Weak Approximation in Risk Theory

David R. Alexander

Department of Mathematics & Statistics
McGill University, Montreal
February 3, 1997

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfilment of the requirements of the degree of Master of Science.

0-612-29644-X

Abstract

The most natural stochastic models for describing the time evolution of the collective risk reserves of an insurance company are jump or point process models. However, there are difficulties in obtaining from such models explicit and tractable expressions for important quantities such as the probability of ruin and these have spawned the development of procedures to approximate point process models. In this thesis, the nature of weak approximations, as put forward by Iglehart (1969) and Furrer, Michna & Weron (1996), is examined closely with a view toward assessing their value. An interpretation of these approximation procedures is given and a method by which the value of weak approximations may be improved is suggested by considering their Lévy-Grigelionis-Jacod characteristics.

Resumé

Le plus naturel des modèles stochastiques servant à décrire l'évolution dans le temps d'un portefeuille de risques d'une compagnie d'assurances est le processus de sauts. Cependant, il est difficile d'obtenir pour ce type de modèles, des expressions explicites et traitables pour des quantités importantes telle que la probabilité de la ruine, ce qui a mené au développement de procédures d'approximation pour processus de sauts. Dans cette thèse, la nature faible des approximations, telle que soulignée par Iglehart (1969) et Furrer et al. (1996), est examinée en profondeur, avec le but d' établir leur valeur. Une interprétation de ces procédures d'approximations est donnée et une méthode servant à améliorer l'approximation est suggérée en considérant les characteristiques de Lévy-Grigelionis-Jacod.

Acknowledgements

First and foremost, I must express to my family my deepest appreciation of their continual and unconditional support. Thank you. I would like to thank Prof. José Garrido for providing me with the opportunity to learn, not just the particulars of Risk Theory, but of probability theory in general, and for his flexibility as the subject of this thesis deviated significantly from that originally planned. I would also like to thank Prof. William Anderson and Prof. Georg Schmidt for patiently coordinating my supervision by Prof. Garrido. I am indebted to Prof. José Garrido, Prof. Niky Kamran, NSERC, the Department of Mathematics and Statistics as well as the Faculty of Graduate Studies and Research at McGill University for providing financial assistance in various forms. Finally, I thank Hansjörg Furrer, Masoud Asgharian Dastenaei, Ruth Grossman, Bill Boshuck, Shaolin Li, Chunzhang Wu, Prof. John Galbraith, Prof. Daniel Stroock, and Prof. John Taylor for stimulating discussions both related and unrelated to this thesis.

Contents

Notation

$\mathbb{N} = \{1, 2, 3, \cdots\}$

$\mathbb{N}_0 = \mathbb{N} \cup \{0\}$

$\mathbb{Z}$ is the set of integers

$\mathbb{R}$ is the set of reals

$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$

$\mathbb{R} \setminus \{0\}$ is the set of reals except 0

$\mathbb{R}_+ = [0, \infty)$

$\overline{\mathbb{R}}_+ = [0, \infty]$

$(\Omega, \mathcal{F}, \mathbf{P})$ is a probability space

$(\Omega, \mathcal{F}, \mathbf{F}, \mathbf{P})$ is a filtered probability space

$\mathbf{F} = \{\mathcal{F}_t\}_{t \in \overline{\mathbb{R}}_+}$ is a filtration

$S_\alpha(\sigma, \beta, \mu)$ is the collection of all α-stable random variables

$S_\alpha(1, \beta, 0)$ is the collection of all standard α-stable random variables

$N(\mu, \sigma^2)$ is the collection of all Normal random variables

$N(0, 1)$ is the collection of all standard Normal random variables

$\mathcal{B}(\mathbb{R})$ is the Borel σ-algebra on $\mathbb{R}$

$\mathcal{B}(\tau)$ is the Borel σ-algebra generated by the topology τ

$\stackrel{\mathcal{D}}{=}$ means equal in distribution

$\stackrel{a.s.}{=}$ means equal $\mathbf{P}$-a.s.

$\mathbf{E}$ is the expectation operator with respect to $\mathbf{P}$

$\mathbf{V}$ is the variance operator with respect to $\mathbf{P}$

$\mathbf{Cov}$ is the covariance operator with respect to $\mathbf{P}$

i.i.d. means independent and identically distributed

$X \coprod Y$means the random variables X and Y are independent

$X \coprod \mathcal{F}$ means the random variable X and the σ-algebra $\mathcal{F}$ are independent

$\xrightarrow{\mathbf{P}\text{-a.s.}}$ means converges to $\mathbf{P}$-a.s.

$\xrightarrow{\mathbf{P}}$ means converges in probability $\mathbf{P}$

$\mathbf{P}\lim$ means the limit in probability $\mathbf{P}$

$\xrightarrow{\mathcal{D}}$ means converges in distribution

$\Rightarrow$ means converges weakly

$\lfloor x \rfloor$ means the floor of x

IIP means independent increment process

$SIIP$ means stationary and independent increment process

$x \wedge y$ means $\min\{x, y\}$

$x \vee y$ means $\max\{x, y\}$

$\mathcal{P}_{\mathbb{R}}$ is the set of Borel probability measures on $\mathbb{R}$

$\mathcal{M}_{\mathbb{R}}$ is the set of Borel measures on $\mathbb{R}$

$\mathcal{M}_{\mathbb{R}\setminus\{0\}}$ is the set of Borel measures on $\mathbb{R} \setminus \{0\}$

$\mathcal{M}_{\mathbb{R}}^{1} = \{M \in \mathcal{M}_{\mathbb{R}} : \int_{\mathbb{R}} 1 dM(x) < \infty\}$

$\mathcal{M}_{\mathbb{R}\setminus\{0\}}^{1} = \{M \in \mathcal{M}_{\mathbb{R}\setminus\{0\}} : \int_{\mathbb{R}\setminus\{0\}} 1 dM(x) < \infty\}$

$\mathcal{M}_{\mathbb{R}\setminus\{0\}}^{1\wedge|x|} = \{M \in \mathcal{M}_{\mathbb{R}\setminus\{0\}} : \int_{\mathbb{R}\setminus\{0\}} (1 \wedge |x|) dM(x) < \infty\}$

$\mathcal{M}_{\mathbb{R}\setminus\{0\}}^{1\wedge x^2} = \{M \in \mathcal{M}_{\mathbb{R}\setminus\{0\}} : \int_{\mathbb{R}\setminus\{0\}} (1 \wedge x^2) dM(x) < \infty\} =$ Lévy measures

Chapter 1

Risk Theory

A risk business is a commercial enterprise that operates under conditions of significant financial uncertainty, manifesting themselves in highly variable rates of income, expenditure, or both. For a risk business to be a successful enterprise, it must accurately analyze the nature of these uncertain conditions and adjust its operations to prepare for the possible occurrence of financially extreme events, which, if not prepared for, would ultimately ruin the business. A broad interpretation of what is meant by the subject area called Risk Theory would include such general concerns. More traditionally, Risk Theory has been a branch of actuarial mathematics concerned with analyzing the operation of insurance businesses and it is from this context that the subject of this thesis is drawn.

