-1\right)\right]-c\theta-\lambda. \tag{2.52}$$

The equation $\psi(\theta) = 0$ is then

$$\frac{1}{1+\rho}\left(M_0(\theta)\right)^{\frac{1}{k}}+\frac{\rho}{1+\rho}M_1(\theta) = \frac{c\theta}{\lambda(1+\rho)}+1. \tag{2.53}$$

Since $M_0(0) = M_1(0) = 1$ it is clear that $\theta = 0$ is a root of (2.53). $M_1(\theta)$ is a convex function of θ being a moment generating function. We now show that $M_0(\theta)^{1/k}$ is also convex. Indeed

$$\frac{d^2}{d\theta^2}M_0^{1/k} = \frac{1}{r}\left(\frac{1}{k}-1\right)M_0^{\frac{1}{k}-2}(M_0')^2+\frac{1}{k}M_0^{\frac{1}{k}-1}M_0'' = \frac{1}{k}M_0^{1/k}\left(\frac{1}{k}\left(\frac{M_0'}{M_0}\right)^2+\frac{M_0''}{M_0}-\left(\frac{M_0'}{M_0}\right)^2\right) > 0$$

as a result of interpreting $\frac{M_0''}{M_0} - \left(\frac{M_0'}{M_0}\right)^2$ as a variance of an Escher–transformed random variable. Therefore (2.53) besides 0 also has a positive root θ^* provided that the following condition holds

$$\frac{1}{1+\rho}\,\frac{1}{k}\,(M_0(0))^{\frac{1}{k}-1}\,M_0'(0) + \frac{\rho}{1+\rho}\,M_1'(0) < \frac{c}{\lambda(1+\rho)}.$$

or, setting $m_i := M_i'(0)$, $i = 1, 2$,

$$\frac{m_0}{k} + \rho\, m_1 < \frac{c}{\lambda} \tag{2.54}$$

We also need to show that $\Psi'(\theta^*) > 0$. Note that the function

$$\widetilde{M}(\theta) := \frac{1}{1+\rho}\,(M_0(\theta))^{\frac{1}{k}} + \frac{\rho}{1+\rho}\,M_1(\theta)$$

is convex. With this definition, from (2.52), we need to show that

$$\psi'(\theta^*) = \lambda(1+\rho)\widetilde{M}'(\theta^*) - c > 0. \tag{2.55}$$

On the other hand, from (2.53),

$$\widetilde{M}(\theta^*) \;=\; 1 + \frac{c\theta^*}{\lambda(1+\rho)}. \tag{2.56}$$

Then (2.55) follows from the convexitiy of $\widetilde{M}(\theta)$ from which we have

$$\widetilde{M}'(\theta^*) > \frac{\widetilde{M}(\theta^*) - 1}{\theta^* - 0} = \frac{c}{\lambda(1+\rho)}.$$

Then

$$\lim_{x\to\infty} x^{-1} \log P\left(\sup_{t>0} \sum_{i=1}^{N(t)} Y_i - ct > x\right) \;=\; -\theta^* \tag{2.57}$$

where θ^* is the unique positive solution of (2.53).

2.8 *Time Inhomogeneous Independent Cycles:* $\psi(\theta, s) = \frac{\mu_i}{\mu_i+\theta}\frac{\lambda_i}{\lambda_i+s}$, *with* $\lambda_i = \lambda$

We start our analysis with $\lambda_i = \lambda$. In this case,

$$\widehat{M}(\theta, s) = \sum_{n=0}^{\infty} \frac{1}{s}\left(1 - \widehat{F}_{n+1}(s)\right)\left(\frac{\lambda}{\lambda+s}\right)^n \prod_{i=1}^{n} \frac{\mu_i}{\mu_i+\theta} = \lambda^{-1} \sum_{n=0}^{\infty} \left(\frac{\lambda}{\lambda+s}\right)^{n+1} \prod_{i=1}^{n} \frac{\mu_i}{\mu_i+\theta}. \tag{2.58}$$

Using partial fractions we obtain $\prod_{i=1}^{n} \frac{\mu_i}{\mu_i+\theta} = \sum_{i=1}^{n} a_{i,n} \frac{\mu_i}{\mu_i+\theta}$ where

$$a_{i,n} = \prod_{\substack{j=1 \\ j\neq i}}^{n} \frac{\mu_j}{\mu_j - \mu_i}. \tag{2.59}$$

Inverting (2.58) with respect to s, we obtain

$$\begin{aligned} M(\theta,t) &= \sum_{n=0}^{\infty}\frac{(\lambda t)^n}{n!}e^{-\lambda t}\sum_{i=1}^{n}\frac{\mu_i}{\mu_i+\theta}a_{i,n} \\ &= e^{-\lambda t}+\sum_{i=1}^{\infty}\frac{\mu_i}{\mu_i+\theta}\left\{\sum_{n=i}^{\infty}\frac{(\lambda t)^n}{n!}a_{i,n}\right\}e^{-\lambda t} \end{aligned} \tag{2.60}$$

In the special case where $\mu_i = \mu i$ taking into account (2.59) we obtain

$$a_{i,n} = \frac{n!}{(n-i)!i!}(-1)^{i-1} = \binom{n}{i}(-1)^{i-1}. \tag{2.61}$$

Using (2.60) we obtain

$$\begin{aligned} M(\theta,t) &= e^{-\lambda t}+\sum_{i=1}^{\infty}\frac{\mu i}{\mu i+\theta}\left\{\sum_{n=i}^{\infty}\frac{(\lambda t)^n}{n!}\binom{n}{i}(-1)^{i-1}\right\}e^{-\lambda t} \\ &= e^{-\lambda t}+\sum_{i=1}^{\infty}\frac{\mu i}{\mu i+\theta}\frac{(-1)^{i-1}}{i!}(\lambda t)^i. \end{aligned} \tag{2.62}$$

The above Laplace transform can be inverted to give the density

$$e^{-\lambda t}\delta(y)+\mu\lambda t e^{-\mu y}e^{-\lambda t e^{-\mu y}}, \quad y\geq 0,$$

and distribution function

$$\begin{cases} e^{-\lambda t e^{-\mu y}} & \text{if} \quad y\geq 0 \\ 0 & \text{if} \quad y<0 \end{cases}$$

which is the distribution function of the positive part of a Gumbel random variable. This distribution has an atom of size $e^{-\lambda t}$ at 0.

The expected value $m(t) := E[\sum_{i=1}^{N_t} Y_i]$ can be obtained from (2.62) as

$$m(t) = -\frac{\partial}{\partial\theta}M(\theta,t)\Big|_{\theta=0} = \sum_{i=1}^{\infty}\frac{1}{\mu i}\frac{(-1)^{i-1}}{i!}(\lambda t)^i = -\frac{1}{\mu}\sum_{i=1}^{\infty}\frac{(-\lambda t)^i}{i\cdot i!}.$$

The derivative of $m(t)$ (which can be obtained by differentiating the above power series w.r.t. t as we may) is

$$m'(t) = -\frac{1}{\mu}\sum_{i=1}^{\infty}\frac{(-\lambda)^i t^{i-1}}{i!} = \frac{1-e^{-\lambda t}}{\mu t}.$$

Hence

$$m(t) = \int_0^t \frac{1-e^{-\lambda u}}{\mu u}du = \frac{1}{\mu}\int_0^{\lambda t}\frac{1-e^{-u}}{u}du,$$

and thus, as expected, it increases essentially logarithmically as a function of t.

2.9 Time Inhomogeneous Cycles: $\psi(\theta,s)=\frac{\mu_i}{\mu_i+\theta}\frac{\lambda_i}{\lambda_i+s}$, with $\mu_i=\mu$

We continue the above analysis with the case $\mu_i=\mu$. Now, (2.27) becomes

$$\widehat{M}(\theta,s)=\sum_{n=0}^{\infty}\left(\frac{\mu}{\mu+\theta}\right)^n\frac{1}{\lambda_{n+1}+s}\prod_{i=1}^{n}\frac{\lambda_i}{\lambda_i+s} \tag{2.63}$$

Inverting (2.63) with respect to s, we obtain

$$M(\theta,t)=\sum_{n=0}^{\infty}\left(\frac{\mu}{\mu+\theta}\right)^n\frac{1}{\lambda_{n+1}}\sum_{i=1}^{n+1}a_{i,n+1}\lambda_i e^{-\lambda_i t} \tag{2.64}$$

In the special case where $\lambda_i=\lambda i$ taking into account (2.59) (replacing μ_i by λ_i) we obtain (using (2.61)) that

$$\sum_{i=1}^{n+1}a_{i,n+1}\lambda_i e^{-\lambda_i t}=\lambda\sum_{i=1}^{n+1}\binom{n+1}{i}(-1)^{i-1}i\left(e^{-\lambda t}\right)^i=\lambda e^{-\lambda t}(n+1)\left(1-e^{-\lambda t}\right)^n$$

and (2.64) becomes

$$\begin{aligned}M(\theta,t) &= \sum_{n=0}^{\infty}\left(\frac{\mu}{\mu+\theta}\right)^n\frac{1}{\lambda(n+1)}\lambda e^{-\lambda t}(n+1)\left(1-e^{-\lambda t}\right)^n\\ &= e^{-\lambda t}\sum_{n=0}^{\infty}\left[\left(\frac{\mu}{\mu+\theta}\right)\left(1-e^{-\lambda t}\right)\right]^n\\ &= e^{-\lambda t}\frac{1}{1-\frac{\mu}{\mu+\theta}\left(1-e^{-\lambda t}\right)}.\end{aligned} \tag{2.65}$$

After simple calculations (2.65) can also be written as

$$M(\theta,t)=e^{-\lambda t}+(1-e^{-\lambda t})\frac{\mu e^{-\lambda t}}{\theta+\mu e^{-\lambda t}}.$$

This is an exponential distribution with an atom at zero and distribution function

$$\begin{cases}1-\left(1-e^{-\lambda t}\right)e^{-y\mu e^{-\lambda t}} & \text{if} \quad y\geq 0\\ 0 & \text{if} \quad y<0\end{cases}$$

As a second special case we can use $\lambda_i=\lambda i+b$. In this case,

$$a_{i,n}=\prod_{\substack{j=1\\ j\neq i}}^{n}\frac{\lambda j+b}{\lambda(j-i)}=\prod_{\substack{j=1\\ j\neq i}}^{n}\frac{j+\beta}{j-i},\qquad \text{where } \beta=\frac{b}{\lambda}.$$

So,

$$\begin{aligned}a_{i,n} &= \prod_{\substack{j=1\\ j\neq i}}^{n}(j+\beta)\frac{(-1)^{i-1}}{(n-i)!(i-1)!}\\ &= \frac{(1+\beta)(2+\beta)\cdots(n+\beta)}{n!}\frac{i}{i+\beta}\binom{n}{i}(-1)^{i-1}\end{aligned}$$

and

$$\begin{aligned}\sum_{i=1}^{n+1} a_{i,n+1}\lambda_i e^{-\lambda_i t} &= \sum_{i=1}^{n+1}\frac{(1+\beta)(2+\beta)\cdots(n+1+\beta)}{(n+1)!}\frac{i}{i+\beta}\binom{n+1}{i}(-1)^{i-1}\lambda(i+\beta)e^{-\lambda(i+\beta)t}\\ &= \frac{(1+\beta)(2+\beta)\cdots(n+1+\beta)}{(n+1)!}\sum_{i=1}^{n+1}\binom{n+1}{i}i(-1)^{i-1}\lambda e^{-\lambda(i+\beta)t}\\ &= \frac{(1+\beta)(2+\beta)\cdots(n+1+\beta)}{(n+1)!}\lambda e^{-\lambda(\beta+1)t}\sum_{i=1}^{n+1}\binom{n+1}{i}i(-e^{-\lambda t})^{i-1}\\ &= \frac{(1+\beta)(2+\beta)\cdots(n+1+\beta)}{(n+1)!}\lambda e^{-\lambda(\beta+1)t}(n+1)(1-e^{-\lambda t})^n.\end{aligned}$$

Using the last expression, (2.64) becomes

$$\begin{aligned}M(\theta,t) &= \sum_{n=0}^{\infty}\left(\frac{\mu}{\mu+\theta}\right)^n\frac{1}{\lambda(n+1)+b}\frac{(1+\beta)(2+\beta)\cdots(n+1+\beta)}{(n+1)!}\lambda e^{-\lambda(\beta+1)t}(n+1)(1-e^{-\lambda t})^n\\ &= \sum_{n=0}^{\infty}\left(\frac{\mu}{\mu+\theta}\right)^n\frac{1}{n+1+\beta}\frac{(1+\beta)(2+\beta)\cdots(n+1+\beta)}{(n+1)!}e^{-\lambda(\beta+1)t}(n+1)(1-e^{-\lambda t})^n\\ &= \sum_{n=0}^{\infty}\left(\frac{\mu}{\mu+\theta}\right)^n\frac{(1+\beta)(2+\beta)\cdots(n+\beta)}{n!}e^{-\lambda(\beta+1)t}(1-e^{-\lambda t})^n\\ &= e^{-\lambda(\beta+1)t}\sum_{n=0}^{\infty}\frac{(1+\beta)(2+\beta)\cdots(n+\beta)}{n!}\left[\frac{\mu}{\mu+\theta}(1-e^{-\lambda t})\right]^n\\ &= \left[\frac{e^{-\lambda t}}{1-\frac{\mu}{\mu+\theta}(1-e^{-\lambda t})}\right]^{\beta+1}\end{aligned}\tag{2.66}$$

that is, a negative binomial sum of exponentials.This enables us to find the density corresponding to $M(\theta,t)$ as

$$\sum_{n=1}^{\infty}\frac{(1+\beta)(2+\beta)\cdots(n+\beta)}{n!}e^{-\lambda(\beta+1)t}(1-e^{-\lambda t})^n\mu\frac{(\mu y)^{n-1}}{(n-1)!}e^{-\mu y}+e^{-\lambda(\beta+1)t}.$$

When, $\lambda_i=\lambda\,(i+m)$, $m\in\mathbb{N}$, then $a_{i,n}=\binom{n+m}{n}\frac{i}{i+m}\binom{n}{i}(-1)^{i-1}$ and

$$\sum_{i=1}^{n+1}a_{i,n+1}\lambda_i e^{-\lambda_i t}=\binom{n+1+m}{m}e^{-(m+1)\lambda t}(n+1)(1-e^{-\lambda t})^n$$

whence we obtain

$$M(\theta,t)=\left[\frac{e^{-\lambda t}}{1-\frac{\mu}{\mu+\theta}(1-e^{-\lambda t})}\right]^{m+1}$$

with density

$$\begin{aligned}&\sum_{n=0}^{\infty}(n+2)(n+3)\cdots(n+m+1)\frac{\left[(1-e^{-\lambda t})\mu y\right]^n}{n!}\mu e^{-\mu y}e^{-(m+1)\lambda t}+e^{-(m+1)\lambda t}\\ &\quad=\mu e^{-\mu y}e^{-(m+1)\lambda t}e^{(1-e^{-\lambda t})\mu y}\sum_{n=0}^{\infty}(n+2)(n+3)\cdots(n+m+1)e^{-(1-e^{-\lambda t})\mu y}\frac{\left[(1-e^{-\lambda t})\mu y\right]^n}{n!}\\ &\qquad+e^{-(m+1)\lambda t}.\end{aligned}\tag{2.67}$$

In order to evaluate the sum $\sum_{n=0}^{\infty}(n+2)(n+3)\cdots(n+m+1)e^{-a}\frac{a^n}{n!}$ where $a := (1-e^{-\lambda t})\mu y$ we use the identity

$$y(y+1)\cdots(y+m-1) = \sum_{k=1}^{m} L(m,k)y(y-1)\cdots(y-k+1)$$

where

$$L(m,k) = \binom{m}{k}\frac{(m-1)!}{(k-1)!} = \binom{m-1}{k-1}\frac{m!}{k!} \tag{2.68}$$

are the (signless) Lah numbers. They count the number of ways a set of m elements can be partitioned into k nonempty linearly ordered subsets.

If the random variable Z is Poisson (a) distributed, then $E[Z(Z-1)\cdots(Z-j+1)] = a^j$ and if we set $y = Z+2$, then

$$E[(Z+2)(Z+3)\cdots(Z+m+1)] = \sum_{k=1}^{m} L(m,k)E[(Z+2)(Z+1)Z\cdots(Z-k+3)]$$

For $k \geq 3$,

$$\begin{aligned}(Z+2)&(Z+1)Z\cdots(Z-k+3)\\ &= k(k-1)Z(Z-1)\cdots(Z-k+3) + 2kZ(Z-1)\cdots(Z-k+3)(Z-k+2)\\ &\quad +Z(Z-1)\cdots(Z-k+3)(Z-k+2)(Z-k+1)\end{aligned}$$

Taking expectations

$$E[(Z+2)(Z+1)Z\cdots(Z-k+3)] = k(k-1)a^{k-2} + 2ka^{k-1} + a^k$$

which also holds for $k=1$ since $E[Z+2] = a+2$ and for $k=2$ since $E[(Z+2)(Z+1)] = E\left[Z^2\right] + 3E[Z] + 2 = a^2 + 4a + 2$. Hence,

$$E[(Z+2)(Z+3)\cdots(Z+m+1)] = \sum_{k=1}^{m} L(m,k)\left[k(k-1)a^{k-2} + 2ka^{k-1} + a^k\right]$$

Using the last equation and (2.68), (2.67) gives the density in the form

$$\begin{aligned}&\mu e^{-\mu y}e^{-(m+1)\lambda t}e^{(1-e^{-t})\mu y}\left\{\sum_{k=1}^{m}\binom{m-1}{k-1}\frac{m!}{k!}\left[(1-e^{-\lambda t})\mu y\right]^k\right.\\ &\quad +\sum_{k=1}^{m}\binom{m-1}{k-1}\frac{m!}{k!}2k\left[(1-e^{-\lambda t})\mu y\right]^{k-1} + \sum_{k=1}^{m}\binom{m-1}{k-1}\frac{m!}{k!}k(k-1)\left[(1-e^{-\lambda t})\mu y\right]^{k-2}\Bigg\}\\ &\quad +e^{-(m+1)\lambda t}\delta(y)\\ &= m!\mu e^{-e^{-\lambda t}\mu y}e^{-(m+1)\lambda t}\left\{(1-e^{-\lambda t})\mu y + 2\right.\\ &\quad +\sum_{k=2}^{m}\frac{\left[(1-e^{-\lambda t})\mu y\right]^k}{k!}\left[\binom{m-1}{k-1} + 2\binom{m-1}{k} + \binom{m-1}{k+1}\right]\Bigg\}\\ &\quad +e^{-(m+1)\lambda t}\delta(y)\end{aligned}$$

which, since $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$ reduces to

$$\mu e^{-e^{-\lambda t}\mu y} e^{-(m+1)\lambda t} m! \left\{ (1-e^{-\lambda t})\mu y + 2 + \sum_{k=2}^{m} \frac{\left[(1-e^{-\lambda t})\mu y\right]^k}{k!} \binom{m+1}{k+1} \right\} + e^{-(m+1)\lambda t}\delta(y)$$

which is a mixture of Erlang distributed random variables.

3. SUPERPOSITION OF MARKOVIAN ON/OFF SOURCES

In this chapter we study the effect of the superposition of a large number of ON/OFF sources, each generating claims very infrequently. Both the ON and the OFF durations are independent exponential random variables with possibly different ON and OFF rates for each source. We also assume that during the ON period the level of activity is not uniform across sources. The above setup could be used as a model for a risk portfolio consisting of a (large) number of contracts, say n, each of which generates claims. We assume that the claims are not settled immediately but instead that they generate financial obligations for the insurer that last for a period of time. This is in fact the case in situations where reimbursement occurs in a number of installments over a period of time, which may be significant compared to the operating time scale, as opposed to one lump sum. In the model which we study we take the approach that the jth contract does not generate a claim over an exponential period of time with rate λ_j and then becomes active for an independent period of time with distribution G_j. The transition from inactivity to activity is caused by the event against which insurance was purchased and during this period the insurer reimburses the insured at a fixed rate, say r_j. During the active period, the jth contract cannot generate another claim, which is a natural assumption in our framework. We also assume independence across the n contracts. Thus the behavior of each contract, considered individually, is described by an alternating renewal process with exponential OFF times and general ON times.

As a first step towards the analysis of such models we will introduce the simplifying assumption that the active periods are also exponentially distributed and all sources are identical, which leads to a simple markovian model. For this model a diffusion approximation will be developed and it will be shown that, under appropriate conditions, the scaled risk process converges to an Ornstein-Uhlenbeck process for which large deviation results are known and can readily be applied in order to compute ruin probabilities.

We also examine the problem of inhomogeneous portfolios by means of structure functions, using an approach that follows the classic work of Bühlmann [28]. More precisely, we assume that a probability measure is given on $\mathbb{R}_+^3$ and that the parameters $(\lambda_{jn}, \mu_{jn}, r_{jn})$ are i.i.d. random elements distributed according to this probability measure. This is studied in more detail in the next chapter.

3.1 Superposition of a large number of identical, independent sources

Consider a continuous time Markov process $\{\xi_t; t \geq 0\}$ with state space the set $\{0, 1\}$ and generator

$$Q = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}.$$

Following standard usage we will denote by $P_i(B) = P(B|\xi_0 = i)$ for $i = 0, 1$, and $B \in \mathcal{B}$. Similarly, E_i denotes expectation with respect to P_i, $i = 0, 1$. The transition probabilities for this system are given by

$$\begin{bmatrix} P_{00}(t) & P_{01}(t) \\ P_{10}(t) & P_{11}(t) \end{bmatrix} = \begin{bmatrix} \frac{\mu}{\lambda+\mu} & \frac{\lambda}{\lambda+\mu} \\ \frac{\mu}{\lambda+\mu} & \frac{\lambda}{\lambda+\mu} \end{bmatrix} + e^{-(\lambda+\mu)t} \begin{bmatrix} \frac{\lambda}{\lambda+\mu} & -\frac{\lambda}{\lambda+\mu} \\ -\frac{\mu}{\lambda+\mu} & \frac{\mu}{\lambda+\mu} \end{bmatrix}. \tag{3.1}$$

Suppose that we have n contracts, all identical, modeled by a family of independent ON/OFF processes, $\{\xi_{jn}(t)\}$, $i = 1, 2, \ldots, n$ with exponential ON and OFF periods with identical rates for the ON periods, μ_n and for the OFF periods λ_n. (At a later stage we will consider the number of sources, n, as a parameter which takes large values and will examine the limit process as $n \to \infty$. This is the reason for indexing the parameters and the processes by n.) For the time being however we shall assume that the parameters λ and μ do not depend on n. Denote the *superposition process*

$$Y_n(t) \ := \ \sum_{j=1}^{n} \xi_{jn}(t), \tag{3.2}$$

and note that, in steady state,

$$Y_n(t) \ \stackrel{\mathrm{d}}{=} \ \text{Binomial}\left(\frac{\lambda}{\lambda+\mu}, n\right).$$

Since in this case all the sources are identical, $\{Y_n(t); t \geq 0\}$ is the classical Ehrenfest model in continuous time where n particles move independently between two compartments. The time each particle spends in the first and the second compartment respectively are independent, exponential random variables with rates λ and μ. Thus, it is expected that a properly rescaled version of $\{Y_n(t); t \geq 0\}$ will converge to an Ornstein-Uhlenbeck process.

In fact we will show that $\frac{1}{\sqrt{n}}\left(Y_n(t) - EY_n(t)\right)$ converges in distribution, as $n \to \infty$, to an Ornstein-Uhlenbeck process. Indeed, $EY_n(t) = nE\xi_{1n}(t)$ and

$$\frac{1}{\sqrt{n}}\left(Y_n(t) - EY_n(t)\right) \stackrel{\mathrm{d}}{\to} \mathcal{N}\left(0, E\xi_{1n}(t)\left(1 - E\xi_{1n}(t)\right)\right) = \mathcal{N}\left(0, p_t\left(1 - p_t\right)\right) \tag{3.3}$$

where $\mathcal{N}(a, b^2)$ denotes the normal distribution with mean a and variance b^2 and $p_t = E\xi_{1n}(t)$. The covariance is

$$\begin{aligned} &\mathsf{Cov}\left(\tfrac{1}{\sqrt{n}}(Y_n(s) - EY_n(s)), \tfrac{1}{\sqrt{n}}\left(Y_n(t) - EY_n(t)\right)\right) \\ &\qquad = \ \frac{1}{n}\mathsf{Cov}\left(Y_n(s), Y_n(t)\right) \ = \ \mathsf{Cov}\left(\xi_{1n}(s), \xi_{1n}(t)\right) \ =: \ c(s,t) \end{aligned} \tag{3.4}$$

Thus the covariance of the scaled process is equal to the covariance of each source. This can be computed in turn from the transition probabilities as follows

$$c(s,t) \;=\; \left(\tfrac{\lambda}{\lambda+\mu} - \tfrac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)s}\right)\left(\tfrac{\mu}{\lambda+\mu}e^{-(\lambda+\mu)(t-s)} + \tfrac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t}\right), \quad s<t. \tag{3.5}$$

Denote the scaled process by

$$Z_s^{(n)} = \frac{Y_n(s) - EY_n(s)}{\sqrt{n}} = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\zeta_{jn}(s), \qquad s \geq 0,$$

where $\zeta_{jn}(s) := \xi_{jn}(s) - E\left[\xi_{jn}(s)\right]$ are the centered ON/OFF processes. Assume further that the processes $\{\xi_{jn}\}$ are all *stationary* and therefore

$$\zeta_{jn}(s) = \begin{cases} -\dfrac{\lambda}{\lambda+\mu} & \text{with prob. } \dfrac{\mu}{\lambda+\mu} \\ \\ \dfrac{\mu}{\lambda+\mu} & \text{with prob. } \dfrac{\lambda}{\lambda+\mu} \end{cases},$$

$E\left[\xi_{jn}(s)\right] = \frac{\lambda}{\lambda+\mu}$, and hence, from (3.3) we have that

$$Z_s^{(n)} \xrightarrow{\mathrm{d}} \mathcal{N}\left(0, \frac{\lambda\mu}{(\lambda+\mu)^2}\right). \tag{3.6}$$

The covariance function in the stationary case is

$$\mathsf{Cov}(Z_s^{(n)}, Z_t^{(n)}) = \tfrac{\lambda\mu}{(\lambda+\mu)^2}e^{-(\lambda+\mu)(t-s)}, \qquad s<t. \tag{3.7}$$

With little additional effort, using the Cramér-Wold device (see Billingsley, [18]), we can establish the corresponding result for the finite dimensional distributions. Let $k \in \mathbb{N}$ and $0 < s_1 < s_2 < \cdots < s_k$, $\theta_i \in \mathbb{R}$, $i = 1, \ldots, k$. We will show that

$$\left(Z_{s_1}^{(n)}, Z_{s_2}^{(n)}, \ldots, Z_{s_n}^{(n)}\right) \xrightarrow{d} \mathcal{N}\left(0, S\right), \tag{3.8}$$

where S is the covariance matrix with elements $S_{il} := c(s_i, s_l)$ given by (3.7), and hence establish that the finite dimensional distributions of the sequence of processes $Z_\cdot^{(n)}$ converges in distribution to a Gaussian process with mean 0 and covariance function $c(s,t)$. To this end define the random variables $U_{jn} := \sum_{i=1}^{k}\theta_i\zeta_{jn}(s_i)$ and set

$$V_n := \frac{1}{\sqrt{n}}\sum_{j=1}^{n}U_{jn} \;=\; \frac{1}{\sqrt{n}}\sum_{i=1}^{k}\theta_i\sum_{j=1}^{n}\zeta_{jn}(s_i).$$

Then $\{V_n\}_{n=1,2,\ldots}$ is an i.i.d. sequence with mean 0 and variance

$$\sigma^2 \;:=\; \sum_{i=1}^{k}\sum_{l=1}^{k}\theta_i\theta_l c(s_i,s_l) \;=\; \theta^T S\theta.$$

Since the variance is finite, from the Central Limit Theorem we have $V_n \xrightarrow{d} \mathcal{N}\left(0, \theta^T S\theta\right)$ and, since the vector θ was arbitrary, (3.8) follows.

In order to complete the proof and show that $Z_\cdot^{(n)}$ converges in distribution to a Gaussian process with mean 0 and covariance function $c(s,t)$ there remains to establish the tightness of the sequence of processes $\left\{Z_\cdot^{(n)}\right\}$. This is done using Hahn's theorem (see e.g. [146, p. 226]).

Lemma 3.1. *Let $\{\xi_t\}$ be a Markovian ON/OFF source, assumed stationary, and suppose that $u < t < s$. Then*

$$E\left[(\xi_t - \xi_u)^2(\xi_s - \xi_t)^2\right] \leq \frac{1}{4}\lambda\mu(s-u)^2, \tag{3.9}$$

$$|E\left[(\xi_u - \xi_t)(\xi_t - \xi_s)\right]| \leq \frac{1}{4}\lambda\mu(s-u)^2, \tag{3.10}$$

and

$$E(\xi_u - \xi_s)^2 \leq \frac{2\lambda\mu}{\lambda+\mu}\lambda\mu(s-u). \tag{3.11}$$

Proof. The product $(\xi_t - \xi_u)(\xi_s - \xi_t)$ is non zero only when $\xi_u = 0, \xi_t = 1, \xi_s = 0$, and $\xi_u = 1, \xi_t = 0, \xi_s = 1$. In both cases the product is equal to -1. Then

$$\begin{aligned} E\left[(\xi_t - \xi_u)^2(\xi_s - \xi_t)^2\right] &= P(\xi_u = 0)P(\xi_t = 1|\xi_u = 0)P(\xi_s = 0|\xi_t = 1) \\ &\quad + P(\xi_u = 1)P(\xi_t = 0|\xi_u = 1)P(\xi_s = 1|\xi_t = 0) \end{aligned}$$

and thus

$$E\left[(\xi_t - \xi_u)^2(\xi_s - \xi_t)^2\right] = \frac{\lambda\mu}{(\lambda+\mu)^2}(1 - e^{-(\lambda+\mu)(t-u)})(1 - e^{-(\lambda+\mu)(s-t)}).$$

Taking into account the fact that

$$0 \leq 1 - e^{-(\lambda+\mu)x} \leq (\lambda+\mu)x \quad \text{when } x > 0 \tag{3.12}$$

we see that

$$E\left[(\xi_t - \xi_u)^2(\xi_s - \xi_t)^2\right] \leq \frac{\lambda\mu}{(\lambda+\mu)^2}(\lambda+\mu)^2(t-u)(s-t) \leq \lambda\mu\frac{1}{4}(s-u)^2$$

where the last inequality follows from the fact that $(t-u)(s-t) \leq \frac{1}{4}(s-u)^2$. The second inequality is established by taking again into account the fact that $\xi_u - \xi_t$ and $\xi_t - \xi_s$ take values in the set $\{-1, 0, 1\}$. Then,

$$|E\left[(\xi_u - \xi_t)(\xi_t - \xi_s)\right]| \leq E\left[|\xi_u - \xi_t|\,|\xi_t - \xi_s|\right] = E\left[(\xi_u - \xi_t)^2(\xi_t - \xi_s)^2\right]$$

and hence we can use the first inequality of the statement of the Lemma. For the third inequality

$$\begin{aligned} E(\xi_s - \xi_u)^2 &= P(\xi_u = 0)P(\xi_s = 1|\xi_u = 0) + P(\xi_u = 1)P(\xi_s = 0|\xi_u = 1) \\ &= \frac{\mu}{\lambda+\mu}\left(\frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu}\,e^{-(\lambda+\mu)(s-u)}\right) + \frac{\lambda}{\lambda+\mu}\left(\frac{\mu}{\lambda+\mu} - \frac{\mu}{\lambda+\mu}\,e^{-(\lambda+\mu)(s-u)}\right) \\ &= \frac{2\lambda\mu}{(\lambda+\mu)^2}\left(1 - e^{-(\lambda+\mu)(s-u)}\right) \leq \frac{2\lambda\mu}{\lambda+\mu}(s-u) \end{aligned}$$

where, in the last step, we have again used (3.12). □

Theorem 3.2. *(Hahn's Theorem). Let $\{\xi_n; n \geq 1\}$ be a sequence of i.i.d. random elements of $D[0,\infty)$ with $E\left[\xi_n^2(t)\right] < \infty$ for all t. Suppose that, for all $0 < T < \infty$, there exist continuous nondecreasing real-valued functions g and f on $[0,T]$ and numbers $\alpha > \frac{1}{2}, \beta > 1$, such that*

$$E(\xi(u) - \xi(s))^2 \leq (g(u) - g(s))^\alpha \tag{3.13}$$

and

$$E\left[\left(\xi(u)-\xi(t)\right)^2\left(\xi(t)-\xi(s)\right)^2\right] \leq \left(f(u)-f(s)\right)^{\beta} \tag{3.14}$$

for all $0 \leq s \leq t \leq u \leq T$ *and* $u-s<1$. *Then*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\xi_i(t)-E\left(\xi_i(t)\right)\right) \xrightarrow{d} Z(t)$$

where $Z(t)$ *is a zero-mean Gaussian process with covariance function*

$$E\left[Z(s)Z(t)\right] \;=\; \mathsf{Cov}\left(\xi(s),\xi(t)\right)$$

and continuous paths with probability 1.

We will thus show that the scaled limit converges to a Gaussian process with covariance given by (3.7) and hence that it is a stationary Ornstein–Uhlenbeck process provided that we can show the following

Proposition 3.3. *The stationary family* $\{\xi_n\}$ *satisfies the conditions (3.13), (3.14), of Hahn's theorem.*

Proof. The proposition follows immediately from Lemma 3.1. (3.13) is satisfied with $g(u)=u$ and $\alpha=1$. Also, (3.14) holds with $f(u)=u$ and $\beta=2$. We conclude that the conditions of Hahn's theorem are satisfied and consequently the convergence in distribution of the family of scaled processes to the Ornstein-Uhlenbeck process is established. □

3.1.1 The integrated Ornstein–Uhlenbeck process

Let $\{X_t; t\geq 0\}$ be a stationary Ornstein-Uhlenbeck process with

$$X_t \;=\; X_0 e^{-at} + \sigma\int_0^t e^{-a(t-u)}dW_u$$

where $\{W_t; t\geq 0\}$ is standard Brownian motion and X_0 is an independent centered normal random variable with variance $\frac{\sigma^2}{2a}$ (in order to have stationarity). Its covariance is, for $s<t$, given by $E[X_sX_t]=E[X_s^2]e^{-a(t-s)}=\frac{\sigma^2}{2a}e^{-a(t-s)}$. The integrated process $\{Z_t; t\geq 0\}$ obtained as $Z_t=\int_0^t X_u du$ is centered Gaussian with covariance obtained as follows $(s<t)$:

$$\begin{aligned}
EZ_sZ_t &= E\left(\int_0^s X_u du\right)^2 + E\left(\int_s^t X_u du\right)\left(\int_0^s X_v dv\right)\\
&= 2\int_{u=0}^{s}\int_{v=0}^{u} E[X_uX_v]dudv + \int_{u=s}^{t}\int_{v=0}^{s} E[X_uX_v]dudv\\
&= \frac{\sigma^2}{a}\int_{u=0}^{s}\int_{v=0}^{u} e^{-a(u-v)}dudv + \frac{\sigma^2}{2a}\int_{u=s}^{t}\int_{v=0}^{s} e^{-a(u-v)}dudv
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\sigma^2}{a^2}\int_{u=0}^{s}(1-e^{-au})du + \frac{\sigma^2}{2a^2}\int_{u=s}^{t}\int_{v=0}^{s} e^{-a(u-v)}dudv \\
&= \frac{\sigma^2}{a^2}s + \frac{\sigma^2}{2a^3}\left(e^{-at}+e^{-as}-e^{-a(t-s)}-1\right) \\
&= \frac{\sigma^2}{a^2}\left(s+\frac{1}{2a}\left(1-e^{-as}\right)\left(1+e^{-a(t-s)}\right)\right). \qquad (3.15)
\end{aligned}
$$

In particular,

$$\mathsf{Var}(Z_t) = \sigma^2\frac{1+at-e^{-at}}{a^3}.$$

3.2 Cumulative input from the superposition of identical sources

Consider now the additive process arising from the superposition of n identical ON-OFF Markov processes $\{\xi_{jn}(t)\}$, assumed independent and stationary. Setting $m := \frac{\lambda}{\lambda+\mu} = E\xi_{jn}$, the centered cumulative input process is then $\sum_{j=1}^{n}\int_0^t \xi_{jn}(s)ds$ and the corresponding scaled process would be

$$S_n(t) \;:=\; \frac{\sum_{j=1}^{n}\int_0^t \xi_{jn}(s)ds \;-\; nmt}{\sqrt{n}}. \qquad (3.16)$$

We can then show that the finite dimensional distributions of $S_n(t)$ converge to the corresponding to those of $\{X(t); t \geq 0\}$ is the integrated Ornstein-Uhlenbeck process with covariance given by (3.15) with $a = \lambda+\mu$ and $\sigma^2 = \frac{2\lambda\mu}{\lambda+\mu}$. Tightness is again established using Hahn's theorem. Thus $S_n(t)$ converges weakly to $X(t)$.

3.3 Superposition of sources with unequal parameters

We now turn to the case where the input process arises from the superposition of a large number of independent ON/OFF processes *with different parameters*. We will thus consider triangular arrays of independent ON/OFF processes, $\{\xi_{jn}(s); s \geq 0\}_{j=1\ldots k_n}^{n=1,2,\ldots}$. As before *we will assume for simplicity the stationarity of the elements of the family, $\xi_{jn}(s)$, under the probability measure P.* Also we will confine ourselves to the case $k_n = n$. Let

$$B_n^2 \;:=\; \sum_{j=1}^{n}\mathsf{Var}\left(\xi_{jn}(s)\right) \;=\; \sum_{j=1}^{n}\frac{\lambda_{jn}\mu_{jn}}{\left(\lambda_{jn}+\mu_{jn}\right)^2}. \qquad (3.17)$$

(The above quantity does not depend on s because of the stationarity assumption.) Define the *scaled processes* via

$$\zeta_{jn}(s) \;=\; \frac{\xi_{jn}(s) - E\left[\xi_{jn}(s)\right]}{B_n}, \qquad s \geq 0. \qquad (3.18)$$

The scaled sequence of superpositions is defined then as

$$X_n(s) \;=\; \sum_{j=1}^{n}\zeta_{jn}(s) \qquad (3.19)$$

and, of course, it inherits the stationarity from the component processes. The covariance of the scaled superposition X_n is, by appealing to the independence between sources in the same row of the array,

$$\begin{aligned} C_n(\tau) &:= \mathsf{Cov}\left(X_n(s+\tau), X_n(s)\right) = \frac{1}{B_n^2}\sum_{j=1}^{n} \mathsf{Cov}\left(\xi_{jn}(s+\tau), \xi_{jn}(s)\right) \\ &= \frac{1}{B_n^2}\sum_{j=1}^{n} \frac{\lambda_{jn}\mu_{jn}}{(\lambda_{jn}+\mu_{jn})^2}\, e^{-(\lambda_{jn}+\mu_{jn})\,\tau}, \quad \tau \geq 0 \end{aligned} \tag{3.20}$$

where B_n^2 is defined in (3.17).

We will explore then the following question. Assuming that the triangular array is such that the sequence of covariance functions $\{C_n(\tau)\}$ converges to a limiting covariance function $c(\tau)$, we intend to study the conditions under which the sequence $\{X_n(s)\}$ converges weakly to a (zero mean) Gaussian process with covariance $c(\tau)$ as well as explore the possibility of other, non-gaussian limits.

3.4 Marginal distributions via the Lindeberg–Feller theorem

We consider triangular arrays of random values $\{\xi_{jn}\}$, $j = 1, 2, \ldots, n$, $n \in \mathbb{N}$, with values in $\{0, 1\}$. We assume that ξ_{jn}, $j = 1, \ldots, n$ are *independent for each* n. Setting $B_n^2 = \sum_{j=1}^{n} \mathsf{Var}(\xi_{jn})$ and considering the corresponding array of normalized random variables $\zeta_{jn} := B_n^{-1}(\xi_{jn} - E\xi_{jn})$. The Lindeberg–Feller theorem applied to the above triangular array of random variables states that the sequence $\{\sum_{j=1}^{n} \zeta_{jn}\}_{n=1,2,\ldots}$ converges weakly to a normal limit if and only if the following condition holds

$$\lim_{n\to\infty} \sum_{j=1}^{n} E\left[\zeta_{jn}^2 \mathbf{1}(|\zeta_{jn}| > \eta)\right] = 0 \quad \text{for all } \eta > 0. \tag{3.21}$$

Consider for instance a triangular array of Bernoulli random variables with

$$\xi_{jn} = \begin{cases} 1 & \text{w.p.} \quad \frac{1}{n^\alpha}, \\ \\ 0 & \text{w.p.} \quad 1 - \frac{1}{n^\alpha} \end{cases}$$

where $\alpha > 0$. Then the corresponding sum of variances is given by

$$B_n^2 = \sum_{j=1}^{n} \frac{1}{n^\alpha}\left(1 - \frac{1}{n^\alpha}\right) = n^{1-\alpha}\left(1 - \frac{1}{n^\alpha}\right) \to \begin{cases} \infty & \alpha < 1 \\ 1 & \alpha = 1 \\ 0 & \alpha > 1 \end{cases}. \tag{3.22}$$

Note that $|\xi_{jn} - E\xi_{jn}| < 1$. If $\alpha \in (0, 1)$ then $B_n^2 \to \infty$ and hence there exists $n_0(\eta)$ such that $\mathbf{1}(|\zeta_{jn}| > \eta) = \mathbf{1}(|\xi_{jn} - E\xi_{jn}| > \eta B_n) = 0$ when $n > n_0(\eta)$. Hence, when $n > n_0(\eta)$, $E\left[\zeta_{jn}^2 \mathbf{1}(|\zeta_{jn}| > \eta)\right] = 0$ and therefore (3.21) holds and the sequence $\{\sum_{j=1}^{n} \zeta_{jn}\}$ converges weakly to the a standard normal.

When $\alpha = 1$ an easy computation shows that, for $\eta < 1$ and n large enough

$$E\left[\zeta_{jn}^2 \mathbf{1}(|\zeta_{jn}| > \eta)\right] = \frac{1}{n}\left(1 - \frac{1}{n}\right)^2 \frac{1}{\frac{n-1}{n}}$$

and thus

$$\lim_{n\to\infty} \sum_{j=1}^{n} E\left[\zeta_{jn}^2 \mathbf{1}\left(|\zeta_{jn}| > \eta\right)\right] \;=\; \lim_{n\to\infty} \frac{n-1}{n} \;=\; 1.$$

This means that (3.21) does not hold and thus the limit is not normal in this case. In fact, it can be shown that the limit is in this case Poisson.

Finally, when $\alpha > 1$ we can again show that (3.21) does not hold and the limit is not normal.

3.5 Triangular arrays of identical sources

In this section we focus exclusively on the case of a triangular array consisting of *identical sources* with OFF rates $\lambda_{jn} = \lambda_n$ for $j = 1, 2, \ldots, n$ and $\lambda_n \to 0$ as $n \to \infty$. ON rates are all equal, $\mu_{jn} = \mu$ for all j, n. In that case, from (3.17), (3.18), and (3.19)

$$B_n^2 = \frac{n\lambda_n\mu}{(\lambda_n + \mu)^2} \tag{3.23}$$

and the covariance (with $\tau > 0$) is

$$\mathsf{Cov}\left(X_n(s), X_n(s+\tau)\right) = e^{-(\lambda_n+\mu)\tau} \to e^{-\mu\tau} \quad \text{as } n \to \infty. \tag{3.24}$$

This would imply convergence to the Ornstein–Uhlenbeck process, provided that the limit is a Gaussian process. We will consider some cases beginning by examining the behavior of the marginal distributions of the normalized superposition process and investigate some cases depending on the behavior of the sequence $\{\lambda_n\}$. (The convergence of finite dimensional distributions and the tightness of the family of processes can be will be considered later.)

Theorem 3.4. *In triangular arrays with identical ON rates equal to μ, described by their OFF rates $\lambda_{jn} = \lambda_n$ for $j = 1, 2, \ldots, n$ the limiting behavior of the marginal distribution of the normalized process resulting from their superposition,*

$$X_n(s) = \frac{\lambda_n + \mu}{\sqrt{n\lambda_n\mu}} \sum_{j=1}^{n} \left(\xi_{jn}(s) - \frac{\lambda_n}{\lambda_n + \mu}\right), \tag{3.25}$$

depends on the limit of the sequence $\{n\lambda_n\}$ as follows.

1. *If $n\lambda_n \to \infty$, as $n \to \infty$ then the sequence of r.v.'s in (3.25) converges weakly to an standard normal random variable.*
2. *If as $n \to \infty$ $n\lambda_n \to \alpha > 0$, then sequence of r.v.'s in (3.25) converges weakly to $\sqrt{\frac{\mu}{\alpha}}Q - \sqrt{\frac{\alpha}{\mu}}$ where Q is Poisson with mean $\frac{\alpha}{\mu}$.*

3. Finally, if $n\lambda_n \to 0$ as $n \to \infty$, then the limiting process is null i.e. $X_n(s) \xrightarrow{p} 0$ as $n \to \infty$.

Proof. We will establish here only the convergence of the marginal distributions by appealing to the Lindeberg–Feller conditions according to which we have convergence to the normal law if and only if

$$\lim_{n\to\infty} \sum_{j=1}^{n} E\left[\zeta_{jn}^2 \mathbf{1}(|\zeta_{jn}| > \eta)\right] = 0 \quad \text{for all } \eta > 0.$$

where ζ_{jn} is defined in (3.18) or, since all sources are identical,

$$\lim_{n\to\infty} nE[\zeta_{jn}^2 \mathbf{1}(|\zeta_{jn}| > \eta)] \;=\; 0.$$

This last condition is equivalent to

$$\lim_{n\to\infty} \frac{n}{B_n^2} E[(\xi_{jn} - E(\xi_{jn}))^2 \mathbf{1}(|\xi_{jn} - E(\xi_{jn})| > \eta B_n)] = 0. \tag{3.26}$$

Since $\xi_{jn} \in \{0, 1\}$ it is easy to see that

$$|\xi_{jn} - E(\xi_{jn})| < 1. \tag{3.27}$$

CASE 1. In the first case, from (3.23) we see that $n\lambda_n \to \infty$ implies that $B_n \to \infty$ and hence, for any given $\eta > 0$ there exists $n_0(\eta)$ such that $n > n_0(\eta)$ implies $B_n\eta > 1$ and hence that $E[(\xi_{jn} - E(\xi_{jn}))^2 \mathbf{1}(|\xi_{jn} - E(\xi_{jn})| > \eta B_n)] = 0$ when $n > n_0(\eta)$. Therefore (3.26) holds and we conclude that in this case (3.25) converges to a standard normal r.v. in distribution.

CASE 2. Then, in view of (3.23), $B_n^2 \to \frac{\alpha}{\mu}$ and thus, because of (3.27), there exists $\eta > 0$ and $n_1(\eta)$ such that, for all $n > n_1(\eta)$, $|\xi_{jn} - E(\xi_{jn})| > \eta B_n$. Then

$$\frac{n}{B_n^2} E[(\xi_{jn} - E(\xi_{jn}))^2 \mathbf{1}(|\xi_{jn} - E(\xi_{jn})| > \eta B_n)] = \frac{n}{B_n^2} E[(\xi_{jn} - E(\xi_{jn}))^2] = 1$$

and hence (3.26) is not satisfied which means that the limit distribution, if it exists, cannot be normal. A direct computation can answer this question. If $\phi_n(t) := Ee^{itX_n(s)}$ then

$$\phi_n(t) := \left(1 + \frac{\lambda_n}{\lambda_n + \mu}\left(e^{it\frac{\lambda_n+\mu}{\sqrt{n\lambda_n\mu}}} - 1\right)\right)^n e^{-it\sqrt{\frac{n\lambda}{\mu}}}.$$

and thus

$$\log \phi_n(t) \;=\; n \log\left(1 + \frac{\lambda_n}{\lambda_n + \mu}\left(e^{it\frac{\lambda_n+\mu}{\sqrt{n\lambda_n\mu}}} - 1\right)\right) - it\sqrt{\frac{n\lambda}{\mu}}.$$

Taking into account that $n\lambda_n \to \alpha$ we have

$$\lim_{n\to\infty} \log \phi_n(t) \;=\; \frac{\alpha}{\mu}\left(e^{it\sqrt{\frac{\mu}{\alpha}}} - 1\right) - it\sqrt{\frac{\alpha}{\mu}}$$

We have thus seen that

$$\phi_n(t) \to \phi(t) := e^{\frac{\alpha}{\mu}\left(e^{it\sqrt{\frac{\mu}{\alpha}}} - 1\right) - it\sqrt{\frac{\alpha}{\mu}}}. \tag{3.28}$$

The right hand side on (3.28) corresponds to the characteristic function of the random variable $X(s) := \sqrt{\frac{\mu}{\alpha}}Q - \sqrt{\frac{\alpha}{\mu}}$ where Q is Poisson with rate $\frac{\alpha}{\mu}$. The continuity theorem for characteristic functions shows that the sequence of random variables $\{X_n(s)\}$ converges weakly to $X(s)$.

CASE 3. Finally, we examine the case where $n\lambda_n \to 0$. Then, from (3.23) it follows that $B_n \to 0$ and hence, arguing as in the previous case, we see that for each $\eta > 0$ threre exists $n_0(\eta)$ such that

$$\left|\xi_n - \frac{\lambda_{jn}}{\mu + \lambda_{jn}}\right| > \eta B_n \quad \text{for all } n > n_0(\eta).$$

Therefore for $n > n_0(\eta)$

$$\frac{n}{B_n^2} E[(\xi_{jn} - E(\xi_{jn}))^2 \mathbf{1}(|\xi_{jn} - E(\xi_{jn})| > \eta B_n)] \;=\; \frac{n}{B_n^2} E[(\xi_{jn} - E(\xi_{jn}))^2] \;=\; 1$$

and hence (3.26) cannot hold.

The limit law can be found using the approach of Gnedenko [69, p. 278-280] or using a direct argument based on characteristic functions as follows:

The limiting behavior can be found then by a direct argument. The Bernoulli random variables $\{\xi_{jn}\}$, $i = 1, 2, \ldots n$, $n = 1, 2, \ldots$, form a triangular array such that $P(\xi_{jn} = 0) = \frac{\mu}{\lambda_n + \mu}$ and $P(\xi_{jn} = 1) = \frac{\lambda_n}{\lambda_n + \mu}$. The characteristic function of the superposition of the normalized variables can be computed as follows:

$$\begin{aligned}
\varphi_n(t) &= E\left[e^{it\frac{1}{B_n}\sum_{j=1}^{n}(\xi_{jn} - E(\xi_{jn}))}\right] = \prod_{j=1}^{n}\left(\frac{\mu}{\lambda_n + \mu} + e^{i\frac{t}{B_n}}\frac{\lambda_n}{\lambda_n + \mu}\right) e^{-i\frac{t}{B_n}\frac{\lambda_n}{\lambda_n + \mu}} \\
&= \left[1 + \frac{\lambda_n}{\lambda_n + \mu}\left(e^{i\frac{t}{B_n}} - 1\right)\right]^n e^{-i\frac{t}{B_n}\frac{n\lambda_n}{\lambda_n + \mu}}
\end{aligned}$$

For all n sufficiently large $\frac{\lambda_n}{\lambda_n + \mu}\left|e^{i\frac{t}{B_n}} - 1\right| < \frac{1}{2}$, hence the logarithm of the characteristic function above is well-defined. Thus, using the expression for B_n given in (3.23), we have

$$\log \varphi_n(t) = n \log\left[1 + \frac{\lambda_n}{\lambda_n + \mu}\left(e^{i\frac{t}{B_n}} - 1\right)\right] - i\frac{t}{\sqrt{\mu}}\sqrt{n\lambda_n}$$

and since $n\lambda_n \to 0$ as $n \to \infty$, we only have to focus on the first term on the right hand side of the above equation. Using the inequality

$$\frac{|x|}{1 + |x|} < |\log(1 + x)| < |x|$$

we obtain

$$\frac{\left|\frac{n\lambda_n}{\lambda_n + \mu}\left(e^{i\frac{t}{B_n}} - 1\right)\right|}{1 + \left|\frac{\lambda_n}{\lambda_n + \mu}\left(e^{i\frac{t}{B_n}} - 1\right)\right|} < n\left|\log\left(1 + \frac{\lambda_n}{\lambda_n + \mu}\left(e^{i\frac{t}{B_n}} - 1\right)\right)\right| < \left|\frac{n\lambda_n}{\lambda_n + \mu}\left(e^{i\frac{t}{B_n}} - 1\right)\right|.$$

As $n \to \infty$ the first and the third term of the above inequality go to zero (Note that $\left|e^{-i\frac{t}{B_n}} - 1\right| < 2$). This means that $\log \varphi_s(t) \to 0$ or $\varphi_s(t) \to 1$. Hence, an appeal to the continuity theorem for characteristic functions completes the proof of the assertion that the limiting process is null. □

Remark. The above theorem establishes only the converges of marginal distributions. Convergence of finite dimensional distributions and tightness, thus establishing weak convergence at the level of processes when the limit is Gaussian, is shown in the sequel. Chapter 4 gives examples and applications. In Case 2 above we have convergence of the marginal distributions to a (scaled and translated) Poisson random variable. Because the covariance is $\mathsf{Cov}(X(s+\tau), X(s)) = e^{-\mu\tau}$ we expect that the sequence of processes $\{X_n(s)\}_{s\in[0,t]}$ converges to a stationary (scaled) $M/M/\infty$ process. This, and other related results, are discussed in Chapter 5.

The following theorem in Billingsley [18] guarantees the tightness of a family of processes in $\mathbb{D}[0,T]$.

Theorem 3.5. *Consider a sequence of processes $\{X_n(t); t \in [0,T]\}$ and a process $\{X(t); t \in [0,T\}$ shuch that, for all finite dimensional distributions we have*

$$(X_n(t_1), \ldots, X_n(t_k)) \Longrightarrow (X(t_1), \ldots, X(t_k)) \tag{3.29}$$

for all $k \in \mathbb{N}$ and $t_i \in [0,T]$ and further that

$$X_T - X_{T-\delta} \Longrightarrow 0 \quad \text{as } \delta \to 0. \tag{3.30}$$

Finally, for any need to show that for $t_1 \le t_2 \le t_3$ the following inequality holds

$$E\left[|X_n(t_2) - X_n(t_1)|^\beta \, |X_n(t_3) - X_n(t_2)|^\beta\right] \le K(t_3 - t_1)^{1+\alpha} \tag{3.31}$$

for some $\alpha > 0$, $\beta > 1$ and $K > 0$.

Suppose now that $\{\lambda_{jn}\}$ is a triangular array of positive real numbers corresponding to the OFF rates of alternating processes with ON periods with rate μ (same for all sources). Thus each $\xi_{jn}(s)$ is an ON/OFF source and we consider the superposition $\sum_{j=1}^n \xi_{jn}(s)$ and the corresponding normalized processes

$$X_n(s) = \frac{1}{B_n} \sum_{j=1}^n \left(\xi_{jn}(s) - \frac{\lambda_{jn}}{\lambda_{jn} + \mu_{jn}} \right). \tag{3.32}$$

with B_n given in (3.17).

Theorem 3.6. *With the definitions of the above paragraph, suppose that the following assumptions hold.*

A1. $\liminf_{n\to\infty} B_n = +\infty,$ (3.33)

A2. $\lim_{n\to\infty} \frac{1}{B_n^2} \sum_{j=1}^n \frac{\lambda_{jn}\mu_{jn}}{(\lambda_{jn}+\mu_{jn})^2} e^{-(\lambda_{jn}+\mu_{jn})t} = c(t)$ *uniformly on* $[0,T]$ *for all* $T > 0$, (3.34)

A3. $\lim_{t\to 0} \frac{c(0) - c(t)}{t} = \gamma > 0.$ (3.35)

A4. $K_1 \le \liminf_{n\to\infty} \frac{1}{B_n^2} \sum_{j=1}^n \lambda_{jn}\mu_{jn} \le \limsup_{n\to\infty} \frac{1}{B_n^2} \sum_{j=1}^n \lambda_{jn}\mu_{jn} \le K_2$ (3.36)

for some positive constants K_1, K_2.

Then $\{X_n(\cdot)\}$ converges weakly to a zero mean Gaussian process $\{X(\cdot)\}$ with covariance function $c(\cdot)$ in the space $\mathbb{D}[0,\infty)$ with the topology of uniform convergence. The limiting Gaussian process X may be taken to have continuous paths with probability one.

Proof. We will first show that there exists a version of $\{X\}$ that has continuous paths with probability 1. Indeed, since the process is Gaussian,

$$|X(t+h) - X(t)|^4 \;=\; 3(c(0) - c(h))^2. \tag{3.37}$$

In view of (3.35) we conclude that there exists $K > 0$ such that $|X(t+h) - X(t)|^4 \;\leq\; K|h|^2$ and hence, by the Kolmogorov-Chentsov theorem there exists a version of the process which has continuous paths with probability 1. This in turn implies that condition (3.30) of Theorem 3.5 is satisfied.

We first consider the one-dimensional case. For each $\eta > 0$ let

$$V_n(\eta) \;= \sum_{j=1}^{n} E[\zeta_{jn}^2(s)\mathbf{1}(|\zeta_{jn}(s)| > \eta)]. \tag{3.38}$$

We will show that

$$V_n(\eta) \to 0 \quad \text{as } n \to \infty \text{ for every } \eta > 0. \tag{3.39}$$

Note that $B_n\,|\zeta_{jn}(s)| = \left|\xi_{jn} - \frac{\lambda_{jn}}{\lambda_{jn}+\mu_{jn}}\right| \leq 1$ and hence, because of A1, for any $\eta > 0$ there exists $n_0(\eta)$ such that $\eta B_n > 1$ when $n \geq n_0(\eta)$ which implies that

$$\mathbf{1}\left(|\zeta_{jn}| > \eta\right) \;=\; \mathbf{1}\left(\left|\xi_{jn} - \frac{\lambda_{jn}}{\lambda_{jn}+\mu_{jn}}\right| > \eta B_n\right) \;=\; 0 \quad \text{when } n > n_0(\eta)\ .$$

Therefore,

$$E\left[\zeta_{jn}^2\mathbf{1}(|\zeta_{jn}| > \eta)\right] \;=\; 0 \quad \text{when } n > n_0(\eta)$$

and thus

$$\lim_{n\to 0} V_n(\eta) \;=\; 0. \tag{3.40}$$

Hence the Lindeberg-Feller condition is satisfied and therefore $X_n \to Z \sim \mathcal{N}(0,1)$.

We now turn to the finite dimensional distributions. Suppose $(\theta_1, \ldots, \theta_m) \in \mathbb{R}^m$ and consider $\sum_{k=1}^m \theta_k X_n(s_k)$. Then

$$\sum_{k=1}^{m} \theta_k X_n(s_k) \;=\; \sum_{k=1}^{m} \theta_k \sum_{j=1}^{n} \frac{\xi_{jn}(s_k) - m_{jn}}{B_n} \;=\; \sum_{j=1}^{n}\sum_{k=1}^{m} \theta_k \zeta_{jn}(s_k).$$

Set

$$D_n^2 \;=\; \sum_{j=1}^{n} \mathsf{Var}\left(\sum_{k=1}^{m} \theta_k \zeta_{jn}(s_k)\right) \;=\; \sum_{j=1}^{n}\sum_{k=1}^{m}\sum_{l=1}^{m} \frac{\theta_k\theta_l}{B_n^2} \frac{\lambda_{jn}\mu_{jn}}{(\lambda_{jn}+\mu_{jn})^2}\, e^{-(\lambda_{jn}+\mu_{jn})|s_k - s_l|}. \tag{3.41}$$

We define $\chi_{jn} := \frac{1}{D_n}\sum_{k=1}^n \theta_k \zeta_{jn}(s_k)$ and we will show that this double array satisfies the Lindeberg-Feller conditions, i.e. that

$$\lim_{n\to\infty} \sum_{j=1}^{n} \chi_{jn}^2 \mathbf{1}(|\chi_{jn}| > \eta) \;=\; 0, \quad \text{for all } \eta > 0. \tag{3.42}$$

Mimicking the proof of the first part we define

$$U_n \;=\; \sum_{j=1}^{n} \chi_{jn}^2 \mathbf{1}(|\chi_{jn}| > \eta).$$

We will establish that there exists $n_3(\eta)$ such that $|\chi_{jn}| < \eta$ for all $n > n_3(\eta)$. Indeed

$$|\chi_{jn}| < \eta \;\Leftrightarrow\; \left|\sum_{k=1}^{m} \theta_k \zeta_{jn}(s_k)\right| > \eta D_n \;\Leftrightarrow\; \left|\sum_{k=1}^{m} \theta_k (\xi_{jn}(s_k) - m_{jn})\right| > \eta B_n D_n.$$

In view of the assumption (3.34) for any given $\epsilon > 0$ there exists n_4 such that $n > n_4$ implies that

$$c(|s_k - s_l|) - \epsilon \;<\; \frac{1}{B_n^2} \sum_{j=1}^{n} \frac{\lambda_{jn}\mu_{jn}}{(\lambda_{jn}+\mu_{jn})^2} e^{-(\lambda_{jn}+\mu_{jn})|s_k - s_l|} \;<\; c(|s_k - s_l|) + \epsilon \tag{3.43}$$

and hence

$$\begin{aligned} \sum_{k,l=1}^{m} \theta_k\theta_l (c(|s_k - s_l|) - \epsilon) \;&<\; \sum_{j=1}^{n} \sum_{k,l=1}^{m} \frac{\theta_k\theta_l}{B_n^2} \frac{\lambda_{jn}\mu_{jn}}{(\lambda_{jn}+\mu_{jn})^2} e^{-(\lambda_{jn}+\mu_{jn})|s_k - s_l|} \\ &<\; \sum_{k,l=1}^{m} \theta_k\theta_l (c(|s_k - s_l|) + \epsilon) \end{aligned} \tag{3.44}$$

Consider now the $m \times m$ covariance matrix $[c(|s_k - s_l|)]$, $k, l = 1, \ldots, m$ which is symmetric and positive definite. Let $0 < \rho_{\min} < \rho_{\max}$ denote the smallest and largest eigenvalue of the covariance matrix. Then (3.44) can be rewritten as

$$(\rho_{\min} - \epsilon)|\theta|^2 \;<\; \sum_{j=1}^{n} \sum_{k,l=1}^{m} \frac{\theta_k\theta_l}{B_n^2} \frac{\lambda_{jn}\mu_{jn}}{(\lambda_{jn}+\mu_{jn})^2} e^{-(\lambda_{jn}+\mu_{jn})|s_k - s_l|} \;<\; (\rho_{\max} + \epsilon)|\theta|^2. \tag{3.45}$$

Therefore

$$(\rho_{\min} - \epsilon)^{1/2}|\theta| B_n \;<\; D_n \;<\; (\rho_{\max} + \epsilon)^{1/2}|\theta| B_n.$$

Then repeating the argument used for the one-dimensional case we establish (3.42) and therefore, using the Cramér-Wold device the convergence of the finite dimensional distributions to those of those of a Gaussian process with covariance function $c(s)$.