A non-life insurance business begins with some initial capital in a reserve fund and policy holders pay regular premiums into the fund in order to be eligible to make a claim if some specific contingency occurs such as fire, car accident, disability, or death. The expenditures associated with claims occur at random times and are of random magnitudes, resulting in sudden financial shocks to the fund. Premium payments are generally smaller and occur more frequently, resulting in a relatively stable rate of income. Life insurance businesses operate in mirrored contrast; after paying regularly into a fund over many years, building up equity or becoming vested, policy holders then draw a regular pension or life insurance annuity until the random occurrence of some policy terminating event. If, at this time, the amount actually paid to the policy holder is less than the amount expected to be paid then the difference is considered as positive income; it is money that no longer needs to be allocated to a particular policy holder and can therefore be freed up in the fund for general use. Here, the initial capital is the equity in the fund at the time payments commence, the expenditures are regular, but income is received in random amounts at random times. This mirror symmetry between the two types of insurance often allows techniques for analyzing non-life insurance businesses to be easily modified for the analysis of life insurance businesses and we therefore now restrict our attention to non-life insurance.

The initial approach to insurance business was that of Individual Risk Theory, first appearing in the mid-1700's, in which each insurance policy was analyzed on an individual basis. An individual's premium for a given period of time was required to be greater than the mathematical expectation of the individual's claims for that period, which was based on the business' experience with the individual. The risk to the insurance business was the deviation of the individual's claims from their expected value and it was recognized (Bernouilli (1738) and DeMoivre (1738)) that an insurance business would eventually be ruined if it failed to include a margin in its favor. In the absence of rigorous methods to quantify this risk, large premiums and conservative estimates of the individual's distribution of claims were employed. If the individual's claims were covered by the premiums charged, market competition compelled the return of a portion of the excess in the form of a dividend or a refund, the remainder being kept in reserve.

As insurance businesses acquired larger numbers of policy holders, their portfolios were, from the point of view of Individual Risk Theory, seen merely as a large collection of individual policies, each treated individually. Collective Risk Theory, on the other hand, viewed portfolios as aggregates of large numbers of independent individual risks. The total assets of the insurance portfolio were considered as a whole, attention being paid only to the income, the claim occurrence times, and the claim severities. The details of any individual's policy or information about which particular policies gave rise to claims is disregarded. If the number of policy holders was sufficiently large and their individual effects were sufficiently small, central limit theorem arguments were used to justify normal approximations to the aggregate distribution of claims over a given period. Considering the business as a collective provided a greater statistical sample and proved to be useful, providing a systematic method for assessing risk (Cramér 1930). However, it soon became apparent that better mathematical tools were needed for situations in which the individual effects were not small, the number of policy holders was not large, and there appeared to be a dynamical dependence on time. In 1903, before the development of a general theory of stochastic processes in the 1930's, Lundberg (1903, 1909) proposed the use of dynamic continuous time stochastic models to address these problems, introducing the idea of Dynamic Collective Risk Theory, which, since the 1930's, has developed significantly, stimulating and making use of many advancements in both statistics and the theory of stochastic processes. For a review of these developments see Janssen (1981).

Further restricting our attention to the Collective Risk Theory of non-life insurance businesses, we consider point processes as the natural models for insurance portfolios. The practical focus of this thesis is the problem of determining from models of this type the probability that the reserve fund becomes negative after some time, called the finite time ruin probability, and the probability that the reserve fund eventually becomes negative, called the ultimate ruin probability. However, point process models have practical computational problems. Even when significant and unrealistic simplifying assumptions are imposed, analytical expressions for the probabilities of ruin and distributions of stopping times, if they exist, are obtained through difficult arguments and are not very tractable for applications. As a means of obtaining mathematically tractable results, Iglehart (1969) showed that a properly defined sequence of such models converges weakly to a Wiener process, thereby enabling one to bring all the powerful tools of stochastic calculus and the corresponding analytical results for various stopping times to bear on the problem. Unfortunately, the performance of the Wiener process approximations is less than ideal. The Wiener process approximation is mediocre, particularly for skewed or heavy tailed claims distributions (Asmussen 1984). Furthermore, the Wiener process approximation is not applicable for infinite variance claims distributions. Recently, Furrer et al. (1996) have extended Iglehart's weak convergence argument to permit the approximation of point process models with highly skewed claims distributions, possibly with infinite variance, by α-stable Lévy processes. However, preliminary numerical results are not quite satisfactory and so a closer examination of weak approximations is required.

We begin this thesis by defining a general point process model and describe some illustrative specializations to indicate some of the analytical difficulties involved in such models. We then describe the theory of convergence of stochastic processes, as required by Iglehart's application and the more recent application of Furrer et al. Background material on α-stable distributions and α-stable processes is also provided. Both Iglehart's and Furrer et al's weak convergence arguments are examined in detail. An interpretation of both weak convergence arguments is provided and a special case of interest is pointed out. This interpretation, together with a closer look at infinite divisibility, provides some insights into weak limit approximations. Furthermore, this interpretation suggests that a deeper study of the structure of weak approximations in terms of their Lévy-Grigelionis-Jacod characteristics may prove fruitful for improving their quality and for finding procedures to statistically fit them to the point process model, an issue of importance not only in Risk Theory but in any application of point process models.

Chapter 2

Point Process Reserve Models

2.1 The Model

We propose a study of the following model $R = \{R_t\}_{t \in \mathbb{R}_+}$ for risk reserves with mixed portfolio composition consisting of $N(\Pi)$ sources of premium income and $N(\chi)$ sources of claims. For $1 \leqslant i \leqslant N(\Pi)$, $I^i = \{I_t^i\}_{t \in \mathbb{R}_+}$ is the accumulated income process due to premium payments from the i^{th} source. The payments have random magnitudes $\{\Pi_k^i\}_{k \in \mathbb{N}}$ and occur at random times $\{T_k^{\Pi^i}\}_{k \in \mathbb{N}}$ where $T_k^{\Pi^i} < T_{k+1}^{\Pi^i}$ for $k \in \mathbb{N}$. The counting process associated with $\{T_k^{\Pi^i}\}_{k \in \mathbb{N}}$ is defined by $N_t^{\Pi^i} = \sum_{k \in \mathbb{N}} 1_{\{T_k^{\Pi^i} \leqslant t\}}$, giving the number of payments during $[0, t]$ and so $I_t^i = \sum_{k=1}^{N_t^{\Pi^i}} \Pi_k^i$. Similarly, for $1 \leqslant j \leqslant N(\chi)$, $\{\chi_k^j\}_{k \in \mathbb{N}}$ and $\{T_k^{\chi^j}\}_{k \in \mathbb{N}}$ are the magnitudes and occurrence times of claims from the j^{th} source, $T_k^{\chi^j} < T_{k+1}^{\chi^j}$ for $k \in \mathbb{N}$, $N_t^{\chi^j} = \sum_{k \in \mathbb{N}} 1_{\{T_k^{\chi^j} \leqslant t\}}$ gives the number of claims in $[0, t]$, and $C_t^j = \sum_{k=1}^{N_t^{\chi^j}} \chi_k^j$ defines the accumulated claims process $C^j = \{C_t^j\}_{t \in \mathbb{R}_+}$. We consider empty sums to be zero. The portfolio consists of a finite number of policies and so from each source we observe a finite number of payments and claims during any bounded time interval $[0, t]$. Furthermore, payments and claims are finite in magnitude and so the processes $\{I^i\}_{i=1}^{N(\Pi)}$ and $\{C^j\}_{j=1}^{N(\chi)}$ are assumed to be almost surely finite on $[0, t]$, i.e., without explosion. Letting R_0 be the initial capital in the reserve fund, possibly random, we then define R to be the superposition of its constituent processes:

$$\begin{aligned} R_t &= R_0 + \sum_{i=1}^{N(\Pi)} I_t^i - \sum_{j=1}^{N(\chi)} C_t^j \\ &= R_0 + \sum_{i=1}^{N(\Pi)} \sum_{k=1}^{N_t^{\Pi^i}} \Pi_k^i - \sum_{j=1}^{N(\chi)} \sum_{k=1}^{N_t^{\chi^j}} \chi_k^j \qquad t \in \mathbb{R}_+ \end{aligned} \tag{2.1}$$

The processes $\{I^i\}_{i=1}^{N(\Pi)}$ and $\{C^j\}_{j=1}^{N(\chi)}$ are piecewise constant, right continuous, and finite over any bounded time interval. Therefore, the income and claims processes, and hence R, are càdlàg (sample paths are almost surely right continuous with finite left limits).