We next establish tightness. Suppose that $0 < u < t < s$. We have

$$E(X_n(t) - X_n(u))^2 (X_n(s) - X_n(t))^2 \;=\; \frac{1}{B_n^4} E \left(\sum_{j=1}^{n} (\xi_{jn}(t) - \xi_{jn}(u))\right)^2 \left(\sum_{j=1}^{n} (\xi_{jn}(s) - \xi_{jn}(t))\right)^2.$$

The expectation on the right can be written as the sum of three non-zero terms,

$$\begin{aligned} I \;&=\; \sum_{j=1}^{n} E\left[(\xi_{jn}(t) - \xi_{jn}(u))^2 (\xi_{jn}(s) - \xi_{jn}(t))^2\right] \\ II \;&=\; \sum_{j=1}^{n} \sum_{k \neq j} E(\xi_{jn}(t) - \xi_{jn}(u))^2 E(\xi_{kn}(s) - \xi_{kn}(t))^2 \\ III \;&=\; 2 \sum_{j=1}^{n} \sum_{k \neq j} E(\xi_{jn}(t) - \xi_{jn}(u))(\xi_{jn}(s) - \xi_{jn}(t))\, E(\xi_{kn}(t) - \xi_{kn}(u))(\xi_{kn}(s) - \xi_{kn}(t)) \end{aligned}$$

Then, taking into account Lemma 3.1 and (**??**) we obtain the bounds

$$\begin{aligned} I &< \frac{1}{4}(s-u)^2 \sum_{j=1}^{n} \lambda_{jn}\mu_{jn} \\ II &< \frac{1}{4}(s-u)^2 \sum_{j=1}^{n}\sum_{k\neq j} \frac{\lambda_{jn}\mu_{jn}}{\lambda_{jn}+\mu_{jn}} \frac{\lambda_{kn}\mu_{kn}}{\lambda_{kn}+\mu_{kn}} \\ III &< \frac{2}{16}(s-u)^4 \sum_{j=1}^{n}\sum_{k\neq j} \lambda_{jn}\mu_{jn}\lambda_{kn}\mu_{kn} \end{aligned}$$

and

$$E(X_n(t)-X_n(u))^2(X_n(s)-X_n(t))^2 \;\leq\; \frac{(s-u)^2}{4B_n^4}\left(\sum_{j=1}^{n}\lambda_{jn}\mu_{jn} + \frac{3}{4}(s-u)^2\sum_{j=1}^{n}\sum_{k\neq j}\lambda_{jn}\mu_{jn}\lambda_{kn}\mu_{kn}\right)$$

Then, under Assumptions (3.33), (3.36), of Theorem 3.6 $\lim_{n\to\infty}\frac{1}{B_n^4}\sum_{j=1}^{n}\sum_{k\neq j}\lambda_{jn}\mu_{jn}\lambda_{kn}\mu_{kn} \leq K_2$. Also, from (3.34) and (3.35) for $\epsilon > 0$ there exists $\delta > 0$ such that $|t| < \delta$ implies

$$\gamma - \epsilon \leq \lim_{n\to\infty}\frac{1}{B_n^2}\sum_{j=1}^{n}\frac{\lambda_{jn}\mu_{jn}}{(\lambda_{jn}+\mu_{jn})^2}\frac{1-e^{-(\lambda_{jn}+\mu_{jn})t}}{t} \leq \gamma+\epsilon. \tag{3.46}$$

From this we conclude that there also exists N_0 such that, for $n \geq N_0$

$$\gamma - 2\epsilon \;\leq\; \frac{1}{B_n^2}\sum_{j=1}^{n}\frac{\lambda_{jn}\mu_{jn}}{(\lambda_{jn}+\mu_{jn})^2}\frac{1-e^{-(\lambda_{jn}+\mu_{jn})t}}{t} \;\leq\; \gamma+2\epsilon. \tag{3.47}$$

We have $\frac{1-e^{-(\lambda_{jn}+\mu_{jn})t}}{t} = \frac{1}{t}\int_0^{(\lambda_{jn}+\mu_{jn})t} e^{-u}du$ and the inequality

$$(\lambda_{jn}+\mu_{jn})e^{-(\lambda_{jn}+\mu_{jn})t} \;\leq\; \frac{1}{t}\int_0^{(\lambda_{jn}+\mu_{jn})t} e^{-u}du \;\leq\; \lambda_{jn}+\mu_{jn}.$$

Therefore, from (3.47),

$$\frac{1}{B_n^2}\sum_{j=1}^{n}\frac{\lambda_{jn}\mu_{jn}}{\lambda_{jn}+\mu_{jn}}e^{-(\lambda_{jn}+\mu_{jn})t} \;\leq\; \gamma+2\epsilon. \tag{3.48}$$

Since the above holds for all $t < \delta$, letting $t \to 0$ in the above we obtain (with $K_4 = \gamma + 2\epsilon$).

$$\frac{1}{B_n^2}\sum_{j=1}^{n}\frac{\lambda_{jn}\mu_{jn}}{\lambda_{jn}+\mu_{jn}} \;\leq\; K_4 \tag{3.49}$$

and hence

$$E(X_n(t)-X_n(u))^2(X_n(s)-X_n(t))^2 \;\leq\; K(s-u)^4$$

where $K = \frac{3}{16}K_2$. This establishes tightness. □

4. STRUCTURE FUNCTIONS AND LIMITS OF SUPERPOSITIONS OF TWO-STATE MARKOVIAN SOURCES

4.1 Structure Functions

Our approach is in the spirit of Bühlmann [28]. The concept of *collective* as a set of risks is defined. Each risk is labeled with a parameter (characteristic quantity) ϑ. This parameter can be defined very generally: in particular it need not be a real number (it can be a real pair, etc.). Briefly: Collective $\Theta = \{\vartheta\}$, the totality of risks identified by the parameter ϑ. The risk ϑ in the collective Θ is then denoted by the symbol ϑ or, in somewhat greater detail risk $\vartheta = [N_t^{(\vartheta)}, F_t^{(\vartheta)}]$.

Collective of risks plays a decisive role in connecting a number of claims processes and the corresponding amount jump distribution with any risk which have distribution functions more or less unknown but we treat them as known.

Consider a collective Θ of risks ϑ characterized by $N_t^{(\vartheta)}$ the homogeneous Poisson process corresponding to ϑ with probabilities $p_k(t) = \frac{(\vartheta t)^k}{k!} e^{-\vartheta t}$ and by the distributions $F_t^{(\vartheta)}$ which are independent of t and ϑ, and such that $F_t^{(\vartheta)} = F(x - y)$. Then the following accumulated claim distribution applies in respect of the risk with parameter ϑ :

$$G_t^{(\vartheta)}(x) = \sum_{k=0}^{\infty} e^{-\vartheta t} \frac{(\vartheta t)^k}{k!} F^{*k}(x).$$

We know therefore the form of the distribution function $G_t^{(\vartheta)}(x)$ (compound Poisson). We think of the parameter ϑ, however, as being unknown. This situation is typical of all problems were we are considering a risk ϑ in a collective Θ.

To the extend that the probability characteristics vary from one risk to another, a description of the collective's structure is necessary. More precisely, if Θ' is some portion of the collective Θ, we require an indication of the probability with which a risk taken at random from the collective falls in the portion Θ'. The function which provides this information is called the *structure function.* In many cases it is possible to determine this function approximately by means of statistical investigations (see Bichsel, [17]). He approximated the structure function from the distribution of the observed number of claims among the group of Swiss automobile drivers. Also, in the case in which there is no prior information for the analysis, the assumption of structure functions can be justified. It then simply amounts to a personal appraisal of the collective's structure by an underwriter.

As we have said previously, the parameter ϑ need not necessarily be a real number (linear

parameter), even though it will be in numerous practical cases. The collective Θ can then be represented by the real line and the structure function will be the distribution function $U_\vartheta^{(t)} = P(\vartheta \leq t)$ on the real axis. Similarly if the parameter is in vector form $\vartheta = (\vartheta_1, \vartheta_2, \ldots, \vartheta_n)$, where ϑ_k are real numbers, the structure function will be a distribution function in real $n-$space.

Let us consider a class of stochastic processes $\{X_t^{(\vartheta)}; t \geq 0\}$. Each event A (set of sampling functions) defined on the process then has a probability $P^{(\vartheta)}(A)$ which depends on ϑ. The weighted stochastic process is the process for which

$$P(A) = \int_\Theta P^{(\vartheta)}(A)\, dU(\vartheta).$$

If we denote by $\{S_t^{(\vartheta)}; t \geq 0\}$, the accumulated claim process of the risk with known parameter ϑ, by weighting over the parameter ϑ by means of the structure function $U(\vartheta)$ we obtain $\{S_t; t \geq 0\}$, the weighted accumulated claim process of the risk in the collective.

We consider portfolios consisting of a large number of processes with delayed claim settlement. Each process can be modeled as an ON/OFF process possibly with different exponential rate of occurrence, claim size and settlement duration. For instance, the settlement duration distribution $F(x;\theta)$ could depend on a parameter θ which differs from one risk to another in the portfolio. In that case the portfolio could be described as a collection of values $\{(\lambda_i, \theta_i);\ i = 1, 2, \ldots n\}$ where n is the number of risks.

4.2 Empirical Structure Functions for Contracts with Identical ON Rates

Since the analysis of the total claim process resulting from the superposition of the individual claims in the portfolio is harder than in the classical compound Poisson model, we consider appropriate diffusion approximations. In the simplest possible case all risks have the same characteristics i.e. $(\lambda_j, \theta_j) = (\lambda, \theta)$, $j = 1, 2, \ldots n$. In general however, either the rate of occurrence λ or the distribution parameter θ, or both, may vary across the portfolio.

In order to approximate the total claim process resulting from such portfolios using limit theorems, we will describe portfolios by means of empirical measures in appropriate spaces (e.g. in $\mathbb{R}^2$ when θ is a real parameter so that each risk can be described by means of a point on the plane (λ, θ)). In the simplest case these empirical measures will be i.i.d. samples from a distribution on the space of parameters which plays the role of a structure function.This section is based on ideas similar to the structure function approach introduced by Bühlmann [28]. We consider portfolios made up of a large number of contracts, say n, with parameters $\{\lambda_{jn},\ j = 1, 2, \ldots, n\}$. We will examine both the case where the values of the parameters λ_{jn} come from a given empirical distribution and the limiting case where the portfolio consists of a whole continuum of contracts. In that case the empirical distribution becomes a probability distribution. This model is directly inspired from classical models in risk theory such as the negative binomial model etc. To formalize this idea we may think of a sequence of empirical measures

$$\Lambda_n(dx) \ := \ \frac{1}{n} \sum_{j=1}^{n} \varepsilon_{\lambda_{jn}}(dx), \qquad n = 1, 2, \ldots, \tag{4.1}$$

where $\varepsilon_\alpha(dx)$ denotes the Dirac measure which has a single atom of size 1 at the point $\alpha \in \mathbb{R}$. This includes the analysis of the previous section as a special case. Interesting models can be obtained in the continuous limit, i.e. if we assume that

$$\Lambda_n(x) \Longrightarrow \Lambda(x) \quad \text{as} \quad n \to \infty, \tag{4.2}$$

where $\Longrightarrow$ denotes weak convergence.

Recall that the sum of variances of the occupation random variables ξ_{jn} is equal to

$$B_n^2 = \sum_{j=1}^{n} \frac{\lambda_{jn}\mu}{(\lambda_{jn}+\mu)^2}.$$

Let us define the empirical probability measure

$$F_n(dx) \;=\; \frac{1}{B_n^2} \sum_{j=1}^{n} \frac{\lambda_{jn}\mu}{(\lambda_{jn}+\mu)^2} \varepsilon_{\lambda_{jn}}(dx). \tag{4.3}$$

(It is easy to check that $F_n(\mathbb{R}^+) = 1$). Furthermore, if $g : \mathbb{R}^+ \to \mathbb{R}^+$ is the function defined by $g(x) = \frac{x\mu}{(\mu+x)^2}$, then

$$F_n(dx) = \frac{n}{B_n^2} g(x) \Lambda_n(dx)$$

where we may write $B_n^2 = n \int_{\mathbb{R}^+} g(x)\Lambda_n(dx)$ and

$$F_n(dx) = \frac{g(x)\Lambda_n(dx)}{\int_{\mathbb{R}^+} g(x)\Lambda_n(dx)}.$$

We will investigate the limit as $\Lambda_n(dx) \Longrightarrow \Lambda(dx)$ and $F_n(dx) \Longrightarrow \frac{g(x)\Lambda(dx)}{\int_{\mathbb{R}^+} g(x)\Lambda(dx)}$. Recall that the lag–$\tau$ covariance of the number of active sources in steady state is given by

$$C(\tau) \;=\; \frac{1}{B_n^2} \sum_{j=1}^{n} \frac{\lambda_{jn}\mu}{(\mu+\lambda_{jn})^2} e^{-(\mu+\lambda_{jn})\tau} \;=\; \int_{\mathbb{R}^+} e^{-(\mu+x)\tau} F_n(x). \tag{4.4}$$

The last equation above expresses the covariance in terms of the empirical probability distribution $F_n(x)$. In the limit, if $\Lambda_n(dx) \Longrightarrow \Lambda(x)$, then $F_n(dx) \Longrightarrow F(x)$, by Helly's second theorem

$$\int_{\mathbb{R}^+} e^{-(\mu+x)\tau} F_n(x) \to \int_{\mathbb{R}^+} e^{-(\mu+x)\tau} F(dx).$$

It remains to argue that the superposition of processes converges weakly to a stationary Gaussian process with covariance function $\int_{\mathbb{R}^+} e^{-(\mu+x)\tau} F(x)$ which can be thought of as a continuous mixture of Ornstein–Uhlenbeck processes.

4.3 Structure Functions arising from i.i.d. parameters

Here we discuss the special case of triangular arrays of ON/OFF processes with constant ON rates, equal to μ. The nth row of the array has OFF rates λ_{jn} that are i.i.d. non-negative random variables. As n grows these rates become uniformly smaller. The following proposition discusses such a model and establishes that the assumptions of Theorem 3.6 are satisfied in this case.

Proposition 4.1. *Suppose that $\{\Lambda_j; k = 1, 2, \ldots\}$ are i.i.d. positive random variables with finite first moment, $E\Lambda$, and suppose that $\lambda_{jn} = n^{-\alpha}\Lambda_j$. Suppose that μ is a positive constant and consider the family of functions $f_t : [0,\infty) \to [0,\infty)$ parametrized by $t \in [0,\infty)$ defined as $f_t(x) = \frac{x\mu}{(x+\mu)^2} e^{-(x+\mu)t}$. In particular*

$$\lim_{n\to\infty} \frac{1}{B_n^2} \sum_{j=1}^{n} f_t(\lambda_{jn}) = e^{-\mu t} \quad \text{uniformly on } [0,T], \text{ w.p.1.} \tag{4.5}$$

Proof. Consider the family of functions $h_t : \mathbb{R}^+ \to \mathbb{R}$ given by

$$h_t(x) = \left(\frac{\mu}{x+\mu}\right)^2 e^{-xt}.$$

Note that we have $0 < h_t(x) \leq 1$ for all $x \in \mathbb{R}^+$ and each $t \geq 0$.

We will first establish that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \Lambda_j h_t\left(\Lambda_j n^{-\alpha}\right) \;=\; \frac{1}{\mu} E[\Lambda] \quad \text{uniformly on } [0,T] \text{ for each } T > 0 \text{ w.p.1.} \tag{4.6}$$

Fix $\delta > 0$. Then, since h_t is, for each t, a strictly decreasing function,

$$\frac{1}{n} \sum_{j=1}^{n} \Lambda_j h_t(\delta) \mathbf{1}\left(\Lambda_j \leq \delta n^{\alpha}\right) \leq \frac{1}{n} \sum_{j=1}^{n} \Lambda_j h_t\left(\Lambda_j n^{-\alpha}\right) \leq \frac{1}{n} \sum_{j=1}^{n} \Lambda_j$$

Now fix $M \in \mathbb{N}$. Then

$$\frac{1}{n} \sum_{j=1}^{n} \Lambda_j h_t(\delta) \mathbf{1}\left(\Lambda_j \leq M^{\alpha}\delta\right) \leq \frac{1}{n} \sum_{=1}^{n} \Lambda_j h_t\left(\Lambda_j n^{-\alpha}\right) \leq \frac{1}{n} \sum_{j=1}^{n} \Lambda_j \qquad \text{for } n \geq M.$$

Letting $n \to \infty$ in the above double inequality we obtain from the Strong Law of the Large Numbers,

$$\begin{aligned} h_t(\delta) E[\Lambda \mathbf{1}(\Lambda \leq M^{\alpha}\delta)] \;&\leq\; \liminf_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \Lambda_j h_t\left(\Lambda_j n^{-\alpha}\right) \\ &\leq\; \limsup_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \Lambda_j h_t\left(\Lambda_j n^{-\alpha}\right) \;\leq\; E\Lambda \qquad \text{w.p.1.} \end{aligned}$$

Now let $M \to \infty$ (while keeping δ fixed) on the left and use dominated convergence to obtain

$$h_t(\delta) E[\Lambda] \;\leq\; \liminf_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \Lambda_j h_t\left(\Lambda_j n^{-\alpha}\right).$$

Finally let $\delta \to 0$ and use the fact that $h_t(0) = 1$ for each t.

Since in this case $B_n^2 = \sum_{j=1}^{n} \frac{n^{-\alpha}\Lambda_j \mu}{(\Lambda_j n^{-\alpha}+\mu)^2} = \frac{n^{-\alpha+1}}{\mu} \frac{1}{n} \sum_{j=1}^{n} \Lambda_j h_0(\Lambda_j n^{-\alpha})$ from (4.6)

$$\lim_{n\to\infty} n^{\alpha-1} B_n^2 \;=\; \frac{E\Lambda}{\mu} \quad \text{w.p. 1.} \tag{4.7}$$

Then

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{B_n^2}\sum_{j=1}^{n} f_t(\lambda_{jn}) &= \lim_{n\to\infty}\frac{1}{n^{\alpha-1}B_n^2}\frac{1}{n}\sum_{j=1}^{n}\Lambda_j n^{-\alpha}h_t(\Lambda_j n^{-\alpha})\frac{e^{-\mu t}}{\mu}\\ &= \frac{\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}\Lambda_j h_t\left(\Lambda_j n^{-\alpha}\right)}{\lim_{n\to\infty} n^{\alpha-1}B_n^2} = \frac{\mu^{-1}E\Lambda\, e^{-\mu t}}{\mu^{-1}E\Lambda} = e^{-\mu t} \text{ w.p. } 1\end{aligned}$$

and this completes the proof. □

The above proposition shows that Assumptions A2 and A3 of Theorem (3.6) holds for any value of $\alpha > 0$.

Proposition 4.2. *In the above framework, Assumption A1 of Theorem 3.6 holds if and only if* $\alpha \in (0,1)$. *A4 holds for any* $\alpha > 0$.

Proof. From (4.7) there exists $K_1 > 0$ such that for all n sufficiently large $K_1 < n^{\alpha-1}B_n^2$, $B_n^2 > K_1 n^{1-\alpha}$. Thus A1 holds iff $\alpha \in (0,1)$. To see that A4 holds for any $\alpha > 0$ note that

$$\frac{1}{B_n^2}\sum_{j=1}^{n}\Lambda_j n^{-\alpha}\mu = \frac{\mu}{n^{1-\alpha}B_n^2}\frac{1}{n}\sum_{j=1}^{n}\Lambda_j \to \mu^2 \text{ w.p. } 1$$

by the Strong Law of Large Numbers for the $\{\Lambda_j\}$ and (4.7). □

4.4 Two–Dimensional Structure Functions

In this section we will consider portfolios that can be described by two dimensional structure functions. We will suppose that $\{(\Lambda_k, \Pi_k)\}$, $k = 1, 2, \ldots$ is a sequence of i.i.d. elements of $\mathbb{R}_+^2$ such that the random variables Λ and $\frac{\Lambda}{\Pi}$ are integrable. If $(\lambda_k^n, \mu_k^n) = (n^{-\alpha}\Lambda_k, \Pi_k)$ with $\alpha \in (0,1)$,

$$B_n^2 = \sum_{k=1}^{n}\frac{\lambda_k^n\mu_k^n}{\left(\lambda_k^n+\mu_k^n\right)^2} = \sum_{k=1}^{n}\frac{\Lambda_k\Pi_k n^{-\alpha}}{\left(\Lambda_k n^{-\alpha}+\Pi_k\right)^2}$$

and

$$\begin{aligned}\frac{1}{B_n^2}\sum_{k=1}^{n}\frac{\lambda_k^n\mu_k^n}{\left(\lambda_k^n+\mu_k^n\right)^2}e^{-\left(\lambda_k^n+\mu_k^n\right)t} &= \frac{\sum_{k=1}^{n}\frac{\Lambda_k\Pi_k n^{-\alpha}}{\left(\Lambda_k n^{-\alpha}+\Pi_k\right)^2}e^{-\left(\Lambda_k n^{-\alpha}+\Pi_k\right)t}}{\sum_{k=1}^{n}\frac{\Lambda_k\Pi_k n^{-\alpha}}{\left(\Lambda_k n^{-\alpha}+\Pi_k\right)^2}}\\ &\xrightarrow{\text{w.p.1}} E\left[\frac{\Lambda_k}{\Pi_k}e^{-\Pi_k t}\right],\end{aligned} \tag{4.8}$$

the last convergence following from an argument similar to that of Proposition 4.1. One may also check, as in Proposition (4.2) that the assumptions of Theorem (3.6) are satisfied. In fact in this case

$$c(t) := \frac{E\left[\frac{\Lambda}{\Pi}e^{-\Pi|t|}\right]}{E\left[\frac{\Lambda}{\Pi}\right]}, \qquad t \in \mathbb{R} \tag{4.9}$$

and we can see in particular that in the context of Theorem 3.6 $\gamma = E[\Lambda]$. Indeed,

$$\frac{1}{t}[c(0) - c(t)] \;=\; E\left[\frac{\Lambda}{\Pi}\frac{1-e^{-\Pi t}}{t}\right] \;=\; E\left[\Lambda\frac{\int_0^{\Pi t} e^{-u}du}{\Pi t}\right] \;\to\; E[\Lambda] \quad \text{as } t \to 0$$

the last conclusion following by using the Dominated Convergence Theorem, the bound $0 \leq \frac{\int_0^{\Pi t} e^{-u}du}{\Pi t} \leq \frac{\int_0^{\Pi t} 1du}{\Pi t} = 1$, and the assumption $E\Lambda < \infty$. The fact that $c(t)$ is a positive definite function and hence a valid covariance function follows immediately from the positive definiteness of $e^{-|t|}$ and the representation

$$c(t) = \frac{\iint \frac{x}{y} e^{-y|t|} F(dx, dy)}{\iint \frac{x}{y} F(dx, dy)},$$

where $F(dx, dy)$ is the joint distribution function of (Λ, Π). The scaled limiting process is Gaussian with covariance function given by (4.9).

4.4.1 Determination of the scaled limiting covariance function

Here we examine the general case where (Λ_k, Π_k) are not necessarily independent and we will obtain an expression for the limiting covariance (4.9) in terms of their joint Laplace transform. In fact the problem we face here is to determine the expectation $E\left[\frac{X}{Y}e^{-tY}\right]$ in terms of the joint Laplace transform of two given random variables, X, Y, on $\mathbb{R}^+ \times \mathbb{R}^+$. This is described in the following

Lemma 4.3. *If $\varphi(s_1, s_2) = E\left[e^{-s_1X - s_2Y}\right]$ is the joint Laplace transform of the random variables X, Y then*

$$E\left[\frac{X}{Y}e^{-tY}\right] \;=\; -\int_t^\infty \frac{\partial}{\partial s_1}\varphi(0, u)du \tag{4.10}$$

and the scaled covariance function is

$$c(t) \;:=\; \frac{E\left[\frac{X}{Y}e^{-tY}\right]}{E\left[\frac{X}{Y}\right]} \;=\; \frac{\int_t^\infty \frac{\partial}{\partial s_1}\varphi(0, u)du}{\int_0^\infty \frac{\partial}{\partial s_1}\varphi(0, u)du}. \tag{4.11}$$

Proof. We start with the identity

$$\frac{1 - e^{-s_2Y}}{Y} = \int_0^{s_2} e^{-Yx}dx$$

and multiply with e^{-s_1X} to obtain

$$\frac{e^{-s_1X}}{Y} - \frac{e^{-s_1X - s_2Y}}{Y} = \int_0^{s_2} e^{-s_1X - xY}dx.$$

Differentiating with respect to s_1 we obtain

$$-\frac{X}{Y}e^{-s_1X} + \frac{X}{Y}e^{-s_1X - s_2Y} = \int_0^{s_2} \frac{\partial}{\partial s_1}(e^{-s_1X - xY})dx.$$

Taking expectations in the above expression and using Fubini's theorem

$$-E\left[\frac{X}{Y}e^{-s_1X}\right]+E\left[\frac{X}{Y}e^{-s_1X-s_2Y}\right]=\int_0^{s_2}\frac{\partial}{\partial s_1}E(e^{-s_1X-xY})dx.$$

In particular, if we evaluate the above at $s_1=0$, we obtain

$$-E\left[\frac{X}{Y}\right]+E\left[\frac{X}{Y}e^{-s_2Y}\right]=\int_0^{s_2}\frac{\partial}{\partial s_1}\varphi(0,x)dx$$

or

$$E\left[\frac{X}{Y}e^{-s_2Y}\right]=E\left[\frac{X}{Y}\right]+\int_0^{s_2}\frac{\partial}{\partial s_1}\varphi(0,x)dx. \tag{4.12}$$

If in the above expression we let $s_2\to\infty$ we obtain

$$E\left[\frac{X}{Y}\right]=-\int_0^{\infty}\frac{\partial}{\partial s_1}\varphi(0,x)dx. \tag{4.13}$$

Substituting (4.13) in (4.12) we have that

$$E\left[\frac{X}{Y}e^{-s_2Y}\right]=-\int_0^{\infty}\frac{\partial}{\partial s_1}\varphi(0,x)dx+\int_0^{s_2}\frac{\partial}{\partial s_1}\varphi(0,x)dx$$

and hence, setting $s_2=t$, we establish (4.10). □

The above lemma will be used repeatedly in the sequel in order to compute the limiting covariance function for various models.

4.5 Examples of Structure functions

In this section we will consider several examples of portfolios consisting of independent risks that can be thought of as having characteristics (Λ_k,Π_k) arising from various types of structure functions. In all cases, one may see that the integrability assumptions on Λ and $\frac{\Lambda}{\Pi}$ are satisfied.

4.5.1 Independent Gamma-distributed components

This is the simplest case conceptually. We assume that Λ_k and Π_k are independent and that $\Lambda_k\sim\Gamma(\alpha_1,\beta_1)$ and $\Pi_k\sim\Gamma(\alpha_2,\beta_2)$ with $\alpha_2>1$. Then according to the above analysis the limiting Gaussian process has covariance given by

$$\begin{aligned}
E\left[\frac{\Lambda_k}{\Pi_k}e^{-\Pi_k t}\right] &= \frac{\alpha_1}{\beta_1}\int_0^{\infty}\frac{1}{x}e^{-xt}\beta_2\frac{(\beta_2x)^{\alpha_2-1}}{\Gamma(\alpha_2)}e^{-\beta_2x}dx\\
&= \frac{\alpha_1}{\beta_1}\frac{\beta_2}{\Gamma(\alpha_2)}\int_0^{\infty}\frac{(\beta_2x)^{\alpha_2-2}}{\Gamma(\alpha_2-1)}\beta_2e^{-(t+\beta_2)x}\Gamma(\alpha_2-1)dx\\
&= \frac{\alpha_1}{\beta_1}\frac{\beta_2}{\Gamma(\alpha_2)}\Gamma(\alpha_2-1)\frac{\beta_2^{(\alpha_2-1)}}{(\beta_2+t)^{(\alpha_2-1)}}\\
&= \frac{\alpha_1}{\beta_1}\frac{\beta_2^{\alpha_2}}{\alpha_2-1}\ \frac{1}{(\beta_2+t)^{(\alpha_2-1)}}.
\end{aligned}$$

Then the scaled covariance function is

$$c(t) \;=\; \left(\frac{\beta_2}{\beta_2+t}\right)^{\alpha_2-1}. \tag{4.14}$$

Observe that the covariance function decays as a power of t and not exponentially. Thus we can expect the limiting Gaussian process to exhibit "long range dependence".

4.5.2 Multivariate Gamma I Components

Kibble [88] and Moran [110] discussed a symmetrical bivariate gamma distribution with joint characteristic function

$$\frac{1}{((1-it_1)(1-it_2)+\gamma^2 t_1 t_2)^{\alpha}},\quad \alpha>0$$

The corresponding moment generating function is due to Wicksell [147]. Vere-Jones [142] has shown that this distribution is infinitely divisible. The marginal distributions of X and Y are both standard gamma distributions with shape parameter α. The correlation coefficient is equal to ω^2. Setting now $s_j=-it_j\beta_j$, $j=1,2$ we obtain the joint Laplace transform as

$$\varphi(s_1,s_2)=\left(\frac{\beta_1\beta_2}{(\beta_1+s_1)(\beta_2+s_2)-\omega^2 s_1 s_2}\right)^{\alpha}.$$

Hence, in order to use (4.10) we need to compute

$$\frac{\partial}{\partial s_1}\varphi(s_1,s_2)=-(\beta_1\beta_2)^{\alpha}\frac{\alpha}{((\beta_1+s_1)(\beta_2+s_2)-\omega^2 s_1 s_2)^{\alpha+1}}(-\omega^2 s_2+\beta_2+s_2)$$

and, in particular,

$$\begin{aligned}\frac{\partial}{\partial s_1}\varphi(0,s_2) &= -\frac{\alpha(\beta_1\beta_2)^{\alpha}}{(\beta_1(\beta_2+s_2))^{\alpha+1}}(-\omega^2 s_2+\beta_2+s_2)\\ &= -\alpha\left(\frac{\beta_2}{\beta_2+s_2}\right)^{\alpha+1}\frac{s_2(1-\omega^2)+\beta_2}{\beta_1\beta_2}.\end{aligned}$$

So,

$$\begin{aligned}E\left[\frac{X}{Y}e^{-tY}\right] &= \int_t^{\infty}\alpha\left(\frac{\beta_2}{\beta_2+x}\right)^{\alpha+1}\frac{x(1-\omega^2)+\beta_2}{\beta_1\beta_2}dx\\ &= \frac{\alpha\beta_2^{\alpha+1}}{\beta_1\beta_2}\int_t^{\infty}\frac{x(1-\omega^2)+\beta_2}{(\beta_2+x)^{\alpha+1}}dx\\ &= \frac{\alpha\beta_2^{\alpha+1}}{\beta_1\beta_2}\left[\frac{1-\omega^2}{\alpha-1}\frac{1}{(\beta_2+t)^{\alpha-1}}+\frac{\beta_2\omega^2}{\alpha}\frac{1}{(\beta_2+t)^{\alpha}}\right]\end{aligned}$$

(for $\alpha>1$) and thus

$$c(t) \;=\; \frac{\frac{1-\omega^2}{\alpha-1}}{\frac{1-\omega^2}{\alpha-1}+\frac{\omega^2}{\alpha}}\left(\frac{\beta_2}{\beta_2+t}\right)^{\alpha-1}+\frac{\frac{\omega^2}{\alpha}}{\frac{1-\omega^2}{\alpha-1}+\frac{\omega^2}{\alpha}}\left(\frac{\beta_2}{\beta_2+t}\right)^{\alpha}. \tag{4.15}$$

We note that in this case, the resulting covariance function is the mixture of two positive definite functions, $\left(\frac{\beta_2}{\beta_2+|t|}\right)^{\alpha-1}$ and $\left(\frac{\beta_2}{\beta_2+|t|}\right)^{\alpha}$, which decay as powers of t. If the correlation coefficient, ω^2, is equal to 1 then we see that we have a greater rate of decay of the covariance function as opposed to the case where $\omega^2=0$.

4.5.3 Multivariate Gamma II Components

Following David and Fix [34], we use linear combinations of three independent Gamma distributed random variables to obtain two non independent random variables with Gamma marginals. In this case the joint Laplace transform has the form

$$\varphi(s_1, s_2) = \left(\frac{\beta_1}{\beta_1 + s_1}\right)^{\alpha_1} \left(\frac{\beta_2}{\beta_2 + s_2}\right)^{\alpha_2} \left(\frac{\beta_1\beta_2}{\beta_1\beta_2 + \beta_2 s_1 + \beta_1 s_2}\right)^{\alpha_0}$$

where $\alpha_0, \alpha_1, \alpha_2 > 0$. Thus,

$$\begin{aligned}\frac{\partial}{\partial s_1}\varphi(s_1, s_2) &= \left(\frac{\beta_2}{\beta_2 + s_2}\right)^{\alpha_2} \left\{-\alpha_1 \frac{\beta_1^{\alpha_1}}{(\beta_1 + s_1)^{\alpha_1+1}} \left(\frac{\beta_1\beta_2}{\beta_1\beta_2 + \beta_2 s_1 + \beta_1 s_2}\right)^{\alpha_0}\right. \\ &\qquad \left. - \left(\frac{\beta_1}{\beta_1 + s_1}\right)^{\alpha_1} \alpha_0 \frac{(\beta_1\beta_2)^{\alpha_0}}{(\beta_1\beta_2 + \beta_2 s_1 + \beta_1 s_2)^{\alpha_0+1}} \beta_2 \right\}\end{aligned}$$

and

$$\begin{aligned}\frac{\partial}{\partial s_1}\varphi(0, s_2) &= -\left(\frac{\beta_2}{\beta_2 + s_2}\right)^{\alpha_2} \left\{\frac{\alpha_1}{\beta_1} \left(\frac{\beta_2}{\beta_2 + s_2}\right)^{\alpha_0} + \frac{\alpha_0}{\beta_1} \left(\frac{\beta_2}{\beta_2 + s_2}\right)^{\alpha_0+1}\right\} \\ &= -\frac{\alpha_1}{\beta_1} \left(\frac{\beta_2}{\beta_2 + s_2}\right)^{\alpha_2+\alpha_0} - \frac{\alpha_0}{\beta_1} \left(\frac{\beta_2}{\beta_2 + s_2}\right)^{\alpha_0+\alpha_2+1}.\end{aligned}$$

So,

$$\begin{aligned}E\left[\frac{X}{Y}e^{-tY}\right] &= \frac{\alpha_1}{\beta_1}\int_t^\infty \left(\frac{\beta_2}{\beta_2 + x}\right)^{\alpha_2+\alpha_0} dx + \frac{\alpha_0}{\beta_1}\int_t^\infty \left(\frac{\beta_2}{\beta_2 + x}\right)^{\alpha_0+\alpha_2+1} dx \\ &= \frac{\frac{\alpha_1}{\beta_1}\beta_2^{\alpha_2+\alpha_0}}{(\alpha_2 + \alpha_0 - 1)(\beta_2 + t)^{\alpha_2+\alpha_0-1}} + \frac{\frac{\alpha_0}{\beta_1}\beta_2^{\alpha_2+\alpha_0+1}}{(\alpha_2 + \alpha_0)(\beta_2 + t)^{\alpha_2+\alpha_0}}\end{aligned}$$

(for $\alpha_2 + \alpha_0 > 1$).

$$c(t) = \frac{\alpha_1(\alpha_2 + \alpha_0)}{(\alpha_1 + \alpha_0)(\alpha_2 + \alpha_0) - \alpha_0} \left(\frac{\beta_2}{\beta_2 + t}\right)^{\alpha_2+\alpha_0-1} + \frac{\alpha_0(\alpha_2 + \alpha_0) - \alpha_0}{(\alpha_1 + \alpha_0)(\alpha_2 + \alpha_0) - \alpha_0} \left(\frac{\beta_2}{\beta_2 + t}\right)^{\alpha_2+\alpha_0}. \tag{4.16}$$

We note that again the limiting covariance function is the limit of two functions that decay according to inverse powers of t that differ by one. Furthermore, again the rate of decay is higher when the correlation between X and Y is 1 (corresponding to the case $\alpha_1 = \alpha_2 = 0$) and lower when the correlation between X and Y is 0 (corresponding to the case $\alpha_0 = 0$).

4.5.4 Multivariate Pareto Components

The bivariate distribution with joint density function Mardia [105]

$$p_{X,Y}(x, y) = \alpha(\alpha + 1)(\theta_1\theta_2)^{\alpha+1}(\theta_2 x + \theta_1 y - \theta_1\theta_2)^{-(\alpha+2)}, \quad x \geq \theta_1 > 0, \quad y \geq \theta_2 > 0, \quad \alpha > 0$$

may be called a bivariate Pareto distribution of the first kind, since the marginal distributions have density functions

$$\begin{aligned}p_X(x) &= \alpha\theta_1^\alpha x^{-(\alpha+1)} \qquad x \geq \theta_1 > 0, \\ p_Y(y) &= \alpha\theta_2^\alpha y^{-(\alpha+1)} \qquad y \geq \theta_2 > 0.\end{aligned}$$

From these marginal Pareto distributions, we readily have (see Eqs. (20.11a) and (20.11b) in Chapter 20 of Johnson, Kotz, and Balakrishnan [81])

$$\begin{aligned} E[X] &= \frac{\alpha}{\alpha-1}\theta_1 \qquad \alpha > 1, \\ E[Y] &= \frac{\alpha}{\alpha-1}\theta_2 \qquad \alpha > 1. \end{aligned}$$

and

$$\begin{aligned} \mathsf{Var}[X] &= \frac{\alpha}{(\alpha-1)^2(\alpha-2)}\theta_1^2 \qquad \alpha > 2, \\ \mathsf{Var}[Y] &= \frac{\alpha}{(\alpha-1)^2(\alpha-2)}\theta_2^2 \qquad \alpha > 2. \end{aligned}$$

The conditional density function of Y, given $X = x$, is

$$\begin{aligned} p_{Y|X}(y \mid x) &= (\alpha+1)\theta_1(\theta_2 x)^{\alpha+1}(\theta_2 x + \theta_1 y - \theta_1\theta_2)^{-(\alpha+2)} \\ &\qquad \theta_1 > 0\,,\; y \geq \theta_2 > 0\,, \alpha > 0. \end{aligned} \tag{4.17}$$

Here, in view of the intractability of the Laplace transform of the Pareto distribution we will use the direct approach to compute $E\left[\frac{X}{Y}e^{-Yt}\right]$. From (4.17) we find that

$$E[Y \mid X = x] = \theta_2\left(1 + \frac{x}{\alpha\theta_1}\right). \tag{4.18}$$

Using (4.18) we obtain that

$$E\left[\frac{X}{Y}e^{-Yt} \mid Y\right] = \frac{e^{-Yt}}{Y}E[X \mid Y] = \frac{e^{-Yt}}{Y}\,\theta_1\left(1 + \frac{Y}{\theta_2\alpha}\right) = \frac{e^{-Yt}}{Y}\theta_1 + e^{-Yt}\frac{\theta_1}{\theta_2\alpha}$$

Hence,

$$\begin{aligned} E\left[\frac{X}{Y}e^{-Yt}\right] &= \theta_1\int_{\theta_2}^{\infty}\frac{e^{-yt}}{y}\alpha\theta_2^{\alpha}y^{-(\alpha+1)}dy \;+\; \frac{\theta_1}{\theta_2\alpha}\int_{\theta_2}^{\infty}e^{-yt}\alpha\theta_2^{\alpha}y^{-(\alpha+1)}dy \\ &= \alpha\theta_1\theta_2^{\alpha}\int_{\theta_2}^{\infty}\frac{e^{-yt}}{y^{\alpha+2}}dy \;+\; \theta_1\theta_2^{\alpha-1}\int_{\theta_2}^{\infty}\frac{e^{-yt}}{y^{\alpha+1}}dy \end{aligned}$$

Then

$$c(t) = \frac{\alpha\theta_2^{\alpha-1}}{\frac{\alpha}{\alpha+1}+\frac{1}{\alpha}}\int_{\theta_2}^{\infty}\frac{e^{-yt}}{y^{\alpha+2}}dy + \frac{\theta_2^{\alpha}}{\frac{\alpha}{\alpha+1}+\frac{1}{\alpha}}\int_{\theta_2}^{\infty}\frac{e^{-yt}}{y^{\alpha+1}}dy. \tag{4.19}$$

The integrals in the last equation can be evaluated in terms of the incomplete Gamma function. We point out in particular that the rate of decay of the covariance function in this case is *exponential.*

4.5.5 Dirichlet Components

The Dirichlet distribution has wide-ranging applications. One of them is approximating multinomial distribution using a joint Dirichlet density function; see Johnson [79]. This is discussed in some detail in Johnson, Kotz, and Balakrishnan [95, Ch. 35, Sec. 5].

Assume that we have the independent random variables W_0, W_1, W_2 that have a Gamma(θ_i) $(i = 0, 1, 2)$ distribution and define $X = \frac{W_1}{W_0+W_1+W_2}$ and $Y = \frac{W_2}{W_0+W_1+W_2}$. The joint distribution of X, Y is given by the formula

$$f(x,y) \;=\; \frac{\Gamma(\theta_0+\theta_1+\theta_2)}{\Gamma(\theta_0)\Gamma(\theta_1)\Gamma(\theta_2)}(1-x-y)^{\theta_0-1}x^{\theta_1-1}y^{\theta_2-1}, \qquad x \geq 0\,,\; y \geq 0\,,\; x+y \leq 1.$$

and is called the standard Dirichlet distribution with parameters $\theta_0, \theta_1, \theta_2$.

As in the Pareto case we will compute directly $E\left[\frac{X}{Y}e^{-Yt}\right]$ by first calculating $E[X \mid Y]$.

$$E\left[X \mid Y, \sum_{i=0}^{2} W_i\right] \;=\; E\left[\frac{W_1}{\sum_{i=0}^{2} W_i} \mid W_2, \sum_{i=0}^{2} W_i\right].$$

We will prove that $E[W_1 \mid W_1 + W_0] = \frac{\theta_1}{\theta_0+\theta_1}(W_1 + W_0)$. Indeed, the joint density of W_0, W_1 is

$$f_{W_0,W_1}(w_0, w_1) = \frac{1}{\Gamma(\theta_0)\Gamma(\theta_1)} w_0^{\theta_0-1} w_1^{\theta_1-1} e^{-w_0} e^{-w_1}$$

and hence the density of their sum is $f_{W_0+W_1}(s) = \frac{1}{\Gamma(\theta_0+\theta_1)} s^{\theta_0+\theta_1-1}e^{-s}$. Thus

$$\begin{aligned} f_{W_1|W_0+W_1}(w_1 \mid s) \;&=\; \frac{f_{W_1,W_0+W_1}(w_1, s)}{f_{W_0+W_1}(s)} \\ &=\; \frac{[\Gamma(\theta_0)\Gamma(\theta_1)]^{-1}\, w_1^{\theta_1-1}e^{-w_1}(s-w_1)^{\theta_0-1}e^{-(s-w_1)}}{[\Gamma(\theta_0+\theta_1)]^{-1}\, s^{\theta_0+\theta_1-1}e^{-s}} \\ &=\; \frac{\Gamma(\theta_0+\theta_1)}{\Gamma(\theta_0)\Gamma(\theta_1)} \frac{w_1^{\theta_1-1}(s-w_1)^{\theta_0-1}}{s^{\theta_0+\theta_1-1}} = \frac{1}{s}\frac{\Gamma(\theta_0+\theta_1)}{\Gamma(\theta_0)\Gamma(\theta_1)} \left(\frac{w_1}{s}\right)^{\theta_1-1} \left(1-\frac{w_1}{s}\right)^{\theta_0-1}. \end{aligned}$$

So,

$$E[W_1 \mid W_1 + W_0 = s] = s\frac{\theta_1}{\theta_0+\theta_1}.$$

Also

$$E\left[\frac{W_1}{\sum_{i=0}^{2} W_i} \mid W_0, \textstyle\sum_{i=0}^{2} W_i\right] = \frac{\theta_1}{\theta_0+\theta_1}\frac{W_1+W_0}{\sum_{i=0}^{2} W_i}$$

$$E[X \mid Y, \textstyle\sum_{i=0}^{2} W_i] = \frac{\theta_1}{\theta_0+\theta_1}\frac{W_1+W_0}{\sum_{i=0}^{2} W_i} = \frac{\theta_1}{\theta_0+\theta_1}\frac{\sum_{i=0}^{2} W_i - W_2}{\sum_{i=0}^{2} W_i} = \frac{\theta_1}{\theta_0+\theta_1}(1-Y)$$

and hence

$$E[X \mid Y] = \frac{\theta_1}{\theta_0+\theta_1}(1-Y).$$

Thus

$$E\left[\frac{X}{Y}e^{-Yt} \mid Y\right] = \frac{e^{-Yt}}{Y}\frac{\theta_1}{\theta_0+\theta_1}(1-Y) = \frac{\theta_1}{\theta_0+\theta_1}\left(\frac{e^{-Yt}}{Y} - e^{-Yt}\right)$$

and since $Y \sim$Beta$(\theta_2, \theta_1+\theta_0)$ with density given by

$$\begin{aligned} f(y) \;&=\; \frac{\Gamma(\theta_0+\theta_1+\theta_2)}{\Gamma(\theta_0)\Gamma(\theta_1)\Gamma(\theta_2)} y^{\theta_2-1} \int_0^{1-y} x^{\theta_1-1}(1-y-x)^{\theta_0-1}dx \\ &=\; \frac{\Gamma(\theta_0+\theta_1+\theta_2)}{\Gamma(\theta_0)\Gamma(\theta_1)\Gamma(\theta_2)} y^{\theta_2-1}(1-y)^{\theta_1+\theta_0-1} \int_0^{1} x^{\theta_1-1}(1-x)^{\theta_0-1}dx \\ &=\; y^{\theta_2-1}(1-y)^{\theta_1+\theta_0-1}\frac{\Gamma(\theta_0+\theta_1+\theta_2)}{\Gamma(\theta_2)\Gamma(\theta_0+\theta_1)} \end{aligned}$$

we finally have

$$E\left[\frac{X}{Y}e^{-Yt}\right] = \frac{\theta_1}{\theta_0+\theta_1}\frac{\Gamma(\theta_0+\theta_1+\theta_2)}{\Gamma(\theta_2)\Gamma(\theta_0+\theta_1)}\left\{\int_0^1 \frac{e^{-yt}}{y}y^{\theta_2-1}(1-y)^{\theta_1+\theta_0-1}dy - \int_0^1 e^{-yt}y^{\theta_2-1}(1-y)^{\theta_1+\theta_0-1}dy\right\}$$

where the above integral converges under the assumption that $\theta_2 - 2 > -1$ or $\theta_2 > 1$. In this case it is equal to

$$\begin{aligned}
&\frac{\theta_1}{\theta_0+\theta_1}\frac{\Gamma(\theta_0+\theta_1+\theta_2)}{\Gamma(\theta_2)\Gamma(\theta_0+\theta_1)}\int_0^1 e^{-yt}(y^{\theta_2-2}-y^{\theta_2-1})(1-y)^{\theta_1+\theta_0-1}dy \\
&\quad= \frac{\theta_1}{\theta_0+\theta_1}\frac{\Gamma(\theta_0+\theta_1+\theta_2)}{\Gamma(\theta_2)\Gamma(\theta_0+\theta_1)}\int_0^1 e^{-yt}(1-y)^{\theta_1+\theta_0-1}y^{\theta_2-2}(1-y)dy \\
&\quad= \frac{\theta_1}{\theta_0+\theta_1}\frac{\Gamma(\theta_0+\theta_1+\theta_2)}{\Gamma(\theta_2)\Gamma(\theta_0+\theta_1)}\int_0^1 e^{-yt}(1-y)^{\theta_1+\theta_0}y^{\theta_2-2}dy.
\end{aligned}$$

This last integral can be thought of as the Laplace transform of a Beta-distributed random variable Z. Thus we have

$$\begin{aligned}
E\left[\frac{X}{Y}e^{-Yt}\right] &= \frac{\theta_1}{\theta_0+\theta_1}\frac{\Gamma(\theta_0+\theta_1+\theta_2)}{\Gamma(\theta_2)\Gamma(\theta_0+\theta_1)}\frac{\Gamma(\theta_2-1)\Gamma(\theta_0+\theta_1+1)}{\Gamma(\theta_0+\theta_1+\theta_2)}E\left[e^{-tZ}\right] \\
&= \frac{\theta_1}{\theta_0+\theta_1}\frac{\Gamma(\theta_2-1)\Gamma(\theta_0+\theta_1+1)}{\Gamma(\theta_2)\Gamma(\theta_0+\theta_1)}E\left[e^{-tZ}\right] = \frac{\theta_1}{\theta_0+\theta_1}\frac{\theta_0+\theta_1}{\theta_2-1}E\left[e^{-tZ}\right].
\end{aligned}$$

We thus end up with the expression

$$E\left[\frac{X}{Y}e^{-Yt}\right] = \frac{\theta_1}{\theta_2-1}E\left[e^{-tZ}\right] \qquad \text{where } Z \sim \text{Beta}(\theta_2-1, \theta_0+\theta_1+1).$$

Therefore the limiting covariance is given by

$$c(t) \;=\; E[e^{-tZ}]. \tag{4.20}$$

The above expectation can be expressed in terms of a confluent Hypergeometric series. Let

$$f_X(x) = \frac{1}{B(p,q)}x^{p-1}(1-x)^{q-1} \qquad 0 \le x \le 1.$$

It is known that $E\left[e^{tX}\right] = {}_1F_1(p; p+q; t)$ where ${}_1F_1(a;b;x)$ is the confluent hypergeometric function defined by the series

$${}_1F_1(a;b;x) = \sum_{j=0}^{\infty}\frac{a_{(j)}}{b_{(j)}}\frac{x^j}{j!}$$

where $a_{(n)} := a(a+1)\cdots(a+n-1) = \frac{\Gamma(a+n)}{\Gamma(a)}$ is the ascending factorial. In particular $E\left[e^{-tZ}\right]$ decays exponentially fast as $t \to \infty$ provided that $\theta_2 > 2$ while it behaves subexponentially if $1 < \theta_2 \le 2$.

4.5.6 Inverted Dirichlet Distribution

Assume that we have the independent random variables W_0,W_1,W_2 that have a Gamma(θ_i) ($i = 0, 1, 2$) distribution and define $X = \frac{W_1}{W_0}$ and $Y = \frac{W_2}{W_0}$. The joint distribution of X, Y is given by the formula

$$f(x,y) \;=\; \frac{\Gamma(\theta_0+\theta_1+\theta_2)}{\Gamma(\theta_0)\Gamma(\theta_1)\Gamma(\theta_2)}\frac{x^{\theta_1-1}y^{\theta_2-1}}{(1+x+y)^{\theta_0+\theta_1+\theta_2}},\quad x>0\,,\; y>0\,,\; \theta_i>0.$$

and is called the standard inverted Dirichlet distribution with parameters (also see Tiao and Guttman [136]).

It is known (see Johnson, Kotz, and Balakrishnan [95, Ch. 49, Se. 2] that

$$E[X \mid Y] = (1+Y)\frac{\theta_1}{\theta_0+\theta_2-1}.$$

So,

$$E\left[\frac{X}{Y}e^{-Yt}\right] = E\left[\frac{1+Y}{Y}e^{-Yt}\right]\frac{\theta_1}{\theta_0+\theta_2-1}$$

$$\begin{aligned} f(y) \;&=\; \int_0^\infty f(x,y)dx = \frac{\Gamma(\theta_0+\theta_1+\theta_2)}{\Gamma(\theta_0)\Gamma(\theta_1)\Gamma(\theta_2)}y^{\theta_2-1}\int_0^\infty \frac{x^{\theta_1-1}}{(1+x+y)^{\theta_0+\theta_1+\theta_2}}dx \\ &=\; \frac{\Gamma(\theta_0+\theta_1+\theta_2)}{\Gamma(\theta_0)\Gamma(\theta_1)\Gamma(\theta_2)}y^{\theta_2-1}\frac{1}{(1+y)^{\theta_0+\theta_2}}\frac{\Gamma(\theta_1)\Gamma(\theta_0+\theta_2)}{\Gamma(\theta_0+\theta_1+\theta_2)} \;=\; \frac{\Gamma(\theta_0+\theta_2)}{\Gamma(\theta_0)\Gamma(\theta_2)}\frac{y^{\theta_2-1}}{(1+y)^{\theta_0+\theta_2}}. \end{aligned}$$

Using this, we have that

$$\begin{aligned} E\left[\frac{X}{Y}e^{-Yt}\right] \;&=\; \frac{\theta_1}{\theta_0+\theta_2-1}\left[\int_0^\infty \frac{1+y}{y}e^{-yt}\frac{y^{\theta_2-1}}{(1+y)^{\theta_0+\theta_2}}dy\right]\frac{\Gamma(\theta_0+\theta_2)}{\Gamma(\theta_0)\Gamma(\theta_2)} \\ &=\; \frac{\theta_1}{\theta_0+\theta_2-1}\left[\int_0^\infty e^{-yt}\frac{y^{\theta_2-2}}{(1+y)^{\theta_0+\theta_2-1}}dy\right]\frac{\Gamma(\theta_0+\theta_2)}{\Gamma(\theta_0)\Gamma(\theta_2)} \\ &=\; \frac{\theta_1}{\theta_0+\theta_2-1}\left[\int_0^\infty e^{-yt}\left(\frac{y}{1+y}\right)^{\theta_2-2}\left(\frac{1}{1+y}\right)^{\theta_0+1}dy\right]\frac{\Gamma(\theta_0+\theta_2)}{\Gamma(\theta_0)\Gamma(\theta_2)} \end{aligned}$$

and setting $\frac{1}{1+y} = u$ we obtain

$$\begin{aligned} E\left[\frac{X}{Y}e^{-Yt}\right] \;&=\; \frac{\theta_1}{\theta_0+\theta_2-1}\left[\int_0^1 e^{-(\frac{1-u}{u})t}(1-u)^{\theta_2-2}u^{\theta_0-1}du\right]\frac{\Gamma(\theta_0+\theta_2)}{\Gamma(\theta_0)\Gamma(\theta_2)} \\ &=\; \frac{\theta_1}{\theta_0+\theta_2-1}\frac{\Gamma(\theta_0+\theta_2)}{\Gamma(\theta_0)\Gamma(\theta_2)}\frac{\Gamma(\theta_0)\Gamma(\theta_2-1)}{\Gamma(\theta_0+\theta_2-1)}E\left[e^{-(\frac{1-U}{U})t}\right] \\ &=\; \frac{\theta_1}{\theta_2-1}E\left[e^{-(\frac{1-U}{U})t}\right] \end{aligned}$$

where U is Beta(θ_0, θ_2-1) distributed. Hence

$$E\left[\frac{X}{Y}e^{-Yt}\right] = \frac{\theta_1}{\theta_2-1}\frac{1}{B(\theta_0,\theta_2-1)}\int_0^\infty \frac{x^{\theta_2-2}}{(1+x)^{\theta_0+\theta_2-1}}e^{-xt}dx.$$

Therefore, the limiting scaled covariance is given by

$$c(t) \;=\; \frac{1}{B(\theta_0,\theta_2-1)}\int_0^\infty \frac{x^{\theta_2-2}}{(1+x)^{\theta_0+\theta_2-1}}e^{-xt}dx. \tag{4.21}$$

5. A CLASS OF PROCESSES WITH INFINITELY DIVISIBLE MARGINALS AND ORNSTEIN–UHLENBECK-TYPE COVARIANCE STRUCTURE

5.1 Introduction

The classical Ornstein–Uhlenbeck process is the solution of the stochastic differential equation

$$dX_t = -aX_t dt + dW_t \tag{5.1}$$

where $\{W_t; t \geq 0\}$ is standard Brownian motion and X_0 is independent of $\{W_t; t \geq 0\}$. If $a > 0$ then the process has a stationary version $\{X_t; t \geq 0\}$ under an appropriate choice of X_0.

One particularly fruitful generalization of the classical Ornstein–Uhlenbeck process consists in replacing the driving Brownian Motion term in (5.1) by a Lévy process.

$$dX_t = -aX_t dt + dZ_t \tag{5.2}$$

Assuming $\{Z_t\}$ to be a Lévy process with bounded variation, e.g. a spectrally positive process the integral equation corresponding to the above equation has the solution

$$X_t \;=\; e^{-at}X_0 + \int_0^t e^{-a(t-s)}dZ_s. \tag{5.3}$$

where the above integral, under these assumptions, is defined path-wise. For the definition in the general case see [123] and [5]. The study of such processes and their applications in many areas, most notably mathematical finance, has attracted a great deal of attention over the last two decades. We refer the reader to Sato [123] for background on Generalized Ornstein–Uhlenbeck (GOU) processes and to Barndorff–Nielsen and Shephard [15] for applications in finance and economics. We note in particular that the class of GOU processes is intimately related to self–decomposable distributions (see [148] and [84]).

In this chapter we propose an alternative generalization of the stationary, Wiener process driven, Ornstein–Uhlenbeck process which is intimately connected to coverage processes and limits of superpositions of ON/OFF processes.

5.2 Towards a family of processes which generalizes the covariance structure of the classical Ornstein-Uhlenbeck process

The stationary version of the process defined by (5.1) is a zero mean Gaussian process $\{X_t\}$ and the joint characteristic function of $(X_{t_1}, X_{t_2}, \ldots, X_{t_n})$ for $0 \le t_1 < t_2 < \cdots < t_n$ is given by

$$E \exp\left(i \sum_{i=1}^{n} \theta_i X_{t_i} \right) = \exp\left(-\frac{1}{4a} \sum_{i=1}^{n} \theta_i^2 \; - \; \frac{1}{2a} \sum_{1 \le i < j \le n} \theta_i \theta_j e^{-a(t_j - t_i)} \right). \tag{5.4}$$

(If we choose $a = 1/2$ then the marginal distribution of the process is standard normal.) The key observation is to rewrite the right hand side of (5.4) in the form

$$\exp\Big(-\frac{1}{4a} \sum_{1 \le i \le j \le n} (\theta_i + \cdots + \theta_j)^2 \Big[\Big(1 - e^{-a(t_j - t_{i-1})} \Big) - \Big(1 - e^{-a(t_j - t_i)} \Big) \tag{5.5}$$

$$- \Big(1 - e^{-a(t_{j+1} - t_{i-1})} \Big) + \Big(1 - e^{-a(t_{j+1} - t_i)} \Big) \Big] \Big).$$

In the above expression t_0 is to be interpreted as $-\infty$ whereas t_{n+1} as $+\infty$. This convention will be important throughout. The equality of (5.4) and (5.5) is established in the following

Proposition 5.1. *Expression (5.5) is equal to the right hand side of (5.4).*

Proof. To show this we start with the following observation. Let $\theta^\top := (\theta_1, \ldots, \theta_n)$ and let $B = [b_{ij}]$, $i, j = 1, \ldots, n$ be a $n \times n$ symmetric matrix. The quadratic form $\theta^\top B \theta$ can always be written uniquely in the form

$$\sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} \theta_i \theta_j = \sum_{1 \le i \le j \le n} (\theta_i + \cdots + \theta_j)^2 a_{ij} \tag{5.6}$$

where the a_{ij} and b_{kl} are uniquely determined from each other by means of the relationships

$$b_{kl} = \sum_{1 \le i \le k \le l \le j \le n} a_{ij}, \text{ for } k \le l \tag{5.7}$$

and

$$a_{ij} = b_{ij} - b_{i,j+1} - b_{i-1,j} + b_{i-1,j+1}, \quad \text{for } 1 \le i \le j \le n, \tag{5.8}$$

where, in the last relationship we have taken $b_{0,j} = b_{i,n+1} = 0$.

Now use the identity (5.6) with $b_{ij} = e^{-a(t_j - t_i)}$, $1 \le i \le j \le n$, and

$$\begin{aligned} a_{ij} &= e^{-a(t_j - t_i)} - e^{-a(t_{j+1} - t_i)} - e^{-a(t_j - t_{i-1})} + e^{-a(t_{j+1} - t_{i-1})} \\ &= \Big(1 - e^{-a(t_j - t_{i-1})} \Big) - \Big(1 - e^{-a(t_j - t_i)} \Big) - \Big(1 - e^{-a(t_{j+1} - t_{i-1})} \Big) + \Big(1 - e^{-a(t_{j+1} - t_i)} \Big). \end{aligned}$$

□

We propose then to replace $\frac{1}{4a}\theta^2$ by the characteristic exponent $\psi(\theta)$ of an infinitely divisible characteristic function, $\phi(\theta) = e^{\psi(\theta)}$, and the exponential distribution function $1 - e^{-at}$ which appears in (5.5) by a general *concave* distribution function H on $\mathbb{R}^+$ as shown in the sequel. This class of processes is defined in the sequel and some of their properties are derived.