To formalize this in a way that permits the use of stochastic calculus, assume that we have a complete, filtered probability space $(\Omega, \mathcal{F}, \mathbf{F}, \mathbf{P})$ which satisfies the usual conditions, namely, the filtration $\mathbf{F} = \{\mathcal{F}_t\}_{t \in \overline{\mathbb{R}}_+}$ is right continuous $(\forall t \in \mathbb{R}_+, \mathcal{F}_t = \cap_{s>t} \mathcal{F}_s)$ where we take $\mathcal{F}_{\infty-} = \bigvee_t \mathcal{F}_t$, $\mathcal{F}_\infty = \mathcal{F}$, $\mathcal{F}$ is $\mathbf{P}$-complete, and $\mathcal{F}_0$ contains all of the $\mathbf{P}$-null sets of $\mathcal{F}$. We assume that all income and claims processes are $\mathbf{F}$-adapted and that R_0 is $\mathcal{F}_0$-measurable so that R is $\mathbf{F}$-adapted. By Kolmogorov's existence theorem (Billingsley 1995) we can always construct a $\mathbf{P}$-complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$ carrying the processes R, $\{I^i\}_{i=1}^{N(\Pi)}$ and $\{C^j\}_{j=1}^{N(\chi)}$, and the random variable R_0 (as well as any additional processes and variables one may require). Furthermore, we can generate the natural filtrations for each of the income and claims processes, augment each with the $\mathbf{P}$-null sets of $\mathcal{F}$, and generate a common filtration $\mathbf{F} = \{\mathcal{F}_t\}_{t \in \overline{\mathbb{R}}_+}$ from the union of the individual filtrations. Since all processes are càdlàg, $\mathbf{F}$ so generated is right continuous and $\mathcal{F}_0$ contains all $\mathbf{P}$-null sets of $\mathcal{F}$; this filtration is the smallest filtration containing all information about the probabilistic evolution of all processes and their interdependence. A space $(\Omega, \mathcal{F}, \mathbf{F}, \mathbf{P})$ constructed in this manner satisfies the usual hypotheses. Thus, we assume hereafter that some model $(\Omega, \mathcal{F}, \mathbf{F}, \mathbf{P})$ has been constructed carrying all processes and variables under consideration and satisfying the usual hypotheses.

To ensure that $\{I^i\}_{i=1}^{N(\Pi)}$ and $\{C^j\}_{j=1}^{N(\chi)}$ are without explosion, it is sufficient to require that the counting processes $\{N^{\Pi^i}\}_{i=1}^{N(\Pi)}$ and $\{N^{\chi^j}\}_{j=1}^{N(\chi)}$ are without explosion. since premiums and claims have finite magnitudes. Since $\{I^i\}_{i=1}^{N(\Pi)}$ and $\{C^j\}_{j=1}^{N(\chi)}$ are càdlàg , $\{T_k^{\Pi^i}\}_{k\in\mathbf{N}}$ and $\{T_k^{\chi^j}\}_{k\in\mathbf{N}}$ are $\mathbf{F}$-stopping times for any $1 \leqslant i \leqslant N(\Pi)$ and $1 \leqslant j \leqslant N(\chi)$. Defining the $\mathbf{F}$-stopping time $T_\infty^{\Pi^i} = \sup_k T_k^{\Pi^i}$, the explosion time of the i^{th} premium counting process N^{Π^i}, we see that $\{(t,\omega) : N_t^{\Pi^i}(\omega) = \infty\} = \{(t,\omega) : t \in [T_\infty^{\Pi^i}(\omega), \infty)\}$. If $\mathbf{P}\{T_\infty^{\Pi^i} < \infty\} > 0$ then, for some $n \in \mathbb{N}$, $\mathbf{P}\{T_\infty^{\Pi^i} < n\} > 0$ and hence $\mathbf{P}\{N_n^{\Pi^i} = \infty\} > 0$. To have $\mathbf{P}\{N_n^{\Pi^i} = \infty\} = 0$ for all $n \in \mathbb{N}$ we insist that $\mathbf{P}\{T_\infty^{\Pi^i} < n\} = 0$ for all $n \in \mathbb{N}$. Thus, we adopt the assumption that $T_\infty^{\Pi^i} \overset{a.s.}{=} T_\infty^{\chi^j} \overset{a.s.}{=} \infty$ for $1 \leqslant i \leqslant N(\Pi)$ and $1 \leqslant j \leqslant N(\chi)$ to avoid explosions.

The model (2.1) is very general but reflects the essential qualities of the phenomenon: continuous time evolution, discretely occurring events, and finite magnitudes. No assumptions have been made on $\{\Pi_k^i\}_{k\in\mathbb{N}}$, $\{\chi_k^j\}_{k\in\mathbb{N}}$, $\{T_k^{\Pi^i}\}_{k\in\mathbb{N}}$, or $\{T_k^{\chi^j}\}_{k\in\mathbb{N}}$ about their distributions or interdependence. $\{\Pi_k^i\}_{k\in\mathbb{N}}$ and $\{\chi_k^j\}_{k\in\mathbb{N}}$ are not assumed to be $\mathcal{F}_0$-measurable; doing so is equivalent to assuming their distributions do not depend on time of occurrence. Assuming only that the income and claims processes are $\mathbf{F}$-adapted allows for time dependent distributions and thereby includes processes with conditionally independent increments such as martingales (Gerber 1979) or Cox processes (Grandell 1991) now being actively investigated in Risk Theory. Thus, (2.1) appears to be the most natural framework in which to describe risk reserves as well as other financial and economic processes.

Any application of (2.1) would require a statistical analysis of past claims data, the formulation of parametric models for claims processes, the estimation of all relevant parameters, and lastly, a decision about what premiums to charge and how often to collect them. The focus of attention here is how to decide on a premium policy for a model of the form (2.1) assuming the form of the claims processes have been specified. A quantity of key practical importance is the probability that R becomes negative during some time interval $[0,T]$ or $[0,\infty)$; this occurs when a claim exceeds available reserves, causing the financial ruin of the business. Define the $\mathbf{F}$-stopping time "time to ruin" by

$$T^r = \inf\{t > 0 : R_t < 0\} \tag{2.2}$$

the probability of ruin in $[0, T]$ by

$$\Psi(R_0, T) = \mathbf{P}\{T^r \in [0, T]\} = \mathbf{P}\{T^r \leqslant T\} \tag{2.3}$$

and the probability of ultimate ruin by

$$\Psi(R_0, \infty) = \mathbf{P}\{T^r \in [0, \infty)\} = \mathbf{P}\{T^r < \infty\} \tag{2.4}$$

A premium policy must be chosen so that $\Psi(R_0, T)$ and $\Psi(R_0, \infty)$ are acceptably small, premiums are competitive, and that any regulatory requirements are met. Ideally, one would like to derive explicit, tractable expressions for $\Psi(R_0, T)$ and $\Psi(R_0, \infty)$ in terms of the premium income processes $\{I^i\}_{i=1}^{N(\Pi)}$ and that an optimal premium policy could then be determined. Unfortunately, (2.1) is too general a framework to determine expressions for $\Psi(R_0, T)$ and $\Psi(R_0, \infty)$. In fact, even when further considerable and unrealistic simplifying assumptions are imposed on the structure of the income and claims processes, only rarely can tractable expressions for ruin probabilities be obtained.