5.3 Infinitely divisible generalized coverage processes with a generalized OU covariance structure

Suppose that $f : [0, \infty) \to \mathbb{R}$ is *concave*, i.e. for all $x, y \in [0, \infty)$ and $\lambda \in [0, 1]$

$$f(\lambda x + (1 - \lambda)y) \ \geq \ \lambda f(x) + (1 - \lambda) f(y).$$

Concave functions have the *decreasing increments* property, namely

$$f(x + h) - f(x) \ \geq \ f(y + h) - f(y) \quad \text{for } 0 \leq x \leq y, \quad h > 0. \tag{5.9}$$

(See for instance [113]). As a consequence of this property we can see that the concave function f is *supermodular* i.e. if $0 \leq x_1 < x_2 < x_3 < x_4$ the following inequality holds

$$f(x_3 - x_1) - f(x_3 - x_2) - f(x_4 - x_1) + f(x_4 - x_2) \ \geq \ 0. \tag{5.10}$$

Indeed, to see that (5.9) implies (5.10) it suffices to take $x := x_3 - x_2$, $y := x_3 - x_1$ and $h := x_4 - x_3$.

The following theorem defines a class of *Generalized Coverage Processes with Infinitely Divisible Marginals* (GCID process). This class includes the $M/GI/\infty$ process and the classic Ornstein-Uhlenbeck process.

Theorem 5.2 (GCID)**.** *Suppose that $\phi(\theta) = e^{\psi(\theta)}$ is the characteristic function of an infinitely divisible law and denote by X a random variable with the corresponding distribution, i.e. $\phi(\theta) = Ee^{i\theta X}$. Suppose that H is a probability distribution function on $[0, \infty)$ such that $H(0) = 0$ and H is concave. Then, for $n \in \mathbb{N}$ and $\theta_i, t_i \in \mathbb{R}$, $i = 1, 2, \ldots, n$, $t_1 \leq t_2 \leq \cdots \leq t_n$, the family of functions $\phi_n(\theta_1, \ldots, \theta_n; t_1, \ldots, t_n)$ given by*

$$\begin{aligned} &\log \phi_n(\theta_1, \ldots, \theta_n; t_1, \ldots, t_n) \\ &\quad = \sum_{1 \leq i \leq j \leq n} \psi(\theta_i + \cdots + \theta_j) \left[H(t_j - t_{i-1}) - H(t_j - t_i) - H(t_{j+1} - t_{i-1}) + H(t_{j+1} - t_i) \right] \end{aligned} \tag{5.11}$$

are characteristic functions of consistent *finite–dimensional distributions. (For each fixed n in the above expression t_0 is to be interpreted as $-\infty$ and t_{n+1} as $+\infty$.) The family of characteristic functions defined in (5.11) corresponds to the finite–dimensional distributions of a stationary process $\{X_t; t \in \mathbb{R}\}$ with marginal distribution $Ee^{i\theta X_t} = \phi(\theta)$ and covariance*

$$\mathsf{Cov}(X_t, X_{t+h}) = \psi''(0) \left[1 - H(|h|) \right]. \tag{5.12}$$

Remark 5.3. *We point out that, for $n = 1$, (5.11) contains only one term, namely*

$$\log \phi_1(\theta_1; t_1) = \psi(\theta_1),$$

while for $n = 2$, after simple calculations, and taking into account the conventions regarding t_0 and t_{n+1}, we obtain

$$\log \phi_2(\theta_1, \theta_2; t_1, t_2) = \psi(\theta_1) H(t_2 - t_1) + \psi(\theta_1 + \theta_2) \left[1 - H(t_2 - t_1) \right] + \psi(\theta_2) H(t_2 - t_1). \tag{5.13}$$

Proof. First, we need to show that the right hand side of (5.11) is in fact the logarithm of a characteristic function of a probability distribution on $\mathbb{R}^n$. To this end, using the Cramér–Wold Theorem (see [18, p. 383]), it is enough to show that

$$\phi_n(\theta_1 u, \ldots, \theta_n u; t_1, \ldots, t_n) =: \Phi(u)$$

is the characteristic function of a probability distribution in $\mathbb{R}$ for any $(\theta_1, \ldots, \theta_n) \in \mathbb{R}^n$. Note that,

$$\log \Phi(u) = \sum_{1 \le i \le j \le n} \psi(u(\theta_i + \cdots + \theta_j)) a_{ij}$$

where

$$a_{ij} := [H(t_j - t_{i-1}) - H(t_j - t_i) - H(t_{j+1} - t_{i-1}) + H(t_{j+1} - t_i)] \ge 0$$

since H is concave and therefore supermodular. Therefore,

$$\Phi(u) = \prod_{1 \le i \le j \le n} e^{\psi(u(\theta_i + \cdots + \theta_j)) a_{ij}}.$$

However, $e^{\psi(u(\theta_i + \cdots + \theta_j))}$ is the characteristic function of an infinitely divisible law since $e^{\psi(\theta)}$ is an infinitely divisible characteristic function, and $e^{\psi(u(\theta_i + \cdots + \theta_j)) a_{ij}}$ is also an infinitely divisible characteristic function which means that $\Phi(u)$ is the characteristic function of a probability distribution in $\mathbb{R}$ as a product of characteristic functions. The stationarity of the process $\{X_t; t \in \mathbb{R}\}$ follows from the fact that the corresponding finite dimensional distributions defined by (5.11) are stationary. Indeed,

$$\log \phi_n(\theta_1, \ldots, \theta_n; t_1 + \tau, \ldots, t_n + \tau) = \log \phi_n(\theta_1, \ldots, \theta_n; t_1, \ldots, t_n),$$

for any $\tau \in \mathbb{R}$, as can be readily verified by (5.11).

In order to establish (5.12) we first note that, in view of the stationarity of $\{X_t; t \in \mathbb{R}\}$,

$$\mathsf{Cov}(X_t, X_{t+h}) = \mathsf{Var}(X_t) - \frac{1}{2} E(X_t - X_{t+h})^2.$$

The variance term above is $\mathsf{Var}(X_t) = \psi''(0)$. To evaluate the second term we set $\theta_2 = \theta$, $\theta_1 = -\theta$ and $h = t_2 - t_1$ in (5.13) to obtain

$$\begin{aligned} \phi_2(-\theta, \theta; t, t+h) &= E\left[e^{i\theta(X_{t+h} - X_t)}\right] = \exp\left\{\psi(0)[1 - H(h)] + \psi(\theta)H(h) + \psi(-\theta)H(h)\right\} \\ &= \exp\left\{H(h)[\psi(\theta) + \psi(-\theta)]\right\} \end{aligned} \tag{5.14}$$

where in the last equation we have taken into account that $\psi(0) = 0$. We have also assumed that $h > 0$. The situation where h is negative can be dealt with similarly. Differentiating twice this expression and evaluating at $\theta = 0$ gives

$$\left.\frac{d^2}{d\theta^2} \phi_2(-\theta, \theta; t, t+h)\right|_{\theta=0} = E(X_t - X_{t+h})^2 = 2H(h)\psi''(0). \tag{5.15}$$

Thus, from the above and the stationarity of the process X_t we obtain that

$$\mathsf{Cov}(X_t, X_{t+h}) = \mathsf{Var}(X_t) - \frac{1}{2} E(X_t - X_{t+h})^2 = \psi''(0)[1 - H(|h|)]$$

which establishes (5.12). □

Remark 5.4. *A special case of particular importance arises when* $H(t) = 1 - e^{-\mu t}$ *for some* $\mu > 0$. *In that case*

$$\log \phi_n(\theta_1, \ldots, \theta_n) = \sum_{1 \le i \le j \le n} \psi(\theta_i + \cdots + \theta_j)(1 - e^{-\mu(t_i - t_{i-1})})e^{-\mu(t_j - t_i)}(1 - e^{-\mu(t_{j+1} - t_j)}). \quad (5.16)$$

Note the correlation structure

$$(1 - e^{-\mu(t_i - t_{i-1})})e^{-\mu(t_j - t_i)}(1 - e^{-\mu(t_{j+1} - t_j)})$$

which corresponds to that of an Ornstein-Uhlenbeck process. In particular, the covariance function is given by $\mathsf{Cov}(X_t, X_{t+h}) = \mathsf{Var}(X_0)e^{-\mu|h|}$.

Lemma 5.5. *The following identity holds for any* n:

$$\sum_{1 \le i \le j \le n} (\theta_i + \cdots + \theta_j)\left[H(t_j - t_{i-1}) - H(t_j - t_i) - H(t_{j+1} - t_{i-1}) + H(t_{j+1} - t_i)\right] = \theta_1 + \cdots + \theta_n. \quad (5.17)$$

Therefore, if $\psi(\theta)$ *contains a drift term, i.e. if* $\psi(\theta) = i\mu\theta + \psi_0(\theta)$ *where*

$$\psi_0(\theta) := \frac{1}{2}\alpha\theta^2 + \int_{\mathbb{R}} (e^{i\theta x} - 1 - i\theta h(x))\nu(dx)$$

with $h(x) = x\mathbf{1}(|x| < 1) + \mathbf{1}(x \ge 1) - \mathbf{1}(x \le -1)$, $\alpha \ge 0$ *and the Lévy measure satisfies* $\nu[x, \infty) = \nu(-\infty, -x]$ *for all* $x > 0$, *then (5.11) becomes*

$$\begin{aligned}\log \phi_n(\theta_1, \ldots, \theta_n) &= \mu(\theta_1 + \cdots + \theta_n) \\ &+ \sum_{1 \le i \le j \le n} \psi_0(\theta_i + \cdots + \theta_j)\left[H(t_j - t_{i-1}) - H(t_j - t_i) - H(t_{j+1} - t_{i-1}) + H(t_{j+1} - t_i)\right].\end{aligned}$$

Proof. Fix i, j in the set $\{1, 2, \ldots, n\}$. (We don't exclude the possibility $i = j$.) Pick an integer r such that $i \le r \le j$. The left hand side of (5.17) can be written as

$$\sum_{r=1}^{n} \theta_r \sum_{1 \le i \le r \le j \le n} \left[H(t_j - t_{i-1}) - H(t_j - t_i) - H(t_{j+1} - t_{i-1}) + H(t_{j+1} - t_i)\right] \quad (5.18)$$

and the inner sum in the left hand side of the above expression can be written as

$$\sum_{1 \le i \le r} \left\{ \sum_{r \le j \le n} \left[H(t_j - t_{i-1}) - H(t_{j+1} - t_{i-1})\right] + \sum_{r \le j \le n} \left[H(t_{j+1} - t_i) - H(t_j - t_i)\right] \right\} \quad (5.19)$$

The two inner sums in (5.19) are telescopic and reduce to

$$H(t_r - t_{i-1}) - H(t_{n+1} - t_{i-1}) = H(t_r - t_{i-1}) - H(\infty) = H(t_r - t_{i-1}) - 1$$

and

$$H(t_{n+1} - t_i) - H(t_r - t_i) = H(\infty) - H(t_r - t_i) = 1 - H(t_r - t_i)$$

respectively. In the above we have taken into account the convention $t_{n+1} = +\infty$. Thus (5.19) becomes

$$\begin{aligned}\sum_{1 \le i \le r} \{H(t_r - t_{i-1}) - H(t_r - t_i)\} &= H(t_r - t_0) - H(t_r - t_r) = H(t_r - (-\infty)) - H(0) \\ &= H(\infty) - 0 = 1 \quad (5.20)\end{aligned}$$

and hence (5.18) reduces to $\sum_{r=1}^{n} \theta_r$ and the proof is complete. □

We end the section by considering a subclass of GCID processes which is related to superpositions of limits of independent ON/OFF markovian processes. In this subclass, the Lévy marginal distribution is spectrally positive and the correlation structure is that of the classical Ornstein-Uhlenbeck process.

Definition 5.6 (ESPC)**.** *Consider the subclass of GCID processes which we will term* Exponential Spectrally Positive Coverage Processes *(ESPC),* $\{X_t\}$*, such that*

(a) $\{X_t\}$ *is stationary.*

(b) For all $t \in \mathbb{R}$ X_t *has spectrally positive Lévy distribution with*

$$\psi(\theta) := \log \mathbb{E}\left[e^{i\theta X_t}\right] = \int_0^\infty (e^{i\theta x} - 1)\nu(dx)$$

where $\int_0^\infty x\nu(dx) < \infty$ *(but* $\int_0^\infty \nu(dx)$ *may or may not be finite).*

(c) The finite dimensional distributions are given by

$$\log \mathbb{E}\left[e^{i(\theta_1 X_{t_1} + \theta_2 X_{t_2} + \cdots + \theta_n X_{t_n})}\right] \tag{5.21}$$
$$= \sum_{1 \le j \le k \le n} \int_0^\infty \left(e^{i(\theta_j + \cdots + \theta_k)x} - 1\right) \nu(dx)\,(1 - e^{-\mu(t_j - t_{j-1})})e^{-\mu(t_k - t_j)}(1 - e^{-\mu(t_{k+1} - t_k)}).$$

The fact that (5.21) indeed defines a consistent familty of distributions is a direct consequence of Theroem 5.2. Note that the mean value of the above process is given by $\mathbb{E}X_t = \int_0^\infty x\nu(dx)$, its variance by $\mathsf{Var}(X_t) = \int_0^\infty x^2\nu(dx)$, and its covariance by $\mathsf{Cov}(X_0, X_{t_1}) = e^{-\mu t_1}\int_0^\infty x^2\nu(dx)$.

5.4 Processes with Poisson and compound Poisson marginals. The $M/G/\infty$ process and its generalizations

In this section we consider processes that fall into the above framework that arise from infinite server queues with Poisson arrivals and general service times. In the first subsection we consider the number of customers in the system and thus obtain a stationary process with Poisson marginals, while in the second we will introduce "marked" $M/G/\infty$ queues where to each customer we associate a real–valued mark. In this fashion we obtain stationary processes with compound Poisson marginal distributions.

5.4.1 The number of customers in the system

Let X_t denote the number of customers at time t in an $M/G/\infty$ queue which has Poisson arrivals with rate λ and service times with distribution G. We will use $\overline{G} := 1 - G$ to designate the corresponding survival function. Let $-\infty = t_0 < t_1 < t_2 < \cdots < t_n < t_{n+1} = +\infty$ and denote by

$$\phi_n(\theta_1, \ldots, \theta_n; t_1, \ldots, t_n) := E[e^{i\theta_1 X_{t_1} + i\theta_2 X_{t_2} + \cdots + i\theta_n X_{t_n}}]$$

the joint characteristic function of $(X_{t_1}, \ldots, X_{t_n})$. Also, for $1 \le i \le j \le n$, denote by Y_{ij} the number of customers who arrived in the system in the time interval $(t_{i-1}, t_i]$ and left in the time interval $(t_j, t_{j+1}]$. Also, let $A_{ij} := \{(x,t) : t_j < x + t \le t_{j+1}; x \ge 0; t_{i-1} < t \le t_i\}$ and

$$\mu(A_{ij}) = \iint_{A_{ij}} \lambda G(dx)dt.$$

Then the collection of random variables $\{Y_{ij}; 1 \le i \le j \le n\}$ are independent, Poisson distributed, with parameters $\mu(A_{ij})$. Furthermore,

$$X_{t_i} = \sum_{j=i}^{n} \sum_{k=1}^{i} Y_{kj}$$

and

$$\begin{aligned}
\phi_n(\theta_1, \ldots, \theta_n; t_1, \ldots, t_n) &= E\left[\prod_{1\le i\le j\le n} e^{i(\theta_i+\cdots+\theta_j)Y_{ij}}\right] = \prod_{1\le i\le j\le n} E\left[e^{i(\theta_i+\cdots+\theta_j)Y_{ij}}\right] \\
&= \prod_{1\le i\le j\le n} e^{-\mu(A_{ij})\left(1-e^{i(\theta_i+\cdots+\theta_j)}\right)} \\
&= e^{-\sum_{1\le i\le j\le n}\mu(A_{ij})\left(1-e^{i(\theta_i+\cdots+\theta_j)}\right)} \qquad (5.22)
\end{aligned}$$

where

$$\begin{aligned}
\mu(A_{ij}) &= \lambda \int_{t=t_{i-1}}^{t_i} \left[\int_{x=t_j-t}^{t_{j+1}-t} dG(x)\right] dt \\
&= \lambda m \left[\frac{1}{m}\int_{t=t_{i-1}}^{t_i} \overline{G}(t_j - t)dt - \frac{1}{m}\int_{t=t_{i-1}}^{t_i} \overline{G}(t_{j+1} - t)dt\right]
\end{aligned}$$

or

$$\mu(A_{ij}) = \rho\left[G_I(t_j - t_{i-1}) - G_I(t_j - t_i) - G_I(t_{j+1} - t_{i-1}) + G_I(t_{j+1} - t_i)\right] \qquad (5.23)$$

where, in the above expression, $\rho = \lambda m$ and we have used the integrated tail distribution given by

$$G_I(x) := \frac{1}{m}\int_0^x \overline{G}(y)dy. \qquad (5.24)$$

Therefore this is also a GCID process with $H(t) = G_I(t)$ and $\psi(\theta) = \rho\left(e^{i\theta} - 1\right)$. The supermodularity of $G_I(t)$ follows from its concavity which, in turn, is a consequence of the fact that it is an integral of an non–increasing function. The marginal distribution of X_t is of course Poisson with mean ρ.

Of particular importance is the case where $G(y) = 1 - e^{-\mu y}$. The resulting process is of the ESPC type with joint characteristic function given by

$$\begin{aligned}
&\log \phi_n(\theta_1, \ldots, \theta_n; t_1, \ldots, t_n) \qquad (5.25) \\
&\quad = \sum_{1\le i\le j\le n} \rho\left(e^{i(\theta_i+\cdots+\theta_j)} - 1\right)\left(1 - e^{-\mu(t_i - t_{i-1})}\right) e^{-\mu(t_j - t_i)}\left(1 - e^{-\mu(t_{j+1} - t_j)}\right).
\end{aligned}$$

This corresponds to the $M/M/\infty$ case and is a special case of and ESPC process (5.21) with $\psi(\theta) = \rho\left(e^{i\theta} - 1\right)$.

5.4.2 The marked $M/G/\infty$ system

The process in the above section can be generalized by considering a marked $M/G/\infty$ queue. Suppose that $\{T_n; n \in \mathbb{Z}\}$ is a Poisson process with intensity λ defined on the real line and $\{(\sigma_n, \xi_n)\,; n \in \mathbb{Z}\}$ a collection of independent, identically distributed elements of $\mathbb{R}^2$. The first component, σ_n, is to be thought of as the service requirement of the n-th customer who arrives in the system (at time T_n), while the second is a real-valued mark corresponding to some attribute of the customer. Thus, with the above notation, we consider the process

$$X_t \;=\; \sum_{n\in\mathbb{Z}} \mathbf{1}\left(T_n \le t \le T_n + \sigma_n\right) \xi_n.$$

In view of the above assumptions the marginal distribution is

$$E\left[e^{i\theta X_t}\right] = e^{\rho[\chi(\theta)-1]}$$

where

$$\chi(\theta) := E e^{i\theta\xi}. \tag{5.26}$$

The joint characteristic functions are given by the following generalization of (5.22)

$$\phi_n(\theta_1, \ldots, \theta_n; t_1, \ldots, t_n) = e^{-\sum_{1\le i\le j\le n} \mu(A_{ij})(1-\chi(\theta_i+\cdots+\theta_j))}. \tag{5.27}$$

where the $\mu(A_{ij})$ are given by (5.23). Thus, (5.22) corresponds to (5.11) with $H(t) = G_I(t)$ given by (5.24) and $\psi(\theta) = \rho[\chi(\theta) - 1]$ given by (5.26). This corresponds to a compound Poisson marginal.

5.5 Triangular arrays of ON/OFF sources and limiting theorems

In this section we give results that show how the class of processes defined by (5.11) can arise as a limit of the superposition of simple ON/OFF processes. Consider for instance a risk portfolio consisting of a large number of contracts, say n, each of which generates claims. We will assume that the jth contract does not generate a claim over an exponential period with rate λ_j and then becomes active for an independent period of time, exponentially distributed, with rate μ_j. When the source is ON, it generates fluid at rate r_j, whereas when it is OFF it does not generate any fluid. We model this as a two–state, continuous time Markov process $\{\zeta_t\}$ with state space $\{0, r\}$ and we consider the effect of the superposition of a large number of such sources, each generating claims very infrequently. For simplicity we will assume here that $\mu_j = \mu$, the same for all sources.

Let us introduce a triangular array of independent ON/OFF processes, $\{\zeta_{jn}(t); t \ge 0\}_{j=1\ldots n}^{n=1,2,\cdots}$ with OFF rates $\{\lambda_{jn}\}_{j=1\ldots n}^{n=1,2,\cdots}$, on rates equal to μ for all j, n, and fluid intensity rates $\{r_{jn}\}_{j=1\ldots n}^{n=1,2,\cdots}$. For simplicity we will assume the stationarity of the elements of the family, $\{\zeta_{jn}(t)\}$, under the probability measure P.

Denote the Dirac measure by $\delta_x(A) = 1$ if $x \in A$ and 0 otherwise for every Borel $A \subset \mathbb{R}$. We will describe the limiting behavior of the scaled input process by considering the corresponding

behavior of the sequence of empirical measures

$$\nu_n(\cdot) = \sum_{j=1}^{n} \lambda_{jn} \delta_{r_{jn}}(\cdot). \tag{5.28}$$

We will show that, when the sequence of measures $\{\nu_n\}$ converges weakly to a measure ν on $(0, \infty)$, under some additional assumptions, the sequence of processes X_n converges weakly to a limit process X of the GCID type defined in Theorem 5.2.

Suppose that the sequence of empirical measures $\{\nu_n\}$ converges weakly to the ($\sigma-$finite) measure ν which satisfies the condition

$$(C.1) \qquad \int_0^\infty x^q \nu(dx) \; < \; \infty \quad q = 1, 2,$$

and in addition the triangular array $\{\lambda_{jn}, r_{jn}\}$ satisfies the conditions

$$(C.2) \qquad \max_{1 \le j \le n} |\lambda_{jn}| \to 0 \quad \text{when } n \to \infty,$$

$$(C.3) \qquad \sup_{n \in \mathbb{N}} \sum_{j=1}^{n} \lambda_{jn} r_{jn}^p \le C, \quad p \in [1, 4] \text{ for some } C > 0.$$

Then the following holds.

Theorem 5.7. *Suppose that, for the array of stationary processes $\{\zeta_{jn}\}$, the sequence of measures $\{\nu_n\}$ converges weakly to a σ–finite measure ν on all subsets of $\mathbb{R}^+$ of the form $[x, \infty)$, $x > 0$, and satisfies conditions C.1, C.2, C.3. Then*

$$X_t^n := \sum_{i=1}^{n} \zeta_{jn}(t) \xrightarrow{d} X_t \tag{5.29}$$

where $\{X_t; t \ge 0\}$ is an ESPC process of the above type with joint distribution given by

$$\log E\left[e^{i(\theta_1 X_{t_1} + \theta_2 X_{t_2} + \cdots + \theta_n X_{t_n})}\right] \tag{5.30}$$
$$= \frac{1}{\mu} \sum_{1 \le j \le k \le n} \int_0^\infty \nu(dx) \left(e^{i(\theta_j + \cdots + \theta_k)x} - 1\right) \left(1 - e^{-\mu(t_j - t_{j-1})}\right) e^{-\mu(t_k - t_j)} \left(1 - e^{-\mu(t_{k+1} - t_k)}\right).$$

Proof. Because of the complexity of the proof we will split it into three parts.

Part 1. Convergence of the marginal distributions.

Set $\xi_{jn}(t) = \zeta_{jn}(t)/r_{jn}$ (so that the processes $\xi_{jn}(t)$ take values in the set $\{0, 1\}$). Then

$$X_t^n := \sum_{j=1}^{n} \zeta_{jn}(t) = \sum_{j=1}^{n} r_{jn} \xi_{jn}(t) \tag{5.31}$$

and we have

$$\begin{aligned}\mathbb{E}[e^{i\theta X_t^n}] &= \mathbb{E}e^{i\theta\sum_{j=1}^n r_{jn}\xi_{jn}} = \prod_{j=1}^n \mathbb{E}[e^{i\theta r_{jn}\xi_{jn}}] = \prod_{j=1}^n \mathbb{E}\left(1+\xi_{jn}(t)\left(e^{i\theta r_{jn}}-1\right)\right)\\ &= \prod_{j=1}^n \left(1+\frac{\lambda_{jn}}{\mu+\lambda_{jn}}\left(e^{i\theta r_{jn}}-1\right)\right).\end{aligned}$$

Since $|\log(1+x)-x| \le K_1|x|^2$ for some $K_1>0$ when $|x|\le 1/2$, it follows that

$$\begin{aligned}\left|\log\mathbb{E}[e^{i\theta X_t^n}] - \sum_{j=1}^n \frac{\lambda_{jn}}{\mu}\left(e^{i\theta r_{jn}}-1\right)\right| &\le \left|\sum_{j=1}^n \log\left(1+\frac{\lambda_{jn}}{\mu+\lambda_{jn}}\left(e^{i\theta r_{jn}}-1\right)\right) - \frac{\lambda_{jn}}{\mu}\left(e^{i\theta r_{jn}}-1\right)\right|\\ &\le \sum_{j=1}^n K_1\left|\frac{\lambda_{jn}}{\lambda_{jn}+\mu}(e^{i\theta r_{jn}}-1)\right|^2 + \left|\frac{\lambda_{jn}^2}{\mu(\lambda_{jn}+\mu)}(e^{i\theta r_{jn}}-1)\right|. \qquad (5.32)\end{aligned}$$

Because of the inequality $\left|e^{i\theta r_{jn}}-1\right| \le |\theta| r_{jn}$ the sum above is bounded by

$$\sum_{j=1}^n\left(\frac{K_1\theta^2}{\mu}\lambda_{jn}^2 r_{jn}^2 + \frac{K_2|\theta|}{\mu^2}\lambda_{jn}^2 r_{jn}\right) \le \left(\max_{1\le j\le n}\lambda_{jn}\right)\left(\frac{K_1\theta^2}{\mu}\sum_{j=1}^n \lambda_{jn}r_{jn}^2 + \frac{K_2|\theta|}{\mu^2}\sum_{j=1}^n\lambda_{jn}r_{jn}\right). \qquad (5.33)$$

Fix $\theta_0>0$. By Assumption, C.C it follows that the sequences $\sum_{j=1}^n \lambda_{jn}r_{jn}$ and $\sum_{j=1}^n\lambda_{jn}r_{jn}^2$ are bounded above by C and hence, taking into account Assumption, C.2, for any θ such that $|\theta|<\theta_0$, the right hand side of (5.33) goes to zero as $n\to\infty$. Hence, since $e^{i\theta x}-1$ is a continuous and bounded function of x, by Helly's second theorem,

$$\lim_{n\to\infty}\sum_{j=1}^n \frac{\lambda_{jn}}{\mu}\left(e^{i\theta r_{jn}}-1\right) = \frac{1}{\mu}\lim_{n\to\infty}\int_0^\infty (e^{i\theta x}-1)\nu_n(dx) = \frac{1}{\mu}\int_0^\infty (e^{i\theta x}-1)\nu(dx). \qquad (5.34)$$

Thus, for any $\theta\in\mathbb{R}$,

$$\lim_{n\to\infty}\log\mathbb{E}[e^{i\theta X_t^n}] = \frac{1}{\mu}\int_0^\infty\left(e^{i\theta x}-1\right)\nu(dx). \qquad (5.35)$$

Part 2. Convergence of the finite dimensional distributions.

Next we will determine the joint distribution of $X_t^n = \sum_{j=1}^n r_{jn}\xi_{jn}(t)$ at times $t_1<t_2<\cdots<t_m$. The characteristic function is given by

$$\begin{aligned}Ee^{i\sum_{k=1}^m\theta_k X_n(t_k)} &= Ee^{i\sum_{j=1}^n r_{jn}\sum_{k=1}^m\theta_k\xi_{jn}(t_k)} = E\prod_{j=1}^n e^{i\sum_{k=1}^m r_{jn}\theta_k\xi_{jn}(t_k)}\\ &= \prod_{j=1}^n Ee^{i\sum_{k=1}^m r_{jn}\theta_k\xi_{jn}(t_k)} \qquad (5.36)\end{aligned}$$

the last equation following from the independence of the sources. Writing $\xi_k := \xi_{jn}(t_k)$ and r in place r_{jn} (for the generic element of the array) and taking into account the fact that

$$e^{i\theta_k r\xi_k} = 1-\xi_k+\xi_k e^{ir\theta_k} = 1+\xi_k(e^{ir\theta_k}-1),$$

due to the fact that ξ_k takes only the values 0 and 1,

$$e^{i\sum_{k=1}^m \theta_k r_{jn}\xi_{jn}(t_k)} = \prod_{k=1}^m e^{i\theta r\xi_k} = \prod_{k=1}^m \Big(1+\xi_k(e^{i\theta_k r}-1)\Big)$$

the quantity inside the expectation can be written as

$$\begin{aligned} Ee^{i\sum_{k=1}^m \theta_k\xi_{jn}(t_k)} &= E\prod_{k=1}^m \Big(1+\xi_k(e^{ir\theta_k}-1)\Big) \\ &= 1+\sum_{k=1}^m \sum_{1\le l_1<l_2<\cdots<l_k\le m} E\left[\xi_{l_1}\cdots\xi_{l_k}\right](e^{ir\theta_{l_1}}-1)\cdots(e^{ir\theta_{l_k}}-1). \end{aligned} \tag{5.37}$$

Set

$$\pi_{jn} := \frac{\lambda_{jn}}{\lambda_{jn}+\mu}, \qquad \alpha_{jn} := \lambda_{jn}+\mu. \tag{5.38}$$

Then

$$\begin{aligned} E\xi_{jn}(t_1) &= \pi_{jn} \\ E\xi_{jn}(t_1)\xi_{jn}(t_2) &= \pi_{jn}\Big(e^{-\alpha_{jn}(t_2-t_1)}+\pi_{jn}\Big(1-e^{-\alpha_{jn}(t_2-t_1)}\Big)\Big) \end{aligned} \tag{5.39}$$

$$\begin{aligned} E[\xi_{l_1}\cdots\xi_{l_k}] &= \pi_{jn}\Big[e^{-\alpha_{jn}(t_{l_2}-t_{l_1})}+\pi_{jn}\Big(1-e^{-\alpha_{jn}(t_{l_2}-t_{l_1})}\Big)\Big]\times\cdots \\ &\quad\times\Big[e^{-\alpha_{jn}(t_{l_k}-t_{l_{k-1}})}+\pi_{jn}\Big(1-e^{-\alpha_{jn}(t_{l_k}-t_{l_{k-1}})}\Big)\Big] \end{aligned} \tag{5.40}$$

We assume that $\pi_{jn} = O(\frac{1}{n})$. Hence,

$$E[\xi_{l_1}\cdots\xi_{l_k}] = \pi_{jn}e^{-\alpha_{jn}(t_{l_k}-t_{l_1})} + o(\frac{1}{n}). \tag{5.41}$$

Using the above equation in (5.40) we obtain

$$Ee^{i\sum_{k=1}^m \theta_k r_{jn}\xi_{jn}(t_k)} = 1+\pi_{jn}\sum_{k=1}^m \sum_{1\le l_1<l_2<\cdots<l_k\le m} e^{-\alpha_{jn}(t_{l_k}-t_{l_1})}(e^{ir\theta_{l_1}}-1)\cdots(e^{ir\theta_{l_k}}-1) + o(\frac{1}{n}). \tag{5.42}$$

$$\begin{aligned} Ee^{i\sum_{k=1}^m \theta_k r_{jn}\xi_{jn}(t_k)} &= 1+\pi_{jn}\sum_{l_1=1}^m (e^{ir\theta_{l_1}}-1) \\ &\quad+\pi_{jn}\sum_{k=2}^m \sum_{1\le l_1<l_2<\cdots<l_k\le m} e^{-\alpha_{jn}(t_{l_k}-t_{l_1})}(e^{ir\theta_{l_1}}-1)\cdots(e^{ir\theta_{l_k}}-1) \end{aligned}$$

The double sum in the above expression can be rearranged as follows

$$\sum_{1\le u<v\le m} e^{-\alpha(t_v-t_u)}(e^{ir\theta_u}-1)(e^{ir\theta_v}-1)\sum_{k=0}^{v-u-1}\sum_{u<l_1<\cdots<l_k<v}(e^{ir\theta_{l_1}}-1)\cdots(e^{ir\theta_{l_k}}-1). \tag{5.43}$$

Using once more the inclusion–exclusion principle we have

$$\sum_{k=0}^{v-u-1}\sum_{u<l_1<\cdots<l_k<v}(e^{ir\theta_{l_1}}-1)\cdots(e^{ir\theta_{l_k}}-1) = e^{ir(\theta_{u+1}+\theta_{u+2}+\cdots+\theta_{v-1})}.$$

The above equation follows easily from the identity

$$\prod_{j=1}^{n}(1+x_j) \;=\; 1+\sum_{k=1}^{n}\sum_{1\le j_1<j_2<\cdots<j_k\le n} x_{j_1}x_{j_2}\cdots x_{j_k}$$

with $x_j = e^{ir\theta_j}-1$, $j=1,\dots,n$. However, we have

$$\begin{aligned}(e^{ir\theta_u}-1)&(e^{ir\theta_v}-1)e^{ir(\theta_{u+1}+\cdots+\theta_{v-1})}\\ &= e^{ir(\theta_u+\cdots+\theta_v)} - e^{ir(\theta_{u+1}+\cdots+\theta_v)} - e^{ir(\theta_u+\cdots+\theta_{v-1})} + e^{ir(\theta_{u+1}+\cdots+\theta_{v-1})}\end{aligned}$$

Thus, (5.43) can be written as

$$\begin{aligned}&\sum_{1\le u<v\le m} e^{-\alpha(t_v-t_u)}\left[e^{ir(\theta_u+\cdots+\theta_v)} - e^{ir(\theta_{u+1}+\cdots+\theta_v)} - e^{ir(\theta_u+\cdots+\theta_{v-1})} + e^{ir(\theta_{u+1}+\cdots+\theta_{v-1})}\right]\\ &= \sum_{1\le u<v\le m} e^{ir(\theta_u+\cdots+\theta_v)}\left[e^{-\alpha(t_v-t_u)} - e^{-\alpha(t_{v+1}-t_u)} - e^{-\alpha(t_v-t_{u-1})} + e^{-\alpha(t_{v+1}-t_{u-1})}\right]\\ &= \sum_{1\le u<v\le m} e^{ir(\theta_u+\cdots+\theta_v)}\left[1-e^{-\alpha(t_u-t_{u-1})}\right]e^{-\alpha(t_v-t_u)}\left[1-e^{-\alpha(t_{v+1}-t_v)}\right]. \qquad (5.44)\end{aligned}$$

Hence

$$\begin{aligned}Ee^{i\sum_{k=1}^m \theta_k r_{jn}\xi_{jn}(t_k)} \;&=\; 1+\sum_{u=1}^{m}\pi_{jn}(e^{ir_{jn}\theta_u}-1)\\ &+\sum_{1\le u<v\le m}\pi_{jn}\left(e^{ir_{jn}(\theta_u+\cdots+\theta_v)}-1\right)\left[1-e^{-\alpha_{jn}(t_u-t_{u-1})}\right]e^{-\alpha_{jn}(t_v-t_u)}\left[1-e^{-\alpha_{jn}(t_{v+1}-t_v)}\right]\end{aligned}$$

Thus, for the process $X_n(t)=\sum_{j=1}^n r_{jn}\xi_{jn}(t)$ we have, in view of the above and (5.36),

$$\begin{aligned}\log Ee^{i\sum_{k=1}^m\theta_k X_n(t_k)} \;&=\; \sum_{j=1}^{n}\log\left(Ee^{i\sum_{k=1}^m\theta_k\xi_{jn}(t_k)}\right)\\ &=\sum_{j=1}^{n}\log\Bigg(1+\sum_{u=1}^{m}\pi_{jn}(e^{ir_{jn}\theta_u}-1)\\ &\quad+\sum_{1\le u<v\le m}\pi_{jn}\left(e^{ir_{jn}(\theta_u+\cdots+\theta_v)}-1\right)\left[1-e^{-\alpha_{jn}(t_u-t_{u-1})}\right]e^{-\alpha_{jn}(t_v-t_u)}\left[1-e^{-\alpha_{jn}(t_{v+1}-t_v)}\right]\Bigg)\\ &=\sum_{u=1}^{m}\sum_{j=1}^{n}\pi_{jn}(e^{ir_{jn}\theta_u}-1)\\ &\quad+\sum_{1\le u<v\le m}\sum_{j=1}^{n}\pi_{jn}\left(e^{ir_{jn}(\theta_u+\cdots+\theta_v)}-1\right)\left[1-e^{-\alpha_{jn}(t_u-t_{u-1})}\right]e^{-\alpha_{jn}(t_v-t_u)}\left[1-e^{-\alpha_{jn}(t_{v+1}-t_v)}\right]\\ &\quad+o(1).\end{aligned}$$

Using again the argument leading to (5.34) we obtain

$$\begin{aligned}\log Ee^{i\sum_{k=1}^m\theta_k X_n(t_k)} \;&\to\; \sum_{u=1}^{m}\int_0^\infty(e^{ix\theta_u}-1)\nu(dx)\\ &+\sum_{1\le u<v\le m}\left[1-e^{-\mu(t_u-t_{u-1})}\right]e^{-\mu(t_v-t_u)}\left[1-e^{-\mu(t_{v+1}-t_v)}\right]\int_0^\infty\left(e^{ix(\theta_u+\cdots+\theta_v)}-1\right)\nu(dx).\end{aligned}$$

Absorbing the first sum into the double sum the above can be written as

$$\sum_{1\le u\le v\le m}\left[1-e^{-\mu(t_u-t_{u-1})}\right]e^{-\mu(t_v-t_u)}\left[1-e^{-\mu(t_{v+1}-t_v)}\right]\int_0^\infty\left(e^{ix(\theta_u+\cdots+\theta_v)}-1\right)\nu(dx).$$

which is equal to the right hand side of (5.30).

Part 3. Tightness.

Having established the convergence of the finite-dimensional distributions we need to also establish the tightness of the family of processes in order to establish convergence in distribution and complete the proof. This is done as follows:

We shall take $\beta = 2$ and establish the theorem for $a = 1$.

$$\begin{aligned}&\mathbb{E}\left[\left(X_n(t_2)-X_n(t_1)\right)^2\left(X_n(t_3)-X_n(t_2)\right)^2\right]\\ &\quad=\ \mathbb{E}\left[\left(\sum_{k=1}^n\xi_{kn}(t_2)-\xi_{kn}(t_1)\right)^2\left(\sum_{k=1}^n\xi_{kn}(t_3)-\xi_{kn}(t_2)\right)^2\right]\end{aligned}$$

$$\begin{aligned}=\ &\mathbb{E}\left[\sum_{k=1}^n\left(\xi_{kn}(t_2)-\xi_{kn}(t_1)\right)^2\left(\xi_{kn}(t_3)-\xi_{kn}(t_2)\right)^2\right] \qquad (5.45)\\ &+\mathbb{E}\left[\sum_{k\ne m}\left(\xi_{kn}(t_2)-\xi_{kn}(t_1)\right)^2\left(\xi_{mn}(t_3)-\xi_{mn}(t_2)\right)^2\right]\\ &+2\mathbb{E}\left[\sum_{k\ne m}\left(\xi_{kn}(t_2)-\xi_{kn}(t_1)\right)\left(\xi_{kn}(t_3)-\xi_{kn}(t_2)\right)\left(\xi_{mn}(t_2)-\xi_{mn}(t_1)\right)\left(\xi_{mn}(t_3)-\xi_{mn}(t_2)\right)\right].\end{aligned}$$

For the first term in (5.45) we have, in view of (3.9) of Lemma 3.1, that

$$\mathbb{E}\left[\sum_{k=1}^n\left(\xi_{kn}(t_2)-\xi_{kn}(t_1)\right)^2\left(\xi_{kn}(t_3)-\xi_{kn}(t_2)\right)^2\right]\ \le\ \frac{\mu(t_3-t_1)^2}{4}\sum_{k=1}^n\lambda_{kn}r_{kn}^4.$$

For the second term in (5.45) we have that $\sum_{k\ne m}\mathbb{E}\left[\left(\xi_{kn}(t_2)-\xi_{kn}(t_1)\right)^2\right]\mathbb{E}\left[\left(\xi_{mn}(t_3)-\xi_{mn}(t_2)\right)^2\right]$ since the sources are independent. Also, from (3.11),

$$\begin{aligned}\mathbb{E}\left(\xi_{kn}(t_2)-\xi_{kn}(t_1)\right)^2\mathbb{E}\left(\xi_{mn}(t_3)-\xi_{mn}(t_2)\right)^2\ &\le\ 4\mu^2(t_3-t_2)(t_2-t_1)\frac{\lambda_{kn}r_{kn}^2}{\lambda_{kn}+\mu}\frac{\lambda_{mn}r_{mn}^2}{\lambda_{mn}+\mu}\\ &\le\ \mu^2(t_3-t_1)^2\frac{\lambda_{kn}r_{kn}^2}{\lambda_{kn}+\mu}\frac{\lambda_{mn}r_{mn}^2}{\lambda_{mn}+\mu}\end{aligned}$$

and thus this second term is bounded by

$$\mathbb{E}\left[\sum_{k\ne m}\left(\xi_{kn}(t_2)-\xi_{kn}(t_1)\right)^2\left(\xi_{mn}(t_3)-\xi_{mn}(t_2)\right)^2\right]\ \le\ \mu^2(t_3-t_1)^2\left(\sum_{k=1}^n\frac{\lambda_{kn}r_{kn}^2}{\lambda_{kn}+\mu}\right)^2.$$

For the third term of (5.45) we note that, because of the independence of sources and (3.10)

$$\begin{aligned}&\left|\mathbb{E}\left[\left(\xi_{kn}(t_2)-\xi_{kn}(t_1)\right)\left(\xi_{kn}(t_3)-\xi_{kn}(t_2)\right)\left(\xi_{mn}(t_2)-\xi_{mn}(t_1)\right)\left(\xi_{mn}(t_3)-\xi_{mn}(t_2)\right)\right]\right| \\ &\qquad \leq \frac{1}{16}\mu^4\lambda_{kn}\lambda_{mn}r_{kn}^2r_{mn}^2(t_3-t_1)^4.\end{aligned}$$

Hence,

$$\begin{aligned}&2\mathbb{E}\left[\sum_{k\neq m}\left(\xi_{kn}(t_2)-\xi_{kn}(t_1)\right)\left(\xi_{kn}(t_3)-\xi_{kn}(t_2)\right)\left(\xi_{mn}(t_2)-\xi_{mn}(t_1)\right)\left(\xi_{mn}(t_3)-\xi_{mn}(t_2)\right)\right] \\ &\leq \frac{\mu^4(t_3-t_1)^4}{8}\sum_{k\neq m}\lambda_{kn}r_{kn}^2\lambda_{mn}r_{mn}^2 \leq \frac{\mu^4(t_3-t_1)^4}{8}\left(\sum_{k=1}^n\lambda_{kn}r_{kn}^2\right)^2.\end{aligned}$$

Therefore,

$$\begin{aligned}\mathbb{E}\left[\left(X_n(t_2)-X_n(t_1)\right)^2\left(X_n(t_3)-X_n(t_2)\right)^2\right] &\leq \frac{\mu(t_3-t_1)^2}{4}\sum_{k=1}^n\lambda_{kn}r_{kn}^4 \\ &\quad+\mu^2(t_3-t_1)^2\left(\sum_{k=1}^n\frac{\lambda_{kn}r_{kn}^2}{\lambda_{kn}+\mu}\right)^2+\frac{\mu^4(t_3-t_1)^4}{8}\left(\sum_{k=1}^n\lambda_{kn}r_{kn}^2\right)^2.\end{aligned}$$

From Assumption C.3 it follows that $\limsup_{n\to\infty}\sum_{k=1}^n\lambda_{kn}r_{kn}^4\leq C$ and $\limsup_{n\to\infty}\sum_{k=1}^n\lambda_{kn}r_{kn}^2\leq C$ and hence, for some $K>0$,

$$\mathbb{E}\left[\left(X_n(t_2)-X_n(t_1)\right)^2\left(X_n(t_3)-X_n(t_2)\right)^2\right]\leq K(t_3-t_1)^2.$$

Finally there remains to establish (3.30) and to this end it suffices to show that $\mathbb{P}(|X_T-X_{T-\delta}|>\epsilon)\to 0$ as $\delta\to 0$ for any $\epsilon>0$. From Markov's inequality we have

$$\mathbb{P}(|X_T-X_{T-\delta}|>\epsilon) \leq \frac{\mathbb{E}\left(X_T-X_{T-\delta}\right)^2}{\epsilon^2} = \frac{1}{\epsilon^2}2H(\delta)\int_0^\infty x^2\nu(dx) \tag{5.46}$$

where in the last inequality we have used (5.14), (5.15), and the fact that $\psi(\theta)=\int_0^\infty\left(e^{i\theta x}-1\right)\nu(dx)$ and $\psi''(0)=-\int_0^\infty x^2\nu(dx)$. Using the fact that $\lim_{\delta\to 0}H(\delta)=0$ in (5.46) proves (3.30). □

It is well known that if the sources are identical then (an appropriately scaled version of) their superposition is an Ehrenfest process and the corresponding limit process obtained is the Ornstein–Uhlenbeck process. Thus the family of processes examined in this paper are a generalization of the classical Ornstein–Uhlenbeck process.

5.6 Further examples of GCID processes

Proposition 5.8. *Suppose that $\{X_t;t\geq 0\}$ is a GCID process with $\psi(\theta)=i\mu\theta-\frac{1}{2}\sigma^2\theta^2$ and given H satisfying the conditions of Theorem 5.2. Then $\{X_t\}$ is a Gaussian process with covariance function*

$$\mathsf{Cov}(X_{t_i},X_{t_j}) = \sigma^2\left(1-H(t_j-t_i)\right),\quad 0\leq t_i\leq t_j. \tag{5.47}$$

If in addition H satisfies $\lim_{h\to 0}H(h)/h=:\ell>0$ then the process has a.s. continuous sample paths.

Proof. From Theorem 5.2 and Lemma 5.5

$$\phi_n(\theta_1,\ldots,\theta_n;t_1,\ldots,t_n) \;=\; i\mu(\theta_1+\cdots+\theta_n)-\sigma^2\sum_{1\le i\le j\le n}(\theta_i+\cdots+\theta_j)^2\,a_{ij} \qquad (5.48)$$

with $a_{ij} = [H(t_j-t_{i-1})-H(t_j-t_i)-H(t_{j+1}-t_{i-1})+H(t_{j+1}-t_i)]$. From (5.7), (5.8),

$$\sum_{1\le i\le j\le n}(\theta_i+\cdots+\theta_j)^2\,a_{ij} \;=\; \sum_{1\le i\le j\le n}\theta_i\theta_j b_{ij}$$

with

$$b_{ij} \;=\; \sum_{k=1}^{i}\sum_{l=j}^{n}[H(t_j-t_{i-1})-H(t_j-t_i)-H(t_{j+1}-t_{i-1})+H(t_{j+1}-t_i)] \;=\; 1-H(t_j-t_i)$$

where in the above we use the telescopic nature of the sums and the convention $t_0 = -\infty$, $t_{n+1} = +\infty$, $H(0) = 0$, $H(+\infty) = 1$. Thus

$$\log\mathbb{E}[e^{i\sum_{k=1}^n\theta_k X_{t_k}}] \;=\; i\mu\sum_{k=1}^n\theta_k-\sum_{j=1}^n\sum_{k=1}^n\theta_j\theta_k\,[1-H(|t_k-t_j|)]\,. \qquad (5.49)$$

Thus $\{X_t\ t\ge 0\}$ is a stationary Gaussian process with covariance given by (5.47).

From the Kolmogorov Continuity Theorem [131, p. 51], a sufficient condition for a process $\{X_t\}$ to have a.s. continuous paths is the existence of positive constants α, β, M, such that $\mathbb{E}|X_{t+h}-X_t|^\alpha \;\le\; Mh^{1+\beta}$. As we have seen

$$\widetilde{\phi}(\theta) \;:=\; \phi_2(-\theta,\theta;t,t+h) \;=\; \mathbb{E}e^{i\theta(X_{t+h}-X_t)} \;=\; e^{H(h)(\psi(\theta)-\psi(-\theta))}.$$

Then $\widetilde{\phi}^{(4)}(0) \;=\; 2H(h)\psi^{(4)}(0)+12H^2(h)\psi''(0)$ and therefore $\widetilde{\phi}^{(4)}(0) \;=\; 12H^2(h)\sigma^2$. By the concavity of H it follows that $H(h)\le\ell h$ for all $h>0$ and hence

$$\mathbb{E}|X_{t+h}-X_t|^4 \;=\; 12H^2(h)\sigma^2 \;\le\; 12\sigma^2\ell^2h^2.$$

This completes the proof. □

The classical Ornstein–Uhlenbeck process is a gaussian GCID process with $H(t) = 1-e^{-\alpha t}$ (with $\alpha>0$) and, since in this case $H(h)/h\to\alpha$ as $h\to 0$, it satisfies the condition in Proposition 5.8 and therefore has (as is well known) continuous sample paths w.p. 1. On the other hand its GCID counterpart with Poisson marginals, the $M/M/\infty$ process, has sample paths that are piecewise constant and have jumps of size 1, a.s. Thus it is clear that path behavior depends both on the nature of the marginal distributions and on the type of correlation structure which is determined by H.

Another instance of a gaussian GCID process which also satisfies the condition of Proposition 5.8 and has therefore a.s. continuous sample paths is the process considered in Slepian [128] and Shepp [126] which we discuss in the following

Example 5.9. If $H(x) = \min(x, 1) = x \wedge 1$ and $\psi(\theta) = -\frac{1}{2}\sigma^2\theta^2$, (5.11) gives

$$\begin{aligned} &\log \phi_n(\theta_1, \ldots, \theta_n; t_1, \ldots, t_n) \qquad (5.50)\\ &= -\frac{1}{2}\sigma^2 \sum_{1\le i\le j\le n} (\theta_i + \cdots + \theta_j)^2 \left[(t_j - t_{i-1}) \wedge 1 - (t_j - t_i) \wedge 1 - (t_{j+1} - t_{i-1}) \wedge 1 + (t_{j+1} - t_i) \wedge 1\right] \end{aligned}$$

Then from Proposition 5.8, (5.50) can be written as

$$-\frac{1}{2}\sigma^2 \left\{ \sum_{i=1}^{n} \theta_i^2 + 2 \sum_{1\le i\le j\le n} \theta_i\theta_j \left[1 - (t_j - t_i) \wedge 1\right] \right\}$$

and thus it represents a Gaussian process with covariance function

$$R(\tau) = \mathsf{Cov}(X_t, X_{t+\tau}) = (1 - |\tau|)^+.$$

It has been pointed out in [128], [126], that the process defined in (5.50) shares with the Ornstein–Uhlenbeck process the property that they are Gaussian processes for which passage times can be computed in closed form. We point out that the corresponding GCID process with Poisson marginals is the $M/D/\infty$ process with service times equal to 1. Indeed, we can check this by refering to section 5.4 and taking the service time distribution in the $M/GI/\infty$ model to be $G(x) = \mathbf{1}(x \ge 1)$ which implies that $G_I(x) = x \wedge 1$ for $x > 0$.

Example 5.10. We take again $H(x) = \min(x, 1) = x \wedge 1$. However, this time we choose $\psi(\theta) = -\log(1 - i\theta)$. Then

$$\begin{aligned} \log \phi_n(\theta_1, \ldots, \theta_n; t_1, \ldots, t_n) &= -\sum_{1\le i\le j\le n} \log[1 - i(\theta_i + \cdots + \theta_j)] \qquad (5.51)\\ &\quad \times \left[(t_j - t_{i-1}) \wedge 1 - (t_j - t_i) \wedge 1 - (t_{j+1} - t_{i-1}) \wedge 1 + (t_{j+1} - t_i) \wedge 1\right]. \end{aligned}$$

This defines a process which has the same covariance structure as the Gaussian process of Example 5.9 but has exponential marginals. In particular, when $t_n - t_1 \le 1$, the right hand side of (5.51) reduces to

$$\begin{aligned} -\sum_{j=1}^{n-1} \log[1 - i(\theta_1 + \cdots + \theta_j)](t_{j+1} - t_j) \; - \; \sum_{i=2}^{n} \log[1 - i(\theta_i + \cdots + \theta_n)](t_i - t_{i-1})\\ - \; \log[1 - i(\theta_1 + \cdots + \theta_n)] \left[1 - (t_n - t_1)\right]. \end{aligned}$$

The distribution of $X_{t+h} - X_t$ can be obtained from (5.14). Its characteristic function is

$$\exp\left(H(h)\left[-\log(1 - i\theta) - \log(1 + i\theta)\right]\right) = \left(\frac{1}{1 + \theta^2}\right)^{H(h)}.$$

This is in fact the difference of two independent Gamma–distributed random variables. Comparing with the Gaussian case, we see that $\mathbb{E}(X_{t+h} - X_t)^4 = 12h(h + 1) = O(h)$ as $h \to 0$. In this case it is not possible to use Kolmogorov's theorem to establish the existence of a version of the process with continuous paths. In particular, note that

$$\phi(\theta_1, \theta_2) = \mathbb{E}\left[e^{i\theta_1 X_{t_1} + i\theta_2 X_{t_2}}\right] = \left(\frac{1}{1 - i(\theta_1 + \theta_2)}\right)^{1 - H(t_2 - t_1)} \left(\frac{1}{1 - i\theta_1}\right)^{H(t_2 - t_1)} \left(\frac{1}{1 - i\theta_2}\right)^{H(t_2 - t_1)}.$$

Thus (5.51) defines a type of multidimensional exponential distribution. If H has finite support, say $[0, \tau]$ then X_t, X_{t+s} are independent for $s > \tau$.

5.7 The Covariance Structure of a Generalized OU Process

In this final section of this chapter we give the covariance structure of the Generalized OU process for the purpose of comparison with the class of processes defined above. Suppose that $\{Z_t\}$ is a Lévy process with characteristic exponent $\chi(\theta) := \frac{1}{t}\log \mathbb{E}e^{i\theta Z_t}$ and consider the Generalized OU process $\{X_t\}$ given in (5.3) which arises as the solution of the Ornstein-Uhlenbeck SDE (5.2).

Proposition 5.11. *Let $\chi(\theta) := \log Ee^{i\theta Z_1}$ denote the characteristic exponent of the Lévy process and*

$$\psi(\theta) \;:=\; \int_0^\theta \frac{\chi(u)}{\alpha u}\, du. \tag{5.52}$$

Then, if X_0 is defined so that the process $\{X_t\}$ is stationary, i.e. $X_0 := \int_0^\infty e^{-\alpha s} dZ'_s$ (where $\{Z'_t\}$ is an independent copy *of the driving Lévy process $\{Z_t\}$, the marginal distribution of the process is given by*

$$Ee^{i\theta X_t} = e^{\psi(\theta)} \qquad \textit{for all } t \geq 0 \tag{5.53}$$

while the finite dimensional distributions of $\{X_t\}$ are given by

$$\begin{aligned} \log Ee^{i\sum_{k=1}^n \theta_k X_{t_k}} \;&=\; \sum_{l=1}^{n-1} \left(\psi\left(\textstyle\sum_{k=l}^n \theta_k e^{-a(t_k-t_l)}\right) \;-\; \psi\left(\textstyle\sum_{k=l+1}^n \theta_k e^{-a(t_k-t_l)}\right)\right) \\ &\quad + \psi(\theta_n). \end{aligned} \tag{5.54}$$

Proof. We begin with the following alternative expression for $\psi(\theta)$:

$$\log E[e^{i\theta X_0}] \;=\; \log E[e^{i\theta\int_0^\infty e^{-\alpha s}dZ_s}] \;=\; \int_0^\infty \chi(\theta e^{-\alpha s})ds \;=\; \int_0^\theta \frac{\chi(u)}{\alpha u}du \;=\; \psi(\theta)\,. \tag{5.55}$$

If $\{X_t\}$ is the solution of the Stochastic Differential Equation (5.2) considered pathwise, and $0 < t_1 < t_2 < \cdots < t_k$ then

$$e^{at_k}X_{t_k} \;=\; e^{at_1}X_{t_1} + \int_{t_1}^{t_k} e^{as}dZ_s. \tag{5.56}$$

Set $\zeta_k = \int_{t_k}^{t_{k+1}} e^{as}dZ_s$. Then

$$\begin{aligned} E[e^{i\theta\zeta_k}] \;&=\; \exp\left(\int_{t_k}^{t_{k+1}} \chi(\theta e^{as})du\right) \;=\; \exp\left(\int_{\theta e^{at_k}}^{\theta e^{at_{k+1}}} \frac{\chi(u)}{\alpha u}du\right) \\ &=\; \exp\left(\psi(\theta e^{\alpha t_{k+1}}) - \psi(\theta e^{\alpha t_k})\right). \end{aligned} \tag{5.57}$$

Suppose also that the Lévy process is defined on the whole real line. To this end we extend a given Lévy process $\{Z_t; t \geq 0\}$ (with $Z_0 = 0$ $\mathbb{P}$–a.s.) by considering a second independent, copy of the process, $\{\tilde{Z}; t \geq 0\}$ and setting

$$Z'_t \;=\; \begin{cases} Z_t & \text{if} \quad t \geq 0 \\ \tilde{Z}_{-t} & \text{if} \quad t < 0 \end{cases}.$$

Then setting $X_0 := \int_{-\infty}^0 e^{\alpha s}dZ_s$, ensures that the process $\{X_t\}$ is stationary. Then

$$\sum_{k=1}^n \theta_k X_{t_k} \;=\; \sum_{k=1}^n \theta_k e^{-at_k}\left(e^{at_1}X_{t_1} + (\zeta_1 + \cdots + \zeta_{k-1})\right)$$

and since the random variables $X_{t_1}, \zeta_1, \ldots, \zeta_{n-1}$ are independent,

$$Ee^{i\sum_{k=1}^n \theta_k X_{t_k}} \;=\; E\exp\left(i\left(\sum_{k=1}^n \theta_k e^{-a(t_k-t_1)}\right)X_{t_1}\right)\prod_{k=1}^{n-1} E\exp\left(i\left(\sum_{l=k+1}^n \theta_l e^{-at_l}\right)\zeta_k\right)\;.$$

From (5.53)

$$\log E\exp\left(i\left(\sum_{k=1}^n \theta_k e^{-a(t_k-t_1)}\right)X_{t_1}\right) \;=\; \psi\left(\sum_{k=1}^n \theta_k e^{-a(t_k-t_1)}\right)$$

while, from (5.57),

$$\log E\exp\left(i\left(\sum_{l=k+1}^n \theta_l e^{-at_l}\right)\zeta_k\right) \;=\; \psi\left(\sum_{l=k+1}^n \theta_l e^{-a(t_l-t_{k+1})}\right) \;-\; \psi\left(\sum_{l=k+1}^n \theta_l e^{-a(t_l-t_k)}\right)\;.$$

From the above

$$\begin{aligned}\log Ee^{i\sum_{k=1}^n \theta_k X_{t_k}} \;=\;& \psi\left(\sum_{k=1}^n \theta_k e^{-a(t_k-t_1)}\right)\\ &+\sum_{k=1}^{n-1}\left(\psi\left(\sum_{l=k+1}^n \theta_l e^{-a(t_l-t_{k+1})}\right) \;-\; \psi\left(\sum_{l=k+1}^n \theta_l e^{-a(t_l-t_k)}\right)\right)\end{aligned}$$

which can be displayed in telescopic form

$$\begin{aligned}\log Ee^{i\sum_{k=1}^n \theta_k X_{t_k}} \;=\;& \psi\left(\sum_{k=1}^n \theta_k e^{-a(t_k-t_1)}\right) \;-\; \psi\left(\sum_{k=2}^n \theta_k e^{-a(t_k-t_1)}\right)\\ &+\psi\left(\sum_{k=2}^n \theta_k e^{-a(t_k-t_2)}\right) \;-\; \psi\left(\sum_{k=3}^n \theta_k e^{-a(t_k-t_2)}\right)\\ &\vdots\\ &+\psi\left(\theta_{n-1}e^{-a(t_{n-1}-t_{n-1})}+\theta_n e^{-a(t_n-t_{n-1})}\right) \;-\; \psi\left(\theta_n e^{-a(t_n-t_{n-1})}\right)\\ &+\psi\left(\theta_n e^{-a(t_n-t_n)}\right).\end{aligned}$$

This simplifies into (5.54). □

Compound Poisson Driven Processes

In the compound Poisson process with positive jumps $\chi(u) = \lambda(\widehat{G}(u)-1)$ where $\widehat{G}(u) = \int_0^\infty e^{iux}dG(x)$ and hence

$$\chi(u) \;=\; \lambda\int_0^\infty \left(e^{iux}-1\right)dG(x)$$

Also, $e^{iux}-1 = iu\int_0^x e^{iuy}dy$

$$\begin{aligned}\psi(\theta) \;&=\; \frac{\lambda}{\alpha}\int_0^\theta \frac{du}{u}\int_0^\infty \left(e^{iux}-1\right)dG(x) \;=\; \frac{i\lambda}{\alpha}\int_0^\theta du\int_{x=0}^\infty\int_0^x e^{iuy}dydG(x)\\ &=\; \frac{i\lambda}{\alpha}\int_{u=0}^\theta du\int_{y=0}^\infty e^{iuy}dy\int_{x=y}^\infty dG(x) \;=\; \frac{i\lambda}{\alpha}\int_{y=0}^\infty \overline{G}(y)\int_{u=0}^\theta e^{iuy}dudy\\ &=\; \frac{\lambda}{\alpha}\int_{y=0}^\infty \overline{G}(y)\int_0^{i\theta y}\frac{e^t}{y}dt \;=\; \frac{\lambda}{\alpha}\int_0^\infty \left(e^{i\theta y}-1\right)\frac{\overline{G}(y)}{y}dy.\end{aligned}$$

Then the joint characteristic function of $(X_{t_1}, \ldots, X_{t_n})$ given by (5.54) becomes

$$\begin{aligned}
\log E e^{i\sum_{k=1}^n \theta_k X_{t_k}} &= \sum_{l=1}^{n-1} \int_0^\infty \left(e^{iy\sum_{k=l}^n \theta_k e^{-a(t_k-t_l)}} - e^{iy\sum_{k=l+1}^n \theta_k e^{-a(t_k-t_l)}} \right) \frac{\lambda}{\alpha} \frac{\overline{G}(y)}{y} dy \\
&\quad + \int_0^\infty \left(e^{iy\theta_n} - 1 \right) \frac{\lambda}{\alpha} \frac{\overline{G}(y)}{y} dy \\
&= \sum_{l=1}^{n} \int_0^\infty e^{iy\sum_{k=l+1}^n \theta_k e^{-a(t_k-t_l)}} \left(e^{iy\theta_l} - 1 \right) \frac{\lambda}{\alpha} \frac{\overline{G}(y)}{y} dy.
\end{aligned}$$

6. SMOOTHED MONTE CARLO ESTIMATES FOR THE TIME-IN-THE-RED IN THE DE FINETTI MODEL

6.1 Introduction

We consider the classical continuous time claim process in insurance risk theory which has the following structure: Claims occur at times $\{t_n; n \in \mathbb{N}\}$ which form a Poisson process with rate λ and corresponding counting process $\{N(t); t \geq 0\}$ where $N(t) = \sum_{k=1}^{\infty} \mathbf{1}(t_k \leq t)$. The claim sizes $\{Z_k; k \in \mathbb{N}\}$ are independent, identically distributed random variables, having common distribution function F with $F(x) = 0$ when $x < 0$ and finite mean μ. Furthermore, N and $\{Z_k\}$ are assumed to be independent. Then, $\sum_{i=1}^{N(t)} Z_i$ represents the accumulated claims up to time t.