2.2 Renewal Models: Deterministic Premiums

Typically, premium payments occur frequently and in small amounts relative to claims, which occur infrequently and in relatively large amounts. It is therefore natural to approximate the income processes by a deterministic function since the claims processes are the dominant sources of variability in R. For the rest of this chapter we consider $R_0 = u > 0$ as constant and the case $N(\Pi) = N(\chi) = 1$, suppressing indexes. Thus,

$$R_t \;=\; u + I_t - C_t \;=\; u + \sum_{k=1}^{N_t^{\Pi}} \Pi_k - \sum_{k=1}^{N_t^{\chi}} \chi_k \qquad t \in \mathbb{R}_+ \tag{2.5}$$

with payments $\{\Pi_k\}_{k\in\mathbf{N}}$ and claims $\{\chi_k\}_{k\in\mathbf{N}}$ occurring at times $\{T_k^\Pi\}_{k\in\mathbf{N}}$ and $\{T_k^\chi\}_{k\in\mathbf{N}}$. To see how one may construct an approximation of $I = \{I_t\}_{t\in\mathbb{R}_+}$, suppose that premium payments are made at times separated by a constant time interval of length Δt, for example, at the end of each month. One would then set $T_k^\Pi = k\Delta t$ for $k \in \mathbb{N}$ and so

$$N_t^\Pi = \sum_{k\in\mathbf{N}} 1_{\{T_k^\Pi \leqslant t\}} = \sum_{k\in\mathbf{N}} 1_{\{k\Delta t \leqslant t\}} = \left\lfloor \frac{t}{\Delta t} \right\rfloor$$

resulting in $I_t = \sum_{k=1}^{\lfloor \frac{t}{\Delta t} \rfloor} \Pi_k$. For $n \in \mathbb{N}$, $I_{n\Delta t} = \sum_{k=1}^{\lfloor \frac{n\Delta t}{\Delta t} \rfloor} \Pi_k = \sum_{k=1}^{n} \Pi_k$ so $\mathbf{E}[I_{n\Delta t}] = \sum_{k=1}^{n} \mathbf{E}[\Pi_k]$. A reasonable approximation to I then would be a smooth function $\pi(t)$ going through the points $\{(n\Delta t, \sum_{k=1}^{n} \mathbf{E}[\Pi_k])\}_{n\in\mathbf{N}_0}$ so that $\pi(n\Delta t) = \mathbf{E}[I_{n\Delta t}]$ for $n \in \mathbb{N}$. For instance, one could choose a piecewise linear function or a polynomial fitted to the points $\{(n\Delta t, \sum_{k=1}^{n} \mathbf{E}[\Pi_k])\}_{n\in\mathbf{N}_0}$. For such a choice for $\pi(t)$ we have for all $t \in \mathbb{R}_+$ that

$$\mathbf{E}[I_t] = \sum_{k=1}^{\lfloor \frac{t}{\Delta t} \rfloor} \mathbf{E}[\Pi_k] = \pi(\left\lfloor \frac{t}{\Delta t} \right\rfloor \Delta t)$$

If Δt is small compared to inter-claim times and to any time horizon of business activity, we can suppose $\lfloor \frac{t}{\Delta t} \rfloor \Delta t \approx t$. Since $\{\Pi_k\}_{k\in\mathbf{N}}$, and hence $\{\mathbf{E}[\Pi_k]\}_{k\in\mathbf{N}}$, are small compared to claim severities, we suppose $\pi(t)$ is sufficiently smooth so that $\lfloor \frac{t}{\Delta t} \rfloor \Delta t \approx t \Rightarrow \pi(\lfloor \frac{t}{\Delta t} \rfloor \Delta t) \approx \pi(t)$ uniformly in t. Our approximation of I is such that $\mathbf{E}[I_t] \approx \pi(t)$, properly describing the trend of I. If, in addition, $\{\mathbf{V}[\Pi_k]\}_{k\in\mathbf{N}}$ are small then the volatility of the process I is low and so I remains near $\mathbf{E}[I_t]$ with a high probability. We can then use the approximation $I_t \approx \pi(t)$ if the observed premium income process is fairly smooth with low volatility. If $\pi(t)$ is differentiable then $\pi'(t)$ represents the instantaneous rate of premium payment.

Now, suppose that the business is operating in an environment in which seasonal variations, variations in the number of policy holders, or any other factors that might induce fluctuations in claim intensities and/or claim severities are all negligible. Suppose also that policy holders expose themselves to risks in a similar but independent manner. Thus, $\{\chi_k\}_{k\in\mathbf{N}}$ can be considered an i.i.d. sequence. The inter-claim times

$\tau_k = T_k^\chi - T_{k-1}^\chi$, $k \in \mathbb{N}$, are also likely to be similar and unrelated and so too can be considered an i.i.d. sequence. Furthermore, the assumed time homogeneity suggests that the severity of claims has little to do with claim occurrence times and so $\{\chi_k\}_{k\in\mathbf{N}}$ can be assumed to be independent of $\{\tau_k\}_{k\in\mathbf{N}}$. With these assumptions, the general model (2.5) takes the form

$$R_t = u + \pi(t) - \sum_{k=1}^{N_t^\chi} \chi_k \quad for \quad t \in \mathbb{R}_+ \tag{2.6}$$

which is an ordinary renewal process. A modification could be made to allow the first inter-claim time τ_1 to be distributed differently than the other inter-claim times, resulting in a modified renewal process. This may better model the more realistic situation in which the ends of accounting periods do not necessarily coincide with $\{T_k^\chi\}_{k\in\mathbb{N}_0}$ (see Janssen (1981) for modified renewal models and Garrido (1987) for modified renewal models subject to interest and inflation). The assumption of i.i.d. claims and inter-claim times as well as their mutual independence are strong and are often unrealistic.

The techniques of renewal theory can be used to find expressions for the ruin probabilities $\Psi(u,T)$ and $\Psi(u,\infty)$, as defined in (2.3) and (2.4). Assume that in (2.6), N^χ is an ordinary renewal process so as to avoid the extra technicalities of modified renewal processes. Let F_τ and F_χ be the inter-claim time and claim severity distribution functions respectively. Define the corresponding probabilities of non-ruin $\Gamma(u,T) = 1 - \Psi(u,T)$ and $\Gamma(u,\infty) = 1 - \Psi(u,\infty)$. Furthermore, define these non-ruin probabilities conditioned on the first claim time τ_1 as well as on the first claim magnitude χ_1 by $\Gamma(u,T|\tau_1 = t)$, $\Gamma(u,T|\chi_1 = c)$, and $\Gamma(u,\infty|\tau_1 = t, \chi_1 = c)$. Using the law of total probability twice and the fact that τ_1 and χ_1 are distributed by F_τ and F_χ respectively,

$$\Gamma(u,T) = \int \Gamma(u,T|\tau_1 = t)dF_\tau(t) = \int\int \Gamma(u,T|\tau_1 = t, \chi_1 = c)dF_\chi(c)dF_\tau(t)$$