In the Cramér–Lundberg model (e.g. see Grandell [73]) the free reserves process X is defined as

$$X_t = u + ct - \sum_{i=1}^{N(t)} Z_i$$

where u is the initial capital and $c > 0$ the premium income per unit time. Then, the profit over the interval $[0, t]$ is $S_t = ct - \sum_{i=1}^{N(t)} Z_i$. The *relative safety loading* ρ is defined by

$$\rho = \frac{c}{\lambda\mu} - 1$$

and gives the expected profit rate per unit time as a percentage of the expected claims that have occurred up to that point. We assume that $\rho > 0$, (positive safety loading). Under this assumption, as $t \to \infty$, X_t almost surely will drift to $+\infty$. The typical measure for the long-term financial stability of the risk business in the Cramér–Lundberg model is the *ruin probability*, which is usually expressed as a function of the initial capital u, and defined as $P\{X_t < 0, \text{ for some } t > 0\}$. If one only considers the operation of the company over a finite horizon, say the interval $[0, T]$, then the finite horizon ruin probability, $P\{X(t) < 0, \text{ for some } t \leq T\}$ may be the relevant performance criterion. For further details we refer the reader to Grandell [73].

While the Cramér–Lundberg model has played a central role in the development of risk theory and of actuarial techniques for the analysis of the long-run stability of an insurance company, it has also been criticized, particularly in connection with the feature of the model that free reserves accumulate without bound as $t \to \infty$. In the 1950's de Finetti [38] proposed an alternative point of view which places in the center of the economic argument not the long run stability of the company but the present value of the dividend stream the firm's operation

generates for shareholders. Arguing along these lines he analyzed a simple model and showed that the optimal dividend strategy before the inevitable ruin occurs is a *barrier strategy*. This means that there exists a level $L > 0$ such that, as long as the free reserves are below it all premium income is added to them, whereas as soon as free reserves exceed it, all additional income from premiums is distributed to the shareholders as dividends. In de Finetti's model, the ruin of the insurance company (i.e. the event that the free reserves will become at some time negative) is a certain event. In his original formulation de Finetti argued that the objective of the insurance company would be to maximize *the present value* of the total amount of dividends distributed to shareholders as opposed to keeping the ruin probability below a given value.

We consider a variation of de Finetti's model where, when free reserves become negative the company is not ruined but is instead allowed to continue its operation by borrowing. The period of time during which the free reserves are negative is usually referred to as *time in the red.*

Time in the red has been studied before, in the context of the classical Cramér–Lundberg model based on the following considerations. As argued in Gerber [65], [66], sometimes the event of ruin has a very small probability and the portfolio is just one out of many in the company. The company may thus have enough funds available to support some negative surplus for some time (or secure support from outside sources) in the hope that the portfolio will recover in the future, allowing the company to keep this business alive. This can be regarded as an investment, since the process will recover in the future. The problem of finding if this recovery is quick enough or not, giving good value or not for the money invested, was studied by dos Reis [47] where the distribution of the number of occasions on which the surplus falls below zero is given and results for the moments of the duration of a single period of negative surplus and the total duration of negative surplus are obtained. Also, Dickson and dos Reis [46] consider the distributions of the duration of a single period of negative surplus and of the total duration of negative surplus. They derive explicit results in some cases and show how to approximate these distributions through the use of a discrete time risk model. A markovian analysis of a discrete model with recursive formulas for computing the time in the red is presented in Wagner [143]. The above analytic results have been obtained under the assumption that the claim occurrence process is Poisson. In general, analytic results are not available and one would have to resort to simulation experiments in order to estimate the time in the red.

To the best of our knowledge, the time in the red process for the de Finetti model has not so far been studied. In this paper we propose a smoothed Monte Carlo estimator which has of course the advantage of being applicable under general stochastic assumptions regarding the claim process.

6.2 Model description and Monte Carlo estimators

6.2.1 The dynamics of the free reserves process

To describe the evolution of the system let, as before, $\{t_n; n = 1, 2, \ldots\}$ denote the epochs when claims occur, Z_n the size of the $n-$th claim, and X_t the size of the free reserves at time t. Also set $t_0 = 0$ and suppose that the initial value of the free reserves process is u. Then the process

$\{X_t; t \geq 0\}$ has piece–wise continuous paths, which we will assume to be *right–continuous* with probability one. Between claim occurrences the sample paths of $\{X_t\}$ increase with rate c until the level L is reached and then they remain constantly equal to L (since additional income from premiums is distributed to the shareholders) until the next claim occurs. When this happens, the free reserves are decreased by the amount of the claim (see figure 1). The evolution of the process $\{X_t\}$ can be described heuristically by the following equations

$$\begin{aligned} \frac{d}{dt}X_t &= c\mathbf{1}(X_t < L), \quad t \in (t_n, t_{n+1}),\ n = 0, 1, 2, \ldots \\ X_{t_n} &= X_{t_n-} - Z_n, \quad n = 1, 2, \ldots, \end{aligned}$$

together with the initial condition $X_0 = u$. In the above equations, as usual, X_{t_n-} denotes the value of free reserves just before the nth claim occurs while X_{t_n} the corresponding value just after the claim. From a mathematical standpoint, $\{X_t\}$ is defined (pathwise) as the unique solution of the integral equation

$$X_t = u + c\int_0^t \mathbf{1}(X_s < L)ds - \sum_{k=1}^{N(t)} Z_k.$$

An explicit solution to the above equation is provided by equation (6.4) of section 5.3.

If the operating horizon is t, the total amount of money given to the shareholders is equal to

$$c\int_0^t \mathbf{1}(X_s = L)ds,$$

while the total time in the red is equal to

$$\int_0^t \mathbf{1}(X_s < 0)ds.$$

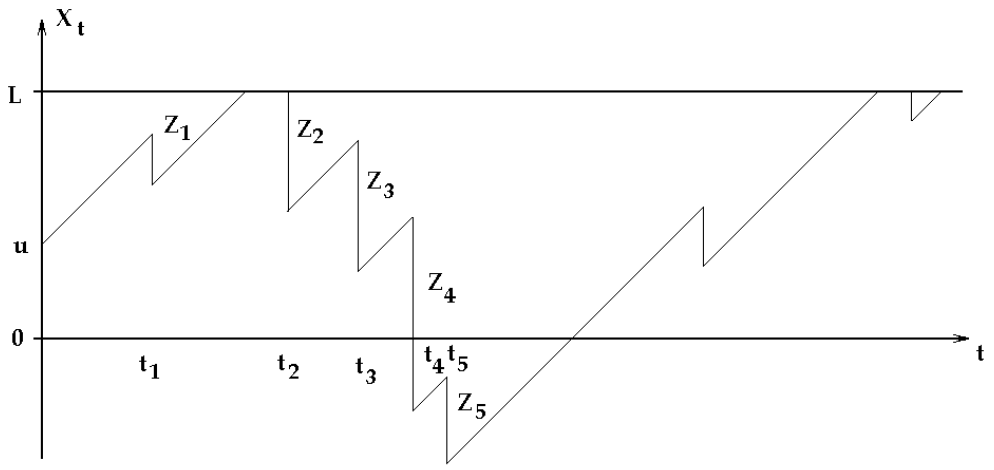
6.2.2 Monte Carlo estimators for the time-in-the-red

A discretized performance criterion which is essentially equivalent to the total time in the red over the horizon $[0, t]$ would be

$$M(t) = \sum_{i=1}^{N(t)} \mathbf{1}\left(X_{t_i} < 0\right), \tag{6.1}$$

the total number of claims up to time t that result in negative free reserves. By an abuse of terminology *we will also be calling $M(t)$ henceforth for convenience "time in the red".*

Since we assume that the net profit condition holds, once the process falls below zero it will remain negative only for a finite period of time before becoming positive again. If we call such periods of negative surpluses *red periods*, then, as long as we have positive loading, red periods are random variables that are finite with probability 1.



The *crude Monte-Carlo estimator* for the time in the red based on counts is $M(t)$, as given in (6.1). In this paper we propose the following *smoothed estimator*

$$K(t) = \sum_{i=1}^{N(t)} \overline{F}(X_{t_i-}). \tag{6.2}$$

where $\overline{F}(x) := 1 - F(x)$ for all $x \in \mathbb{R}$.

Remark: Recall that X_{t_i-} is the size of the reserves, *just prior to the occurrence of the ith claim* and $\overline{F}(X_{t_i-})$ is the conditional probability that the process will fall below level zero after the occurrence of the ith claim, given the size of the free reserves just before the claim occurs. Also note that, by definition, $\overline{F}(x) = 1$ when $x \leq 0$.

The statistical properties of the smoothed estimator will be examined in section 6.3 where it will be shown that it is superior to the crude estimator.

6.2.3 Stochastic Monotonicity of the free reserves process.

We begin with a lemma that establishes the stochastic monotonicity of the free reserves process with respect to the initial capital.

Lemma 6.1 (Stochastic Monotonicity)**.** *If $\{X_t(u)\}$ is the risk process with initial capital u, then $u_1 \leq u_2$ implies $X_t(u_1) \leq_{\text{st}} X_t(u_2)$ for all t.*

Proof. The lemma is an immediate consequence of the corresponding stochastic monotonicity result for queueing systems (see Stoyan [130]). With

$$S_t := ct - \sum_{k=1}^{N(t)} Z_k \tag{6.3}$$

we have the following representation for the free reserves process

$$X_t(u) = \min\left\{u + S_t, L + \inf_{0\le v\le t}[S_t - S_v]\right\}. \tag{6.4}$$

From the above representation it is clear that $u_1 \le u_2$ implies $X_t(u_1) \le X_t(u_2)$ w.p.1. for all $t \ge 0$. □

Having established the stochastic monotonicity of the free reserves process as a function of the initial reserves, we can now use state the following corollary which establishes the monotonicity of φ whose definition is given in (6.12) and repeated here for convenience.

Corollary 6.2. *Suppose* $1 \le i < j$. *The function*

$$\varphi(\overline{u}) := E[\overline{F}(X_{t_j-})\mathbf{1}(t_j \le t)\big|\, X_{t_i} = \overline{u}]$$

is a decreasing function of $\overline{u}$.

Proof. Use the Strong Markov property to argue that

$$E[\overline{F}(X_{t_j-})\mathbf{1}(t_j \le t)\big|\, X_{t_i} = \overline{u}] = E[\overline{F}(X_{t_{j-i}-})\mathbf{1}(t_{j-i} \le t)\big|\, X_0 = \overline{u}]$$

and appeal to the stochastic monotonicity lemma above recalling that $\overline{F}$ is a decreasing function. □

6.3 Statistical Properties of the Smoothed Estimator $K(t)$

Here we formulate and prove our main result, namely that the smoothed estimator is unbiased and has lower variance than the crude estimator for all t. Denote by $(\Omega, \mathcal{F}, P)$ the probability space on which the free reserves process has been defined and by $\mathcal{F}_t = \sigma - \{X_s; s \le t\}$ the σ–field generated by the process X up to time t. The filtration $\{\mathcal{F}_t; t \ge 0\}$ represents thus the history of the process. For background on the theory of processes we refer the reader to Métivier [106]. We recall that $\mathcal{F}_{t+} = \bigcap_{t'>t} \mathcal{F}_{t'}$ and that $\mathcal{F}_{t-} = \bigvee_{t'<t} \mathcal{F}_{t'}$, the σ–field generated by all $\mathcal{F}_{t'}$ with $t' < t$. In accordance with the "usual assumptions" the filtration $\{\mathcal{F}_t\}$ is right–continuous and hence $\mathcal{F}_{t+} = \mathcal{F}_t$.

Recall that, if T is an $\mathcal{F}_t$–stopping time then the stopped σ–field, $\mathcal{F}_T$, is defined as $\mathcal{F}_T := \{A \in \mathcal{F} : A \cap \{T \le t\} \in \mathcal{F}_t \text{ for all } t \ge 0\}$. We also define $\mathcal{F}_{T-}$ as the σ–field generated by the collection of sets $\{A \in \mathcal{F} : A \cap \{T < t\} \in \mathcal{F}_t \text{ for all } t \ge 0\}$.

In our case, the times when claims occur, $\{t_n\}$, form an increasing sequence of stopping times with respect to the filtration $\{\mathcal{F}_t\}$. The corresponding σ–fields $\mathcal{F}_{t_i}$ represent the information available up to the epoch of occurrence of the ith claim, *including the size of the ith claim, Z_i*. Because of the simple structure of our process which evolves according to a deterministic law between claim occurrences and our assumption of right–continuity for the sample paths, it is easy to see that the corresponding σ–fields $\mathcal{F}_{t_i-}$ contain all the information up to the epoch of occurrence of the ith claim, *excluding the size of the ith claim.*

In the course of the proof of our main result we will need the following simple

Lemma 6.3. *Suppose that f, g, are increasing functions, $\mathbb{R} \to \mathbb{R}$ and Z a real random variable. Then*

$$E[f(Z)g(Z)] \geq Ef(Z)Eg(Z)$$

provided the expectations exist.

For a proof see Ross [122].

We are now ready to state our main result.

Theorem 6.4. *The smoothed estimator for the time in the red, given by (6.2) is unbiased and has lower variance than the crude estimator (6.1) i.e. $EK(t) = EM(t)$ for all $t \geq 0$ and*

$$\mathsf{Var}(K(t)) \leq \mathsf{Var}(M(t)) \quad \textit{for all } t \geq 0. \tag{6.5}$$

Proof. We first establish the unbiasedness of the smoothed estimator by showing that its expectation is equal to that of the crude estimator which is obviously unbiased. Indeed we have

$$\begin{aligned}
EM(t) &= E\left[\sum_{i=1}^{\infty} \mathbf{1}(X_{t_i} < 0, t_i \leq t)\right] \\
&= \sum_{i=1}^{\infty} E\left[E\left[\mathbf{1}(t_i \leq t)\mathbf{1}(X_{t_i} < 0)\right] | \mathcal{F}_{t_i-}\right] \\
&= \sum_{i=1}^{\infty} E\left[E\left[\mathbf{1}(X_{t_i} < 0) | \mathcal{F}_{t_i-}\right] \mathbf{1}(t_i \leq t)\right] \\
&= \sum_{i=1}^{\infty} E\left[\mathbf{1}(t_i \leq t)\overline{F}(X_{t_i-})\right] \\
&= E\left[\sum_{i=1}^{\infty} \overline{F}(X_{t_i-})\mathbf{1}(t_i \leq t)\right] = EK(t)
\end{aligned}$$

The interchange between the sum and the expectation in the second and in the last equality can be justified easily using an argument based on the monotone convergence theorem. The fourth equality holds because

$$\begin{aligned}
E\left[\mathbf{1}(X_{t_i} < 0) | \mathcal{F}_{t_i-}\right] &= E\left[\mathbf{1}(X_{t_i-} - Z_i < 0) | \mathcal{F}_{t_i-}\right] = E\left[\mathbf{1}(Z_i > X_{t_i-}) | \mathcal{F}_{t_i-}\right] \\
&= \overline{F}(X_{t_i-}).
\end{aligned}$$

This establishes the unbiasedness of the smoothed estimator.

Next we will establish (6.5). We should point out that when the claim distribution is deterministic the smoothed estimator becomes the same as the crude one. Thus the inequality above cannot be strict in all cases. It is however possible to show that, if the deterministic claim case

is excluded, then the inequality (6.5) becomes strict. We begin with the relationship

$$\begin{aligned} E[M^2(t)] &= E\left[\sum_{i,j} \mathbf{1}(X_{t_i} < 0)\mathbf{1}(X_{t_j} < 0)\mathbf{1}(t_i \le t, t_j \le t)\right] \\ &= EM(t) + 2E\left[\sum_{i<j} \mathbf{1}(X_{t_i} < 0)\mathbf{1}(X_{t_j} < 0)\mathbf{1}(t_j \le t)\right]. \end{aligned}$$

(Both indices, i and j here and in the sequel range of course from 1 to infinity.) Write also the corresponding relationship

$$\begin{aligned} E[K^2(t)] &= E\left[\sum_{i,j} \overline{F}(X_{t_i-})\overline{F}(X_{t_j-})\mathbf{1}(t_i \le t, t_j \le t)\right] \\ &= E\left[\sum_{i} \overline{F}^2(X_{t_i-})\mathbf{1}(t_i \le t)\right] + 2E\left[\sum_{i<j} \overline{F}(X_{t_i-})\overline{F}(X_{t_j-})\mathbf{1}(t_j \le t)\right] \\ &= EK(t) + E\left[\sum_{i} \left(\overline{F}^2(X_{t_i-}) - \overline{F}(X_{t_i-})\right)\mathbf{1}(t_i \le t)\right] \\ &\qquad +2E\left[\sum_{i<j} \overline{F}(X_{t_i-})\overline{F}(X_{t_j-})\mathbf{1}(t_j \le t)\right]. \end{aligned}$$

However, we have already established that $EM(t) = EK(t)$ and

$$E\left[\sum_{i} \left(\overline{F}^2(X_{t_i-}) - \overline{F}(X_{t_i-})\right)\mathbf{1}(t_i \le t)\right] \le 0$$

since each one of the terms inside the sum is negative or zero. Thus, in order to establish (6.5), it suffices to prove that

$$E\left[\sum_{i<j} \mathbf{1}(X_{t_i} < 0)\mathbf{1}(X_{t_j} < 0)\mathbf{1}(t_j \le t)\right] \ge E\left[\sum_{i<j} \overline{F}(X_{t_i-})\overline{F}(X_{t_j-})\mathbf{1}(t_j \le t)\right]$$

or equivalently

$$E\left[\sum_{i<j} \left[\mathbf{1}(X_{t_i} < 0)\mathbf{1}(X_{t_j} < 0) - \overline{F}(X_{t_i-})\overline{F}(X_{t_j-})\right]\mathbf{1}(t_j \le t)\right] \ge 0$$

or

$$\sum_{i<j} E\left[E\left[\left(\mathbf{1}(X_{t_i} < 0)\mathbf{1}(X_{t_j} < 0) - \overline{F}(X_{t_i-})\overline{F}(X_{t_j-})\right)\mathbf{1}(t_j \le t) \mid \mathcal{F}_{t_i}\right]\right] \ge 0. \tag{6.6}$$

Since X_{t_i-}, $X_{t_i} \in \mathcal{F}_{t_i}$, the inner expectation of the typical term in the above summation can be expressed as

$$\begin{aligned} &E\left[\mathbf{1}(X_{t_i} < 0)\mathbf{1}(X_{t_j} < 0)\mathbf{1}(t_j \le t) - \overline{F}(X_{t_i-})\overline{F}(X_{t_j-})\mathbf{1}(t_j \le t)\middle|\, \mathcal{F}_{t_i}\right] \\ &\quad = E\left[\mathbf{1}(X_{t_i} < 0)\mathbf{1}(X_{t_j} < 0)\mathbf{1}(t_j \le t) - \mathbf{1}(X_{t_i} < 0)\overline{F}(X_{t_j-})\mathbf{1}(t_j \le t)\right. \\ &\qquad \left. +\mathbf{1}(X_{t_i} < 0)\overline{F}(X_{t_j-})\mathbf{1}(t_j \le t) - \overline{F}(X_{t_i-})\overline{F}(X_{t_j-})\mathbf{1}(t_j \le t)\middle|\, \mathcal{F}_{t_i}\right] \\ &\quad = \mathbf{1}(X_{t_i} < 0)E\left[\left(\mathbf{1}(X_{t_j} < 0) - \overline{F}(X_{t_j-})\right)\mathbf{1}(t_j \le t)\middle|\, \mathcal{F}_{t_i}\right] \\ &\qquad + \left(\mathbf{1}(X_{t_i} < 0) - \overline{F}(X_{t_i-})\right)E\left[\overline{F}(X_{t_j-})\mathbf{1}(t_j \le t)\middle|\, \mathcal{F}_{t_i}\right]. \end{aligned}$$

However, $i < j$ implies that $\mathcal{F}_{t_i} \subset \mathcal{F}_{t_j-}$ and hence, taking into consideration that $X_{t_j} = X_{t_j-} - Z_j$,

$$\begin{aligned} E\left[\left(\mathbf{1}(X_{t_j} < 0) - \overline{F}(X_{t_j-})\right)\mathbf{1}(t_j \le t)\middle|\mathcal{F}_{t_i}\right] \\ &= E\left[E\left[\left(\mathbf{1}(X_{t_j-} < Z_j) - \overline{F}(X_{t_j-})\right)\mathbf{1}(t_j \le t)\middle|\mathcal{F}_{t_j-}\right]\middle|\mathcal{F}_{t_i}\right] \\ &= E\left[E\left[\mathbf{1}(X_{t_j-} < Z_j)\middle|\mathcal{F}_{t_j-}\right]\mathbf{1}(t_j \le t) - \overline{F}(X_{t_j-})\mathbf{1}(t_j \le t)\middle|\mathcal{F}_{t_i}\right] \\ &= 0 \end{aligned}$$

where, in the next to the last equation we have used the fact that $\mathbf{1}(t_j \le t) \in \mathcal{F}_{t_j-}$. In the last equation we have also used the fact that $E\left[\mathbf{1}(X_{t_j-} < Z_j)\middle|\mathcal{F}_{t_j-}\right] = \overline{F}(X_{t_j-})$. So, in order to establish (6.6), it suffices to show that

$$\sum_{i<j} E\left[\left(\mathbf{1}(X_{t_i} < 0) - \overline{F}(X_{t_i-})\right)E\left[\overline{F}(X_{t_j-})\mathbf{1}(t_j \le t)\,|\,\mathcal{F}_{t_i}\right]\right] \ge 0$$

or equivalently

$$\sum_{i<j} E\left[\left(\mathbf{1}(X_{t_i} < 0) - \overline{F}(X_{t_i-})\right)\overline{F}(X_{t_j-})\mathbf{1}(t_j \le t)\right] \ge 0. \tag{6.7}$$

But $X_{t_i} = X_{t_i-} - Z_i$ and $X_{t_j-} = X_{t_i} + Y_{i,j}$ with

$$Y_{i,j} := -\sum_{k=i+1}^{j-1} Z_k + c\int_{t_i}^{t_j} \mathbf{1}(X_s < L)ds,$$

where the first term in the above sum is the amount paid due to claims and the second is the total income from premiums that is added to the free reserves. With this notation the typical term in the sum (6.7) can be written as

$$E\left[\left(\mathbf{1}(X_{t_i-} - Z_i < 0) - \overline{F}(X_{t_i-})\right)\overline{F}(X_{t_i} + Y_{i,j})\mathbf{1}(t_j \le t)\right]. \tag{6.8}$$

We will show that (6.8) is non–negative and this will establish (6.7) and thus the second part of the theorem. To this end it is enough to show that the following conditional expectation is non–negative:

$$E\left[\left(\mathbf{1}(X_{t_i-} - Z_i < 0) - \overline{F}(X_{t_i})\right)\overline{F}(X_{t_i} + Y_{i,j})\mathbf{1}(t_j \le t)\middle|\mathcal{F}_{t_i-}\right] \ge 0. \tag{6.9}$$

In order to prove (6.9) we have to check two cases:

1. $X_{t_i-} \le 0$. Then $\mathbf{1}(X_{t_i-} - Z_i < 0) = 1$ and $\overline{F}(X_{t_i-}) = 1$ with probability 1, so the left hand side of (6.9) vanishes.

2. $X_{t_i-} > 0$. In this case, write the left hand side of (6.9) as an iterated expectation

$$E\left[E\left[\left(\mathbf{1}(X_{t_i-} - Z_i < 0) - \overline{F}(X_{t_i-})\right)\overline{F}(X_{t_i} + Y_{i,j})\mathbf{1}(t_j \le t)\middle|\mathcal{F}_{t_i}\right]\middle|\mathcal{F}_{t_i-}\right] \tag{6.10}$$

(remember that $\mathcal{F}_{t_i-} \subset \mathcal{F}_{t_i}$). Using the Strong Markov property, the inner expectation in (6.10) can be written as

$$\begin{aligned} \left(\mathbf{1}(X_{t_i-} - Z_i < 0) - \overline{F}(X_{t_i-})\right)E\left[\overline{F}(X_{t_i} + Y_{i,j})\mathbf{1}(t_j \le t)\middle|\mathcal{F}_{t_i}\right] \\ = \left(\mathbf{1}(X_{t_i-} - Z_i < 0) - \overline{F}(X_{t_i-})\right)E\left[\overline{F}(X_{t_i} + Y_{i,j})\mathbf{1}(t_j \le t)\middle|X_{t_i}\right] \end{aligned} \tag{6.11}$$

Let

$$\varphi(\overline{u}) := E\left[\overline{F}(X_{t_i} + Y_{i,j})\mathbf{1}(t_j \le t)\middle| X_{t_i} = \overline{u}\right]. \tag{6.12}$$

As was previously shown in corollary 1, φ is an increasing function of $\overline{u}$. But then, (6.10) can be written as

$$\begin{aligned} E\left[\left(\mathbf{1}(X_{t_i-} - Z_i < 0) - \overline{F}(X_{t_i-})\right)\varphi(X_{t_i-} - Z_i)\middle|\mathcal{F}_{t_i-}\right] \\ \ge E\left[\mathbf{1}(X_{t_i-} - Z_i < 0) - \overline{F}(X_{t_i-})\middle|\mathcal{F}_{t_i-}\right] E\left[\varphi(X_{t_i-} - Z_i)|\mathcal{F}_{t_i-}\right] \\ = 0. \end{aligned} \tag{6.13}$$

The inequality above is a consequence of corollary 1 of the previous section with

$$f(x) := \mathbf{1}(x > X_{t_i}) - \overline{F}(X_{t_i})$$

and

$$g(x) := \varphi(X_{t_i-} - x)$$

being the corresponding increasing functions. Then the last equality in (6.13) follows immediately from the fact that $P(X_{t_i-} - Z_i < 0 \mid X_{t_i-}) = \overline{F}(X_{t_i-})$ This concludes the proof of (6.9) and hence the proof of the theorem. □

The fact that the smoothed estimator has lower variance is of course not surprising. What is interesting however, is the extent to which the simple type of smoothing we propose reduces variance. This is shown in the simulation results presented in the next section.

6.4 Simulation results

Simulation experiments were conducted in order to evaluate in practice the performance of the above algorithm. For different values of the initial capital u and the ceiling L, 10000 iterations were performed and 1000 claim epochs were created. Positive loading values of $\rho = 0.03$, 0.05, and 0.1 were considered. In all cases, the time and reserve axes where scaled so that $c = \lambda = 1$. Experiments were conducted for two different claim size distributions as follows

1. The exponential distribution with c.d.f.

   $$F(x) = 1 - e^{-x/\mu}, \;\; x > 0$$

   and mean $E(X) = \mu$, so that μ was set equal to 0.97, 0.952, and 0.909 respectively in order to have the above values for ρ. Results are shown in Table 1.

2. The Pareto distribution with c.d.f.

   $$F(x) = 1 - \left(1 + \frac{x}{b}\right)^{-a}, \;\; x > 0$$

   where $EX = \frac{b}{a-1}$ and $a > 1$. Using $a = 1.1$, 1.5, 2, 5, and 10 we obtain various values for b in order to have the above values for ρ. Results are tabulated in Tables 2,3, and 4.

It is worth noting that the smoothed estimator we propose outperforms the crude Monte Carlo estimator often by an order of magnitude or more in terms of its variance, particularly in the case of the Pareto distribution which has heavy tails.

u	L	μ	E[M(t)]	E[K(t)]	Var[M(t)]	Var[K(t)]
5	**7**	0.97	0.034	0.0336	0.03554	0.00077
5	**8**	0.97	0.0112	0.0119	0.01167	0.0001
5	**10**	0.97	0.0015	0.0015	0.0014979	0.00000163
5	**11**	0.97	0.0007	0.00053	0.000699	0.0000002
8	**10**	0.97	0.0015	0.0015	0.0015	0.00000162
8	**11**	0.97	0.0006	0.000584	0.000599	0.00000021
10	**12**	0.97	0.0001	0.00016	0.0001	0.00000003
5	**7**	0.952	0.0232	0.0226	0.02386	0.00041
5	**8**	0.952	0.0031	0.0028	0.0033	0.00000625
5	**9**	0.952	0.0024	0.0027	0.00259	0.0000061
5	**10**	0.952	0.0009	0.0009	0.000899	0.00000075
5	**11**	0.952	0.0001	0.0003	0.0001	0.00000009
5	**12**	0.952	0.0001	0.00011	0.0001	0.00000001
8	**10**	0.952	0.0012	0.001	0.001198	0.00000077
8	**11**	0.952	0.0004	0.00034	0.00039988	0.00000009
10	**12**	0.952	0.0001	0.00011	0.0001	0.00000001
5	**7**	0.909	0.0093	0.0093	0.01	0.00007756
5	**8**	0.909	0.0032	0.00314	0.00319	0.00000854
5	**9**	0.909	0.0029	0.00276	0.002891	0.00000607
5	**10**	0.909	0.0002	0.0003	0.0001999	0.00000001
5	**11**	0.909	0.0002	0.0001	0.0001999	0.00000001
8	**10**	0.909	0.0001	0.0003	0.0001	0.00000011
8	**11**	0.909	0.0002	0.0001	0.0001998	0.00000001

Tab. 6.1: Exponential distribution.