Noting that

$$\Gamma(u,T|\tau_1=t,\chi_1=c)=\begin{cases}0 & \text{if } u+\pi(t)-c<0\\ 1 & \text{if } t>T\\ \Gamma(u+\pi(t)-c,T-t) & \text{if } t\in[0,T]\ \&\ u+\pi(t)-c>0\end{cases}$$

we then obtain

$$\begin{aligned}\Gamma(u,T) &= \int_0^T\int_0^{u+\pi(t)}\Gamma(u+\pi(t)-c,T-t)dF_\chi(c)dF_\tau(t)+\int_T^\infty dF_\tau(t)\\ &= \int_0^T\int_0^{u+\pi(t)}\Gamma(u+\pi(t)-c,T-t)dF_\chi(c)dF_\tau(t)+1-F_\tau(T)\end{aligned}\tag{2.7}$$

Finally, letting $T\to\infty$ gives an equation for the probability of ultimate non-ruin:

$$\Gamma(u,\infty)=\int_0^\infty\int_0^{u+\pi(t)}\Gamma(u+\pi(t)-c,\infty)dF_\chi(c)dF_\tau(t)\tag{2.8}$$

Depending on $\pi(t)$, F_χ and F_τ, solving (2.7) and (2.8) for Γ may pose a significant problem. Using (2.7) and (2.8) in an optimization scheme where $\pi(t)$ is the unknown to be found may pose even greater problems. In short term policies, competitive market conditions may influence the choice of $\pi(t)$. In long term plans, such as life insurance, market conditions are less restrictive and so it is in these situations where an optimization of (2.7) and (2.8) may be useful when feasible.

Suppose $\mu=\mathbf{E}[\chi_k]$ is finite and let $M(t)=\mathbf{E}[N_t^\chi]$ be the renewal function of the process N^χ. Then,

$$\mathbf{E}[R_t]=u+\pi(t)-\mathbf{E}\left[\sum_{k=1}^{N_t^\chi}\chi_k\right]=u+\pi(t)-\mathbf{E}[N_t]\mathbf{E}[\chi_k]=u+\pi(t)-\mu M(t)$$

One can define a net accumulated premium policy $\pi_0(t)=\mu M(t)$ so that $\mathbf{E}[R_t]=u+\pi_0(t)-\mu M(t)=u$ ensuring that the premium income $\pi_0(t)$ exactly offsets claims on average. Since deviations from the expected behaviour will occur, the business must protect itself by adding a safety loading factor $\theta(t)>0$ to $\pi_0(t)$ to set a gross aggregate premium policy $\pi(t)=[1+\theta(t)]\pi_0(t)=[1+\theta(t)]\mu M(t)$. This gross premium will ensure that the average net income is strictly positive, allowing for the accumulation of reserve capital, but certain choices of $\theta(t)$ may still result in unacceptably high probabilities of ruin, again requiring an optimization involving (2.7) and (2.8). The form $\pi(t) = [1+\theta(t)]\mu M(t)$, or perhaps $\pi(t) = [1+\theta]\mu M(t)$ where $\theta > 0$ is a constant, may facilitate optimization but in general, exact expressions for $M(t)$ are infinite sums of convolutions and are therefore unlikely to lead to simplifications.

A minimal requirement ensuring that $\mathbf{P}\{T^r = \infty\} > 0$, ie, there is some chance the business will always be able to cover its claims, is that $\mathbf{P}\lim_{t\to\infty} \frac{R_t}{t} > 0$. This essentially says that for very large t, R_t is like $u + at$ in law for some $a > 0$ and that enough sample paths remain near $u + at > 0$, and hence above zero.

Since $\mu = \mathbf{E}[\chi_k] > 0$ and $\lambda^{-1} = \mathbf{E}[\tau_k] > 0$ are finite and $\lim_{t\to\infty} \frac{M(t)}{t} = \lambda$ we have that

$$\mathbf{P}\lim_{t\to\infty} \frac{R_t}{t} = \mathbf{P}\lim_{t\to\infty}\left[\frac{u}{t} + (1+\theta(t))\mu\frac{M(t)}{t} - \frac{1}{t}\sum_{k=1}^{N_t^\chi}\chi_k\right] = (1+\theta(\infty))\mu\lambda - \mu\lambda = \theta(\infty)\mu\lambda$$

And, requiring that $\mathbf{P}\lim_{t\to\infty} \frac{R_t}{t} > 0$ means requiring that $\theta(\infty) = \lim_{t\to\infty}\theta(t) > 0$, which is certainly satisfied if $\theta(t) = \theta > 0$, a constant.

Writing (2.6) in terms of the safety loading factor $\theta(t)$ and the renewal function $M(t)$ we have

$$R_t = u + (1+\theta(t))\mu M(t) - \sum_{k=1}^{N_t^\chi}\chi_k \quad for \quad t \in \mathbb{R}_+ \tag{2.9}$$

Thus, once the claims process has been specified, μ and $M(t)$ are, in principle, known, the remaining free parameters being $\theta(t)$ and u. In practice, however, u may be determined by circumstances or by law so that only $\theta(t)$ remains as a decision variable which itself may be further subject to market conditions.

2.3 The Classical Poisson Model

The classical Poisson risk reserve model is a very specialized version of (2.6) in which N^X is a Poisson process with rate $\lambda \in (0, \infty)$, ie, the inter-claim times are i.i.d. and exponentially distributed with mean λ^{-1} and common distribution function $F_\tau(t) = (1 - e^{-\lambda t})1_{[0,\infty)}(t)$. In this case, $M(t) = \mathbf{E}[N_t^X] = \lambda t$ represents the average number of claims in $[0, t]$. It is also assumed that $\pi(t) = \pi t$ where $\pi > 0$ is a constant aggregate rate of premium payment reflecting the assumption of a constant number of policy holders, constant portfolio composition, and premium payments are regular. Using a constant safety loading factor $\theta > 0$, we have that $\pi(t) = \pi t = (1 + \theta)\mu\lambda t$ so that $\pi = (1 + \theta)\mu\lambda$ is the aggregate gross premium rate with $\pi_0 = \mu\lambda$ being the net aggregate premium rate. (2.9) then becomes

$$R_t = u + (1 + \theta)\mu\lambda t - \sum_{k=1}^{N_t^X} \chi_k \quad for \quad t \in \mathbb{R}_+$$

Specializing (2.7) yields

$$\begin{aligned} \Gamma(u,T) &= \int_0^T \int_0^{u+\pi t} \Gamma(u + \pi t - c, T - t) dF_\chi(c) dF_\tau(t) + 1 - F_\tau(T) \\ &= \int_0^T \int_0^{u+\pi t} \Gamma(u + \pi t - c, T - t) dF_\chi(c) \lambda e^{-\lambda t} dt + e^{-\lambda T} \end{aligned} \tag{2.10}$$

Letting $T \to \infty$ we obtain

$$\begin{aligned} \Gamma(u,\infty) &= \int_0^\infty \int_0^{u+\pi t} \Gamma(u + \pi t - c, \infty) dF_\chi(c) \lambda e^{-\lambda t} dt \\ &= \frac{\lambda}{\pi} e^{\frac{\lambda u}{\pi}} \int_u^\infty e^{-\frac{\lambda x}{\pi}} \int_0^x \Gamma(x - c, \infty) dF_\chi(c) dx \end{aligned} \tag{2.11}$$

where we have used the substitution $x = u + \pi t$.

Assuming (2.11) and (2.10) are differentiable in u and that $dF_\chi(c) = f_\chi(c)dc$ we obtain by differentiating

$$\begin{aligned}\frac{\partial \Gamma}{\partial u}(u,T) &= \int_0^T \Gamma(0,T-t) f_X(u+\pi t)\lambda e^{-\lambda t} dt \\ &+ \int_0^T \int_0^{u+\pi t} \frac{\partial \Gamma}{\partial u}(u+\pi t - c, T-t) f_X(c) dc \lambda e^{-\lambda t} dt \qquad (2.12)\end{aligned}$$

and

$$\frac{\partial \Gamma}{\partial u}(u,\infty) = \frac{\lambda}{\pi}\Gamma(u,\infty) - \frac{\lambda}{\pi}\int_0^u \Gamma(u-c,\infty) dF_X(c) \qquad (2.13)$$

(2.12) is a Volterra integro-differential equation having a limited number of analytical solutions and for which numerical solutions exist only for certain choices of F_X. In general, (2.12) is not a practical method of computing the probability of ruin in finite time (see for example Linz (1985)). (2.12) and (2.13) can be put into the form of renewal equations and solved, in principle, by Laplace transforms where the difficulty is in the inversion (Feller 1966). On the other hand, as long as F_X has a density, (2.13) can clearly be handled numerically, for example, by an Euler scheme plus a discretization of the integral or by the technique of product integration (Ramsay & Usabel 1997). Also, we may, depending on F_X, differentiate again with respect to u to obtain a second order ODE with delay, possibly being able to solve. Finally, in a direct Monte-Carlo simulation of small ruin probabilities there is a large relative error unless a very large sample of paths are simulated (Asmussen 1984) which can be computationally intensive.

2.4 Renewal Models: Stochastic Premiums

We suggest here a simple method of allowing for stochastic premiums while retaining mathematical tractability. Describe claims and premium income together in a single point process as follows. Let the occurrence times of premium income and claims be given by a $\mathbf{P}$-a.s. strictly increasing sequence $\{T_k\}_{k\in\mathbb{N}}$ of $\mathbf{F}$-stopping times as before where $T_0 = 0$. Suppose, as before, that $T_\infty = \sup_k T_k = \infty$ $\mathbf{P}$-a.s. so that the associated counting process N is without explosion. Suppose the premium and claim magnitudes are given by a single sequence $\{X_k\}_{k\in\mathbf{N}}$ of random variables whose range is $\mathbb{R}$. A positive value of the X_k would represent premium income at time T_k and a negative value would represent a claim. The risk reserve process could then be described by

$$R_t = u + \sum_{k=1}^{N_t} X_k \quad for \quad t \in \mathbb{R}_+$$

One would want the distributions F_k of the X_k to be such that there is no mass at zero as well as having the probability of a claim at T_k, given by $P_k^X = \mathbf{P}\{X_k < 0\} = \int_{-\infty}^{0} dF_k(x) > 0$, and the probability of a premium at T_k, given by $P_k^\Pi = \mathbf{P}\{X_k > 0\} = \int_0^{-\infty} dF_k(x) > 0$, to match observed behaviour or to reflect assumptions about anticipated variations in income and claims. In addition, one would want the "shape" of $F_k(x)$ for $x < 0$ to match the observed distribution of claims and the "shape" of $F_k(x)$ for $x > 0$ to match the observed distribution of premiums. For instance, if $f_k^X(x), x \geqslant 0$, and $f_k^\Pi(x), x \geqslant 0$, are the observed density functions of the claims and premium magnitudes, respectively, then, since $P^\Pi + P^X = 1$, one could set

$$f_k(x) = P_k^X f_k^X(-x) 1_{(-\infty,0)}(x) + P_k^\Pi f_k^\Pi(x) 1_{(0,\infty)}(x)$$

and

$$F_k(x) = \int_{-\infty}^{x} f_k(t)dt$$

If one supposes that the $\{X_k\}_{k\in\mathbf{N}}$ are i.i.d. and that N is an ordinary renewal process then renewal theory techniques can again be applied to determine ruin probabilities. It should be noted that such an assumption amounts to allowing the income to occur at purely random times, possibly an undesirable feature. Again,

$$\Gamma(u,T) = \int\int \Gamma(u,T|X_1 = c, \tau_1 = t) dF_X(c) dF_\tau(t)$$

and

$$\Gamma(u,T|\tau_1 = t, X_1 = c) = \begin{cases} 0 & \text{if } u + c < 0 \text{ and } t \in [0,T] \\ \Gamma(u+c, T-t) & \text{if } u + c \geqslant 0 \text{ and } t \in [0,T] \\ 1 & \text{if } t > T \end{cases}$$

so we then obtain

$$\begin{aligned}\Gamma(u,T) &= \int_0^T \int_{-u}^{\infty} \Gamma(u+c, T-t) dF_X(c) dF_\tau(t) + \int_T^{\infty} dF_X(c) dF_\tau(t) \\ &= \int_0^T \int_{-u}^{\infty} \Gamma(u+c, T-t) dF_X(c) dF_\tau(t) + 1 - F_\tau(T)\end{aligned}$$

Finally, letting $T \to \infty$ gives an equation for the probability of ultimate non-ruin:

$$\begin{aligned}\Gamma(u,\infty) &= \int_0^{\infty} \int_{-u}^{\infty} \Gamma(u+c, \infty) dF_X(c) dF_\tau(t) \\ &= \int_{-u}^{\infty} \Gamma(u+c, \infty) dF_X(c)\end{aligned}$$

yielding expressions for ruin probabilities of a similar level of tractability as (2.10) and (2.11).

2.5 Generalizations

The special cases of renewal models discussed above describe the time homogeneous evolution of a risk reserve fund. There are situations in which both claim severities and claim occurrence times fluctuate either deterministically or randomly. For example, the south eastern US seaboard experiences hurricanes more frequently and of greater scale in the summer months than in the winter months resulting in a greater number of more sizeable claims during the summer. Icy road conditions in the winter may lead to similar seasonal fluctuations in the size and frequency of claims associated with automobile accidents. Thus, there is a need for more realistic models incorporating such fluctuations. As already mentioned, Cox processes are being investigated by Grandell (1991), periodic variations in arrival intensity by Chukova, Dimitrov & Garrido (1993), and piecewise continuous Markov processes by M.H.A. Davis and Paul Embrechts. These approaches, while providing more realistic models, will unlikely yield more tractable expressions than the simpler models. The inadequate performance of the simpler models (Seal 1983) and the expected increase in complexity of the more realistic point process models adds impetus to the search for useful approximations.

Chapter 3

Approximation of Point Process Reserve Models

3.1 The Classical Normal Approximation

As seen in Chapter 2, the determination of ruin probabilities for point process models poses difficulties, even for the simpler classical Poisson model. One way of treating this problem is to find a good approximation to the point process model which allows the explicit determination of ruin probabilities. Initial approaches involved normal approximations of the claim magnitudes experienced over a given time interval. A natural extension of this is to perform such a normal approximation at each instant of time. Let $R_t = u + \pi t - \sum_{k=1}^{N_t^X} \chi_k$ for $t \in \mathbb{R}_+$ be a classical Poisson reserve model with rate $\lambda \in (0, \infty)$, $\{\chi_k\}_{k \in \mathbb{N}}$ are i.i.d. with $\mu = \mathbf{E}[\chi_k] > 0$ and $\sigma^2 = \mathbf{V}[\chi_k] > 0$, $u > 0$ is the initial capital and $\pi > 0$ is the aggregate premium rate. Let $\tilde{R}$ denote the approximating process based on performing a normal approximation of R at each time t. If the laws of R and $\tilde{R}$ are to be "close" enough that their relevant macroscopic properties are the same then a necessary requirement is that the first two moments match at all times: $\mathbf{E}[R_t] = \mathbf{E}[\tilde{R}_t]$ and $\mathbf{V}[R_t] = \mathbf{V}[\tilde{R}_t]$ $\forall t \in \mathbb{R}_+$. $\mathbf{E}[R_t] = u + (\pi - \mu\lambda)t$

and $\mathbf{V}[R_t]$ can be computed via the formula $\mathbf{V}[R_t] = \mathbf{E}[\mathbf{V}[R_t|N_t]] + \mathbf{V}[\mathbf{E}[R_t|N_t]]$. For each $n \in \mathbb{N}$,

$$\mathbf{E}[R_t|N_t^X = n] = \mathbf{E}\left[u + \pi t - \sum_{k=1}^{n} \chi_k\right] = u + \pi t - n\mu$$

and, knowing that $\mathbf{E}[N_t^X] = \mathbf{V}[N_t^X] = \lambda t$, we have

$$\mathbf{V}[\mathbf{E}[R_t|N_t^X]] = \mathbf{V}[u + \pi t - N_t^X \mu] = \mu^2 \lambda t$$