[1] Zero entries here and elsewhere for the crude Monte Carlo estimator signify that in 10,000 iterations no claims resulted in time in the red i.e. all sample paths were strictly positive. Thus both estimates for the mean $EM(t)$ and for the variance $\mathsf{Var}(M(t))$ are in these cases zero. For the smoothed Monte Carlo estimator zero entries mean that the corresponding values are equal to zero to eight significant digits.

u	L	a	b	E[M(t)]	E[K(t)]	Var[M(t)]	Var[K(t)]
10	**15**	1.1	0.097	0.3815	0.3733	0.4874	0.013724
		1.5	0.485	0.933	0.925	1.4146	0.177
		2	0.97	0.6716	0.6714	0.937	0.1187
		5	3.883	0.0381	0.0424	0.03925	0.0008676
		10	8.73	0.0039	0.0037	0.003885	0.00000798
10	**20**	1.1	0.097	0.338	0.335	0.438935	0.009738
		1.5	0.485	0.7236	0.7286	1.0321	0.09812
		2	0.97	0.4756	0.4693	0.57386	0.05045
		5	3.883	0.016	0.016	0.01673	0.0001
		10	8.73	0.0004	0.0006	0.00039988	0.0000002
10	**40**	1.1	0.097	0.25	0.25	0.31048	0.004027
		1.5	0.485	0.4135	0.4139	0.49916	0.020125
		2	0.97	0.18	0.18	0.19036	0.005117
		5	3.883	0.0011	0.0011	0.0011	0.00000044
		10	8.73	0	0.00000124	0	0
10	**50**	1.1	0.097	0.23	0.23	0.2625	0.00293
		1.5	0.485	0.334	0.333	0.37468	0.01107
		2	0.97	0.138	0.138	0.1446	0.00226
		5	3.883	0.0003	0.0004	0.0003	0.00000007
		10	8.73	0	0	0	0
20	**50**	1.1	0.097	0.229	0.229	0.256	0.003
		1.5	0.485	0.33	0.33	0.3825	0.011
		2	0.97	0.137	0.138	0.14317	0.00224
		5	3.883	0.0003	0.0004	0.0003	0.00000006
		10	8.73	0	0	0	0
20	**75**	1.1	0.097	0.188	0.188	0.207	0.00155
		1.5	0.485	0.23	0.227	0.25	0.0035
		2	0.97	0.075	0.076	0.0765	0.00048
		5	3.883	0	0	0	0
		10	8.73	0	0	0	0
50	**80**	1.1	0.097	0.18	0.18	0.2	0.00138
		1.5	0.485	0.2078	0.21	0.2244	0.003
		2	0.97	0.0702	0.069	0.072	0.00036
		5	3.883	0.0685	0.069	0.0686	0.00036
		10	8.73	0	0	0	0
50	**100**	1.1	0.097	0.16	0.16	0.187	0.00096
		1.5	0.485	0.1739	0.1688	0.1818	0.00153
		2	0.97	0.0509	0.0494	0.0513	0.00014
		5	3.883	0	0	0	0
		10	8.73	0	0	0	0

Tab. 6.2: Pareto with $\rho = 0.03$

u	L	a	b	$\mathbf{E}[\mathbf{M}(\mathbf{t})]$	$\mathbf{E}[\mathbf{K}(\mathbf{t})]$	$\mathbf{Var}[\mathbf{M}(\mathbf{t})]$	$\mathbf{Var}[\mathbf{K}(\mathbf{t})]$
10	**15**	1.1	0.095	0.36	0.36	0.453	0.0121
		1.5	0.476	0.83	0.84	1.2158	0.1534
		2	0.952	0.5797	0.5865	0.807	0.0979
		5	3.81	0.0329	0.0323	0.034	0.00058
		10	8.57	0.0027	0.0025	0.00269298	0.00000433
10	**20**	1.1	0.095	0.325	0.323	0.417	0.00913777
		1.5	0.476	0.68	0.68	0.9627	0.08529263
		2	0.952	0.4123	0.4098	0.50055	0.04075988
		5	3.81	0.0116	0.0119	0.0116	0.00007
		10	8.57	0.0004	0.0004	0.0003998	0.00000012
10	**40**	1.1	0.095	0.24	0.24	0.27988	0.00359
		1.5	0.476	0.38	0.38	0.461	0.017
		2	0.952	0.173	0.168	0.18575	0.0043
		5	3.81	0.0007	0.0009	0.0006995	0.0000003
		10	8.57	0	0	0	0
20	**50**	1.1	0.095	0.2198	0.22	0.251713	0.0026
		1.5	0.476	0.3008	0.31	0.34035	0.00996232
		2	0.952	0.126	0.124	0.13153	0.0019
		5	3.81	0.0003	0.0003	0.00029994	0.00000005
		10	8.57	0	0	0	0
50	**100**	1.1	0.095	0.1587	0.1572	0.16713	0.0008553
		1.5	0.476	0.1559	0.1585	0.159611	0.00137
		2	0.952	0.0457	0.045	0.04721	0.000132
		5	3.81	0.0001	0.00006	0.0001	0
		10	8.57	0	0	0	0

Tab. 6.3: Pareto with $\rho = 0.05$

u	L	a	b	E[M(t)]	E[K(t)]	Var[M(t)]	Var[K(t)]
10	**15**	1.1	0.0909	0.339	0.333	0.432	0.0103
		1.5	0.4545	0.692	0.6886	1.02624	0.1015
		2	0.909	0.416	0.416	0.5453	0.052055
		5	3.636	0.0153	0.0162	0.01506	0.00017677
		10	8.1818	0.0012	0.00118	0.0012	0.00000105
10	**20**	1.1	0.0909	0.295	0.297	0.357	0.00692
		1.5	0.4545	0.549	0.553	0.7316	0.05757
		2	0.909	0.2884	0.2922	0.34946	0.02306
		5	3.636	0.0057	0.0062	0.005668	0.0000215
		10	8.1818	0.0001	0.0001	0.0001	0.00000003
10	**40**	1.1	0.0909	0.231	0.227	0.26768	0.003
		1.5	0.4545	0.314	0.317	0.37496	0.0122
		2	0.909	0.134	0.128	0.14315	0.002733
		5	3.636	0.0003	0.0004	0.0003	0.0000001
		10	8.1818	0	0	0	0
10	**50**	1.1	0.0909	0.2048	0.2065	0.2246	0.002235
		1.5	0.4545	0.263	0.262	0.30026	0.007088
		2	0.909	0.095	0.095	0.09758	0.001275
		5	3.636	0.0001	0.0002	0.0001	0.00000002
		10	8.1818	0	0	0	0
20	**50**	1.1	0.0909	0.203	0.205	0.235	0.002084
		1.5	0.4545	0.255	0.262	0.2767	0.00707269
		2	0.909	0.0996	0.0954	0.1043	0.00125
		5	3.636	0.0002	0.00021	0.0002	0.00000002
		10	8.1818	0	0	0	0
20	**75**	1.1	0.0909	0.165	0.169	0.17986	0.0011758
		1.5	0.4545	0.188	0.182	0.19958	0.00241
		2	0.909	0.052	0.054	0.053	0.0003
		5	3.636	0	0	0	0
		10	8.1818	0	0	0	0
50	**80**	1.1	0.0909	0.167	0.164	0.17479	0.00103
		1.5	0.4545	0.163	0.169	0.17598	0.00199
		2	0.909	0.05	0.05	0.0528	0.00022
		5	3.636	0	0	0	0
		10	8.1818	0	0	0	0
50	**100**	1.1	0.0909	0.143	0.146	0.1527	0.000712
		1.5	0.4545	0.137	0.136	0.1473	0.00107
		2	0.909	0.033	0.035	0.0327	0.00009547
		5	3.636	0.0001	0.0001	0.0001	0
		10	8.1818	0	0	0	0

Tab. 6.4: Pareto with $\rho = 0.1$

BIBLIOGRAPHY

[1] Albrecher, H. and O. J. Boxma. (2004). A ruin model with dependence between claim sizes and claim intervals. *Insurance: Mathematics and Economics* **35** (2), 245-254.

[2] Albrecher, H. Borst, S.C., Boxma, O.J. and J. Resing. (2011). Ruin excursions, the $G/G/\infty$ queue, and tax payments in renewal risk models. *J. Appl. Prob.* Spec. Vol. **48A**, 3–14.

[3] Aldous, D. (1989). *Probability Approximations via the Poisson Clumping Heuristic.* Applied Mathematical Sciences 77, Springer Verlag.

[4] Anick D., D. Mitra and M.M. Sondhi (1982). Stochastic theory of a data handling system with multiple sources. *Bell System Tech. J.* **61**, 1871-1894.

[5] Applebaum, D. (2009). *Lévy Processes and Stochastic Calculus,* 2nd Ed, Cambridge University Press.

[6] Asmussen, S. (1984). Approximations for the probability of ruin within finite time. *Scandinavian Actuarial Journal,* 31-57.

[7] Asmussen, S. (2003). *Applied Probability and Queues,* 2nd Edition, Springer Verlag.

[8] Asmussen, S. and H. Albrecher (2010). *Ruin Probabilities.* 2nd edition, World Scientific, Singapore.

[9] Asmussen, S. and Kluppelberg, C. (1996). Large Deviation Results for Subexponential Tails, with Applications to Insurance Risk. *Stoch. Proc. Appl.,* **64,** 103-125.

[10] Asmussen S., Schmidli H. and V. Schmidt (1999). Tail Probabilities for Non-Standard Risk and Queueing Processes with Subexponential Jumps. *Advances in Applied Probability* **31**, 2, 422–447.

[11] Asmussen, S. and M. Steffensen. (2020). *Risk and Insurance. A Graduate Text.* Springer.

[12] Baccelli, F. and Brémaud, P. (2003). *Elements of Queueing,* 2nd edition. Springer–Verlag.

[13] Bachelier, L. (1900). Theorie de la Speculation. *Ann. Sci. Ecole Norm. Sup.* **III-17,** 21-86.

[14] Beard, R.E., Pentikäinen, T. and Pesonen, R. (1984). *Risk Theory,* 3rd edition. Chapman and Hall, London.

[15] Barndorff-Nielsen O.E. and N. Shephard, (2001). Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics, *J. Roy. Statist. Soc. Ser. B Statist. Methodol.* **63** (2), pp. 167–241.

[16] Beirlant, J. and Teugels, J.L. (1992). Modeling Large Claims in Non-life Insurance. *Insurance: Math. Econom.,* **11**, 17-29.

[17] Bichsel, F. (1964). Erfahrungstarifierung in der Motorfahrzeughaftpflichtversicherung. MVSV 64, 119-129.

[18] Billingsley, P. (1999). *Convergence of Probability Measures,* 2nd Edition, John Wiley.

[19] Bingham, N.H., Goldie, C.M., and Teugels, J.L. (1987). *Regular Variation,* Cambridge University Press.

[20] Bjork, T. and Grandell, J. (1988). Exponential inequalities for ruin probabilities in the Cox case. *Scandinavian Actuarial Journal,* 77-111.

[21] Boogaert P. and J. Haezendonck. (1989). Delay in claim settlement. *Insurance: Mathematics and Economics,* **8**, 321–330.

[22] D.D. Botvich, N.G. Duffield. (1995). Large deviations, the shape of the loss curve, and economies of scale in large multiplexers. *Queueing Systems*, **20**, 293–320.

[23] Bowers, N.L., Gerber, H.U., Hickman, J.C., Jones, D.A. and Nesbitt, C.J. (1987). *Actuarial Mathematics.* The Society of Actuaries, Schaumburg, Illinois.

[24] Breiman, L. (1967). First exit times from a square root boundary. In *Proc. 5th Berkeley Symp. Math. Statist. Prob.,* **2**, 9–16.

[25] Breiman, L. (1968). *Probability,* Addison-Wesley.

[26] Brémaud, P. (2000). An insensitivity property of Lundberg's estimate for delayed claims. *Journal of Applied Probability,* **37**, 3, 914–917

[27] Bucklew, J.A. (1991). *Large Deviation Techniques in Decision, Simulation and Estimation,* John Wiley.

[28] Bühlmann, H. (1970). *Mathematical Methods in Risk Theory.* Springer, Berlin.

[29] P. Carmona, F. Petit, M. Yor, (2001). Exponential functionals of Lévy processes, in: O.E. Barndorff-Nielsen, T. Mikosch, S. Resnick (Eds.), *Lévy Processes, Theory and Applications,* pp. 41–55, Birkhäuser, Boston.

[30] Chang, C-S. (1994). Stability, Queue Length, and Delay of Deterministic and Stochastic Queueing Networks, *IEEE Transactions on Automatic Control,* **39**, 913–931.

[31] Chung, K.L. (1974). *A Course in Probability Theory, 2nd edition.* Academic Press, New York.

[32] Cramér, H. (1930). *On the Mathematical Theory of Risk.* Skandia Jubilee Volume, Stockholm.

[33] Cramér, H. (1954). *The Elements of Probability Theory and Some of its Applications.* Almqvist och Wiksell, Uppsala, and Wiley, New York. Reprint: Ed. Huntington: Krieger 1973.

[34] David, F. N., and Fix, E. (1961). In *Proceedings of the Fourth Berkeley Symposium,* **1** , 177-197, Berkeley, California: University of California Press.

[35] Daykin, C.D., Pentikäinen, T. and Pesonen, M. (1994). *Practical Risk Theory for Actuaries.* Chapman and Hall, London.

[36] Debicki K. and M. Mandjes (2003). Exact overflow asymptotics for queues with many Gaussian inputs *J. Appl. Probab.* **40** (3) 704–720.

[37] Debicki, K. and T. Rolski (1995). A Gaussian fluid model. Queueing Systems **20**, 433-452.

[38] De Finetti, B. (1957). Su un' impostazione alternativa della teoria colletiva del rischio. *Transactions of the XV International Congress of Actuaries,* **2,** 433-443.

[39] Delbaen, F. and Haezendonck, J.M. (1987). Classical risk theory in an economic environment: *Insurance: Math. Econom.,* **6**, 85-116.

[40] Delbaen, F. and Haezendonck, J.M. (1989). A martingale approach to premium calculation principles in an arbitrage free market: *Insurance: Math. Econom.,* **8**, 269-277.

[41] Dembo, A. and Zeitouni, O. (1997). *Large Deviations Techniques and Applications,* 2nd Edition, Springer Verlag.

[42] Downton, F. (1970). Bivariate Exponential Distributions in Reliability Theory. *Journal of the Royal Statistical Society, Series B,* **32**(3), 408-417.

[43] Krzysztof Debicki. (2002). Ruin probability for Gaussian integrated processes. *Stochastic Processes and their Applications,* **98**, 151–174.

[44] Krzysztof Debicki and Michel Mandjes. (2003). Exact overflow asymptotics for queues with many gaussian inputs, *J. Appl. Prob.* **40**, 704–720.

[45] Krzysztof Debicki and Tomasz Rolski. (1995). A Gaussian fluid model. *Queueing Systems,* **20**, 433–452.

[46] Dickson, D.C.M. and Egidio dos Reis, A.D. (1996). On the distribution of the duration of negative surplus. *Scandinavian Actuarial Journal,* **2**, 148-164.

[47] dos Reis, Egidio, A.D. (1993). How long is the surplus below zero? *Insurance: Mathematics and Economics* **12,** 23-38.

[48] Duffield, N.G., and O'Connell, N. (1995). Large Deviations and Overflow Probabilities for the General Single-server Queue, with Applications, *Math. Proc. Camb. Phil. Soc.,* **118**, 363-374.

[49] Dufresne, D. and Gerber, H.U. (1991). Risk theory for the compound Poisson process that is perturbed by diffusion. *Insurance: Math. Econom.,* **10**, 51–59.

[50] Durbin, J. (1985). The first passage density of a continuous Gaussian process to a general boundary. *J. Appl. Probab.,* **22**, 431-453.

[51] Dussauchoy, A. and Berland, R. (1975). A multivariate gamma distribution whose marginal laws are gamma, in *Statistical Distributions in Scientific Work,* **1** (G. P. Patil, S. Kotz, and J. K. Ord, eds.), 319-328, Dordecht, The Netherlands: D. Reidel.

[52] Embrechts, P., Grandell, J. and Schmidli, H. (1993). Finite-time Lundberg inequalities in the Cox case. *Scandinavian Actuarial Journal,* 17-41.

[53] Embrechts, P. and Kluppelberg, C. (1993). Some Aspects of Insurance Mathematics. *Theory Probab. Appl.,* **38,** 374-416.

[54] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modeling Extremal Events,* Springer, Berlin.

[55] Embrechts, P. and Veraverbeke, N. (1982). Estimates for the Probability of Ruin with Special Emphasis on the Possibility of Large Claims. *Insurance: Math. Econom.,* **1,** 55-72.

[56] Ethier, S.N. and T.G. Kurtz (1986). *Markov processes: Characterization and Convergence.* John Wiley and Sons.

[57] Furrer, H. (1998). Risk Processes Perturbed by a-stable Levy Motion. *Scand. Actuar. J.,* **1,** 59-74.

[58] Furrer, H., Michna, Z. and Weron, A. (1997). Stable Levy Approximation in Collective Risk Theory. *Insurance Math. Econ.,* **20**, 97-114.

[59] Furrer, H.J. and Schmidli, H. (1994). Exponential inequalities for ruin probabilities of risk processes perturbed by diffusion, *Insurance: Math. Econom.,* **15,** 23-36.

[60] Ganesh, A., Neil O'Connell and Damon Wischik. (2004). *Big Queues,* Lecture Notes in Mathematics, 1838, Springer.

[61] Gaenssler P. (1983). *Empirical Processes.* IMS Lecture Notes - Monograph Series.

[62] Gaenssler P. and W. Stute. (1979). Empirical Processes: A survery of results for independent and identically distributed random variables. The Annals of Probability, **7**, 2, 193–243.

[63] Gerber, H.U. (1970). An extension of the renewal equation and its application in the collective theory of risk. *Scandinavian Actuarial Journal,* 1970, 3-4, 205–210.

[64] Gerber, H. U. (1979). *An Introduction to Mathematical Risk Theory* S.S. Huebner Foundation Monograph Series No.8 Distributed by R. Irwin, Homewood, Il.

[65] Gerber, H.U. (1990). *Life Insurance Mathematics.* Springer, Berlin.

[66] Gerber, H.U. (1990). When does the surplus reach a given target? *Insurance: Mathematics and Economics* **9,** 115-119.

[67] Glynn, P. W. and W. Whitt (1993). Limit theorems for cumulative processes. *Stochastic Processes and their Applications* **47** (2), 299-314.

[68] Glynn, P.W. and W. Whitt (1994). Logarithmic asymptotics for steady-state tail probabilities in a single-server queue. *Studies in Appl. Prob.*, 131-156.

[69] Gnedenko, B. (1969). *Theory of Probability,* Mir Publishers, Moscow.

[70] Gnedenko, B.V. and Kolmogorov, A.N. (1954). *Limit distributions for sums of independent random variables,* Addison-Wesley, Massachusetts.

[71] Boris V. Gnedenko B.V. and V.Yu. Korolev. (1996). *Random Summation Limit Theorems and Applications,* CRC Press, Boca Raton.

[72] Grandell, J. (1977). A Class of Approximations of Ruin Probabilities. *Scand. Actuar. J.,* 37-52.

[73] Grandell, J. (1991). *Aspects of Risk Theory,* Springer-Verlag.

[74] Greenwood, P.E. and Perkins, E. (1983). A conditional limit theorem for random walk, and Brownian local time on square root boundaries. *Ann. Probab.,***11**, 227-261.

[75] Hadwiger, H. (1940). Über die Wahrscheinlichkeit des Ruins bei einer grossen Zahl von Geschäften. *Archiv für mathematische Wirtschafts- und Sozialforschung,* **6**, 131-135.

[76] D. Heath, S. Resnick, and G. Samorodnitsky. (1998). Heavy Tails and Long Range Dependence in On/Off Processes and Associated Fluid Models, *Mathematics of Operations Research,* **23** (1), 145-165.

[77] Hurst, H.E. (1950). Long-Term Storage Capacity of Reservoirs. *Proc. Amer. Soc. Civil Eng.*, **76,** 11, 770-808.

[78] Iglehart, D. (1969). Diffusion Approximations in Collective Risk Theory. *J. Appl. Probab.,* **6,** 285-292.

[79] Johnson, N.L. (1960). An approximation to the multinomial distribution: some properties and applications, *Biometrika,* **47,** 93-103.

[80] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions,* Vol. 1, second edition, New York: John Wiley & Sons.

[81] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions,* Vol. 2, second edition, New York: John Wiley & Sons.

[82] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1997). *Discrete Multivariate Distributions,* New York: John Wiley & Sons.

[83] Johnson, N.L., Kotz, S. and Kemp, A. W. (1992). *Univariate Discrete Distributions,* second edition, New York: John Wiley & Sons.

[84] Jurek, Z.J. and W. Vervaat (1983). An Integral Representation for Selfdecomposable Banach Space Valued Random Variables. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **62**, 247–262.

[85] Karatzas, I. and Shreve, S.E. (1988). *Brownian Motion and Stochastic Calculus.* Springer, Berlin.

[86] Karlin, S. and Taylor, H.M. (1981). *A Second Course in Stochastic Processes.* Academic Press, New York.

[87] Kesidis, G., Walrand, J. and Chang, C.S. (1993). Effective Bandwidths for Multiclass Markov Fluids and Other ATM Sources, *IEEE/ACM Trans. Networking*, **1**, No. 4, 424-428.

[88] Kibble, W.F. (1941). A two-variate gamma type distribution, *Sankhya,* **5,** 137-150.

[89] Klüppelberg, C., and Mikosch, T. (1995). Explosive Poisson Shot Noise Processes with Applications to Risk Reserves. *Bernoulli,* **1,** 125-147.

[90] Klüppelberg, C., and Mikosch, T. (1995). Modeling Delay in Claim Settlement. *Scandinavian Actuarial Journal,* 154-168.

[91] Klüppelberg, C., and Mikosch, T. (1997). Large Deviations of Heavy-tailed Random Sums with Applications in Insurance and Finance. *J. Appl. Probab.,* **34**, 293-308.

[92] Klüppelberg, C., and Mikosch, T. (1998). Delay in Claim Settlement and Ruin Probability Approximation, *Scandinavian Actuarial Journal,* **2**, 154–168.

[93] Konstantopoulos, T. and Si-Jian Lin (1998). Macroscopic models for long-range dependent network traffic. *Queueing Systems*, **28**, 215–243.

[94] Korevaar, J. (2004). *Tauberian Theory. A Century of Developments.* Springer Verlag.

[95] Kotz, S., Balakrishnan, N., and Johnson, N.L. (2000). *Continuous Multivariate Distributions: Models and Applications*, New York: John Wiley & Sons.

[96] Kulkarni, V. and Rolski, T. (1994). Fluid model driven by an Ornstein-Uhlenbeck process, *Probability in Engineering and Informational sciences,* **8,** 403-417.

[97] Lerche, H. R. (1986). *Boundary Crossing of Brownian Motion,* Lecture Notes in Statist. 40. Berlin: Springer-Verlag.

[98] Lindner, A. and Maller, R. (2005). Lévy integrals and the stationarity of generalised Ornstein–Uhlenbeck processes *Stochastic Processes and their Applications*, **115**, 170–1722.

[99] Lindvall, T. (1973). Weak convergence of probability measures and random functions in the function space $\mathbb{D}(0,\infty)$, *J. Appl. Prob.,* **10**, 109-121.

[100] Lundberg, F. (1903). *Approximerad Framstallning av Sannolikhetsfunktionen. Aterforsakring av Kollektivrisker.* Akad. Afhandling. Almqvist och Wiksell, Uppsala.

[101] Mandjes, M. (2007). *Large Deviations for Gaussian Queues: Modelling Communication Networks*, John Wiley.

[102] Makatis, G. and M.A. Zazanis (2004). Smoothed Monte Carlo Estimators for the Time-in-the-Red in Risk Processes, *Mathematical Methods of Operations Research,* **59** (2) 329-342.

[103] Macci, C., Stabile, G., and Torrisi G. L. (2005). Lundberg parameters for non standard risk processes. *Scandinavian Actuarial Journal*, **6**, 417–432.

[104] Marcus, M.B. and Shepp, L.A. (1972). Sample behaviour of Gaussian processes. *Proceedings of the Sixth Berkeley Symposium.*

[105] Mardia, K. V. (1962). Multivariate Pareto distributions, *Annals of Mathematical Statistics,* **33,** 1008-1015.

[106] Métivier, M. (1984). *Stochastic Processes and Stochastic Integration.* deGruyter, Amsterdam.

[107] Mikosch, T. (2004). *Non-Life Insurance Mathematics. An Introduction with Stochastic Processes*, Springer Verlag.

[108] Mikosch, T., Resnick, S., Rootzen, H. and A. Stegeman, (2002). Is network traffic approximated by stable Lévy motion or fractional brownian motion? *The Annals of Applied Probability,* **12** (1), 23–68.

[109] T. Mikosch and G. Samorodnitsky, (2007). Scaling Limits for Cumulative Input Processes, *Mathematics of Operations Research,* **32** (4), 890–919.

[110] Moran, P.A.P. (1967). Testing for correlation between non-negative variates, *Biometrika,* **54,** 385–394.

[111] Müller A. and G. Pflug. (2001). Asymptotic ruin probabilities for risk processes with dependent increments. *Insurance: Mathematics and Economics* **28**, 381–392.

[112] Neuhaus, W. (1992). IBNR models with random delay distributions, *Scandinavian Actuarial Journal*, (2) 97–107.

[113] Niculescu, Constantin P. and Lars-Erik Persson (2006). *Convex Functions and Their Applications: A Contemporary Approach.* Springer.

[114] Norros, I., Roberts, J.W., Simonian, A. and Virtamo, J. (1991). The superposition of variable bitrate sources in ATM multiplexers. *IEEE J. Selected Areas in Commun.* **9**, 378-387.

[115] Norberg, R. (1986). A Contribution to Modelling of IBNR Claims *Scandinavian Actuarial Journal,* 155–2030.

[116] Palmowski, Z. and T. Rolski (1998). The superposition of Alternating On–Off Flows and a Fluid Model. *The Annals of Applied Probability*, **8**(2) 524-540.

[117] G. Pang, J. Yang, and Y. Zhou (2017). On functional limit theorems for the cumulative times in alternating renewal processes *Operations Research Letters*, **45**(2), 145-150.

[118] Pollard, D. (1984). *Convergence of Stochastic Processes,* Springer, Berlin.

[119] Revuz, D. and Yor, M. (1991). *Continuous Martingales and Brownian Motion.* Springer Verlag, Berlin and New York.

[120] Rolski, T., Schmidli, H., Schmidt, V. and Teugels, J. L. (1999). *Stochastic Processes for Insurance and Finance.* Wiley, Chichester.

[121] Ross, S.M. (1983). *Stochastic Processes.* Wiley, New York.

[122] Ross, S.M. (1990). *A Course in Simulation.* Macmillan.

[123] Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Cambridge, U.K.

[124] Schmidli, H. (1994). Diffusion approximations for a risk process with the possibility of borrowing and investment. *Commun. Statist.– Stochastic Models*, **10** 2, 365–388.

[125] Schmidli, H. (1995). Cramér–Lundberg approximations for ruin probabilities of risk processes perturbed by diffusion. *Insurance: Mathematics and Economics,* **16**, 135–149.

[126] Shepp L.A. (1971). First passage time for a particular gaussian process. *The Annals of Mathematical Statistics*, **42,** (3) 946-951.

[127] Shepp L.A. and D. Slepian (1976). First-passage time for a particular stationary periodic gaussian process. *J. Appl. Prob.* **13,** 27-38.

[128] Slepian D. (1961). First Passage Time for a Particular Gaussian Process *Annals of Mathematical Statistics,* **32,** (2) 610-612.

[129] Skorokhod, A.V. (1957). Limit theorems for stochastic processes, *Theory Prob. Appl.***1**, 261-290.

[130] Stoyan, D. (1984). *Stochastic Ordering.* John Wiley.

[131] Daniel W. Stroock, S. R. Srinivasa Varadhan (1997). *Multidimensional Diffusion Processes.* Springer, Berlin.

[132] Szczotka, W. (1980). Central limit theorem in $\mathbb{D}[0,\infty)$ for breakdown processes. *Probab. Math. Statist.* **1**, 49-57.

[133] Täecklind, S. (1942). Sur le risque de ruine daus les jeux inequitables. *Skand. Aktuar. Tidskr.,***1942,** 1-42.

[134] Taylor, G.C. (1979). Probability of ruin under inflationary conditions or under experience ratings, *Astin Bulletin*, **16,** 149-162.

[135] Taylor, G.C. (1979). Probability of ruin with variable premium rate, *Scand. Actuar. J.,* **1980,** 57-76.

[136] Tiao, G.C., and Guttman, I.J. (1965). The inverted Dirichlet distribution with applications, *Journal of the American Statistical Association,* **60,** 793-805; Correction, **60,**1251-1252.

[137] Thorin, O. (1974). On the Asymptotic Behaviour of the Ruin Probability for an Infinite Period when the Epochs of Claims Form a Renewal Process. *Scand. Actuar. J.,* 29-50.

[138] Thorin, O. (1982). Probability of Ruin. *Scand. Actuar. J.,* 65-102.

[139] Trufin, J., Albrecher, H. and M. Denuit. (2011). Ruin problems under IBNR dynamics. *Appl. Stochastic Models Bus. Ind.* **27**, 619–632.

[140] van der Vaart A. W. and J. A. Wellner. (2023). *Weak Convergence and Empirical Processes.* 2nd ed. Springer.

[141] Veraverbeke, N. (1993). Asymptotic Estimates for the Probability of Ruin in a Poisson Model with Diffusion. *Math. Econom.,* **13,** 57-62.

[142] Vere-Jones, D. (1967). The infinite divisibility of a bivariate gamma distribution, *Sankhya, Series A,* **27,** 421-422.

[143] Wagner, C. (2002). Time in the red in a two state Markov model. *Insurance: Mathematics and Economics*, **31**, 365-372.

[144] Whitt, W. (1980). Some useful functions for functional limit theorems, *Math. Oper. Res.* **5**, 67-85.

[145] Whitt W. (2000). Limits for cumulative input processes to queues. *Probability in the Engineering and Information Sciences* **14**, 123-150.

[146] Whitt W. (2003). *Stochastic Process Limits. An Introduction to Stochastic-Process Limits and Their Application to Queues.* Springer-Verlag, New York.

[147] Wicksell, S.D. (1933). On correlation functions of type III, *Biometrika,* **25,** 121-133.

[148] Wolfe S.J. (1982). On a continuous analogue of the stochastic difference equation $X_n = \rho X_{n-1} + B_n$. *Stoch. Proc. Appl.* **12**, 301–312.

[149] Walter Willinger, Murad S. Taqqu, Robert Sherman, Daniel V. Wilson (1997). Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level. *IEEE ACM Transactions on Networking.*

[150] Yuen, Kam C., Guo Junyi, and Kai W. Ng, (2005). On ultimate ruin in a delayed-claims risk model. *J. Appl. Prob.* **42**, 163–174.

[151] Zaphiropoulos, Y. and M.A. Zazanis. (2010). Discrete Time Risk Processes with After-Effects and Association, *Stochastic Models*, **26**, 1, 27–45.