Also,

$$\mathbf{V}[R_t|N_t^X = n] = \mathbf{V}\left[u + \pi t - \sum_{k=1}^{n} \chi_k\right] = n\sigma^2$$

and so

$$\mathbf{E}[\mathbf{V}[R_t|N_t^X]] = \mathbf{E}[\sigma^2 N_t^X] = \sigma^2 \lambda t$$

Hence, $\mathbf{V}[R_t] = \lambda t(\sigma^2 + \mu^2)$. The obvious normal approximation would then be $u + (\pi - \mu\lambda)t \pm \sqrt{\lambda t(\sigma^2 + \mu^2)}Z$ where $Z \sim N(0,1)$. But, $\sqrt{t}Z \sim N(0,t)$, which resembles a standard Wiener process W. We are therefore led to the approximation $\tilde{R}_t = u + (\pi - \mu\lambda)t - \sqrt{\lambda(\sigma^2 + \mu^2)}W_t$, partially justified by the fact that R and $\tilde{R}$ have the same trend and volatility at each time and that W and C are both stationary independent increment processes, sharing a similar structure. The utility of the Wiener diffusion approximation is that $W_t \sim N(0,t)$. Thus, "$R \approx \tilde{R}$" suggests that $\mathbf{P}\{R_t \leqslant x\} \approx \mathbf{P}\{\tilde{R}_t \leqslant x\}$ and so we have

$$\begin{aligned}
\mathbf{P}\{\tilde{R}_t \leqslant x\} &= \mathbf{P}\{u + (\pi - \mu\lambda)t + \sqrt{\lambda(\sigma^2 + \mu^2)}W_t \leqslant x\} \\
&= \mathbf{P}\left\{W_t \leqslant \frac{x - u - (\pi - \mu\lambda)t}{\sqrt{\lambda(\sigma^2 + \mu^2)}}\right\} = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\frac{x-u-(\pi-\mu\lambda)t}{\sqrt{\lambda(\sigma^2+\mu^2)}}} e^{-\frac{s^2}{2t}}\, ds \\
&= \frac{1}{A\sqrt{2\pi t}} \int_{-\infty}^{x} e^{-\frac{[s-u-Bt]^2}{2A^2 t}}\, ds \qquad (3.1)
\end{aligned}$$

where the transformation $s \mapsto \frac{s-u-(\pi-\mu\lambda)t}{\sqrt{\lambda(\sigma^2+\mu^2)}}$ was used as well as the substitutions $A = \sqrt{\lambda(\sigma^2 + \mu^2)}$ and $B = \pi - \mu\lambda$. The distribution of $T^r(\tilde{R}) = \inf\{t > 0 : \tilde{R}_t < 0\}$

is also known explicitly:

$$\begin{aligned}\Psi(u,T) &= \mathbf{P}\{T^r(\tilde{R}) \leqslant T\} = \frac{A}{u\sqrt{2\pi}} \int_0^T s^{-\frac{3}{2}} e^{-\frac{1}{2}[\frac{u^2}{As}+\frac{2Bu}{A}+(\frac{Bu}{A})^2 s]} ds \\ &= \Phi\left(\frac{-BT-u}{A\sqrt{T}}\right) + e^{\frac{-2Bu}{A^2}} \Phi\left(\frac{BT-u}{A\sqrt{T}}\right) \end{aligned} \tag{3.2}$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-s^2/2} ds$ is the distribution function of the standard normal distribution. The integral representation can be found in Karlin & Taylor (1981) and the evaluation of the integral via Laplace transforms can be found in Darling & Siegert (1953). Letting $T \to \infty$ then yields the probability of ultimate ruin

$$\Psi(u,\infty) = \mathbf{P}\{T^r(\tilde{R}) < \infty\} = e^{\frac{-2Bu}{A^2}} = e^{\frac{-2(\pi-\mu\lambda)u}{\lambda(\sigma^2+\mu^2)}} \tag{3.3}$$

(3.2) and (3.3) lend themselves easily to an optimization scheme for determining an optimal premium policy π.

The above derivation relies on the ease of computing the mean and variance of the Poisson process N^{χ} and the Wiener process W. For general counting processes N^{χ}, computing the mean and the variance may be non-trivial making the matching of moments difficult. Also, unless N^{χ} is Poisson, R is not an independent increment process and so is structurally different from $\tilde{R}$. Thus, there is a serious drawback, even in the Poisson case: there is no satisfactory justification of the assumption that $\tilde{R}$ provides a good approximation of R since matching the first two moments does not guarantee that the long run behaviour or functionals of these processes match. The more modern approach to this latter problem, based on the theoretical work of Prohorov (1956), Skorokhod (1956), and Billingsley (1968) and then applied by Iglehart (1969), is to construct a sequence of point processes $\{R^{(n)}\}_{n\in\mathbf{N}}$ such that $R^{(n)}$ converges to a limit process $R^{(\infty)}$ in a sense strong enough to ensure that for useful functionals f, $f(R^{(n)})$ converge to $f(R^{(\infty)})$. The hope is that our original point process model R is "close" in law to $R^{(\infty)}$ so that $f(R)$ will be close to $f(R^{(\infty)})$ and, most importantly, that $f(R^{(\infty)})$ has an explicit, tractable expression which could then serve as a useful approximation to $f(R)$. The natural setting for considering the limits of sequences of càdlàg stochastic processes and their functionals is the space of càdlàg functions endowed with a probability structure which we now describe.

3.2 Weak Convergence in $D[0,T]$ and $D[0,\infty)$

For any $T \in \mathbb{R}_+$, let $D[0,T] = \{x : [0,T] \to \mathbb{R} : x \textit{ is } \text{càdlàg}\}$. Let Λ_T be the set of maps $\lambda : [0,T] \to [0,T]$ that are onto, strictly increasing, and continuous with $\|\lambda\|_{\Lambda_T} < \infty$ where $\|\lambda\|_{\Lambda_T} = \sup_{0 \leqslant s < t \leqslant T} \left|\ln\left(\frac{\lambda(t)-\lambda(s)}{t-s}\right)\right|$. For $x, y \in D[0,T]$ define $d_T(x,y) = \inf_{\lambda \in \Lambda_T}\left\{\|\lambda\|_{\Lambda_T} \vee \sup_{s\in[0,T]}\{|x(s) - y \circ \lambda(s)| \wedge 1\}\right\}$. For $T = 1$, d_1 is a metric on $D[0,1]$ and $(D[0,1], d_1)$ is a complete, separable metric space (Billingsley 1968). It is clear that the same is true for $(D[0,T], d_T)$ for any $T \in \mathbb{R}_+$. Let $D = D[0,\infty) = \{x : \mathbb{R}_+ \to \mathbb{R} : x \textit{ is } \text{càdlàg}\}$. Let Λ be the set of maps $\lambda : \mathbb{R}_+ \to \mathbb{R}_+$ that are onto, strictly increasing, and Lipschitz continuous with $\|\lambda\|_\Lambda < \infty$ where $\|\lambda\|_\Lambda = \sup_{0 \leqslant s < t < \infty} \left|\ln\left(\frac{\lambda(t)-\lambda(s)}{t-s}\right)\right|$. For $x, y \in D$, $\lambda \in \Lambda$, and $T \in \mathbb{R}_+$, set $\rho(x,y,\lambda,T) = \sup_{s\in[0,T]}\{|x(s) - y \circ \lambda(s)| \wedge 1\}$ and defining the metric $d(x,y) = \inf_{\lambda\in\Lambda}\left\{\|\lambda\|_\Lambda \vee \int_{\mathbb{R}_+} \rho(x,y,\lambda,T)e^{-T}dT\right\}$ on D we have that (D,d) is a complete, separable metric space (Ethier & Kurtz 1986). The following theorem characterizes convergence in (D,d).

Theorem 1 (Ethier & Kurtz, 1986, pg. 125) *Let $\{x_n\}_{n\in\mathbb{N}} \subset D$ and $x \in D$. $d(x_n, x) \to 0$ if and only if the following three conditions hold for each $t \in \mathbb{R}_+$ and all sequences $\{t_n\}_{n\in\mathbb{N}} \subset \mathbb{R}_+$ such that $t_n \to t$:*
1) $|x_n(t_n) - x(t)| \wedge |x_n(t_n) - x(t-)| \to 0$
2) If $|x_n(t_n) - x(t)| \to 0$, $s_n \geqslant t_n$, $s_n \to t$ then $|x_n(s_n) - x(t)| \to 0$
3) If $|x_n(t_n) - x(t-)| \to 0$, $s_n \geqslant t_n$, $s_n \to t$ then $|x_n(s_n) - x(t-)| \to 0$

However, as Pollard (1984) mentions, if the limit x is continuous or lies in some separable subset of D then convergence with respect to the uniform metric given by $m_T(x_n, x) = \sup_{s\in[0,T]} |x_n(s) - x(s)|$ on all compacts $[0,T]$ is equivalent to convergence with respect to d. Thus, convergence in the space $(C[0,\infty), m)$, where

$m(x,y) = \sup_{t\in\mathbb{R}_+} |x(t) - y(t)|$, implies convergence in $(C[0,T], m_T)$ for all T and therefore in (D,d). Here we want a framework that will allow for discontinuous limits and so we always work within (D,d).

The metric topology τ on D (resp. $D[0,T]$) determined by d (resp. d_T) is Skorokhod's J_1 topology. Let $\mathcal{B}(\tau)$ be the Borel σ-algebra on D (resp. $D[0,T]$) generated by τ. Our risk reserve model R is an $\mathbb{R}$-valued stochastic process whose sample paths are in D. Considering our reserve model as a random element of D, assume a probability model $(\Omega, \mathcal{F}, \mathbf{P})$ and let $R : \Omega \to D$ be a $\mathcal{F}/\mathcal{B}(\tau)$-measurable map. The distribution of R is given by the induced probability measure $P = \mathbf{P}R^{-1}$ on $\mathcal{B}(\tau)$, called the law of the process R. For processes $\{R^{(n)}\}_{n\in\mathbf{N}}$ and $R^{(\infty)}$ on D let $\{P^{(n)} = \mathbf{P}{R^{(n)}}^{-1}\}_{n\in\mathbf{N}}$ and $P^{(\infty)} = \mathbf{P}{R^{(\infty)}}^{-1}$ be the corresponding induced probability measures. The sequence of laws $\{P^{(n)}\}_{n\in\mathbf{N}}$ is defined to converge weakly to $P^{(\infty)}$, written $P^{(n)} \Rightarrow P^{(\infty)}$, if for all bounded, continuous functions $f : D \to \mathbb{R}$ we have $\lim_{n\to\infty} \int_D f(x)dP^{(n)}(x) = \int_D f(x)dP^{(\infty)}(x)$, or equivalently, $\lim_{n\to\infty} \mathbf{E}[f(R^{(n)})] = \mathbf{E}[f(R^{(\infty)})]$. In this case we say $R^{(n)}$ converges in distribution to $R^{(\infty)}$ and denote this by $R^{(n)} \Rightarrow R^{(\infty)}$ or by $R_t^{(n)} \Rightarrow R_t^{(\infty)}$ where in this context $R_t^{(n)}$ is taken to mean the process $R^{(n)}$, not just its value at time t.

For $k \in \mathbb{N}$ and $\{t_1, \dots, t_k\} \subset \mathbb{R}_+$ let the projections $\pi_{t_1,\cdots,t_k} : D \to \mathbb{R}^k$ be defined by $\pi_{t_1,\cdots,t_k}(R^{(n)}) = (R_{t_1}^{(n)}, \cdots, R_{t_k}^{(n)})$. If we consider only the finite dimensional distributions $\mathbf{P}{R^{(n)}}^{-1}{\pi_{t_1,\cdots,t_k}}^{-1}$ of a process $R^{(n)}$ induced on $\mathcal{B}(\mathbb{R}^k)$ by $\pi_{t_1,\cdots,t_k}(R^{(n)})$ then weak convergence of $\mathbf{P}{R^{(n)}}^{-1}{\pi_{t_1,\cdots,t_k}}^{-1}$ to $\mathbf{P}{R^{(\infty)}}^{-1}{\pi_{t_1,\cdots,t_k}}^{-1}$ is just the ordinary weak convergence of distribution functions (Billingsley 1968). $\mathbf{P}{R^{(n)}}^{-1} \Rightarrow \mathbf{P}{R^{(\infty)}}^{-1}$ implies the weak convergence of all finite dimensional distributions. However, weak convergence of all finite dimensional distributions $\mathbf{P}{R^{(n)}}^{-1}{\pi_{t_1,\cdots,t_k}}^{-1}$ to $\mathbf{P}{R^{(\infty)}}^{-1}{\pi_{t_1,\cdots,t_k}}^{-1}$ does not determine the weak convergence of $\mathbf{P}{R^{(n)}}^{-1}$ to $\mathbf{P}{R^{(\infty)}}^{-1}$. Thus, weak convergence for laws of processes is a stronger notion that the weak convergence of finite dimensional distributions, even for processes with $\mathbf{P}$-a.s. continuous sample paths, and turns out to be the mode of convergence that is strong enough for the convergence of functionals. We now collect some theorems on weak convergence for later use.

Theorem 2 (Billingsley, 1995, pg. 331) *Let $\{X^{(n)}\}_{n\in\mathbb{N}}$ and I be processes in $D[0,1]$ where I is the non-random identity process in $D[0,1]$ defined by $I(t,\omega) = t$ for all $(t,\omega) \in [0,1]\times\Omega$ and where $\lambda > 0$ is a constant. Then, $X^{(n)} \Rightarrow \lambda I \iff X^{(n)} \xrightarrow{\mathbf{P}} \lambda I$.*

Theorem 3 (Billingsley, 1968, pg. 27) *Let $\{X^{(n)}\}_{n\in\mathbb{N}}$, X, and $\{Y^{(n)}\}_{n\in\mathbb{N}}$ be processes in $D[0,1]$. If $X^{(n)} \Rightarrow X$ and $Y^{(n)} \xrightarrow{\mathbf{P}} \lambda I$ for some $\lambda > 0$ then $(X^{(n)}, Y^{(n)}) \Rightarrow (X, \lambda I)$. In other words, the induced probability measures $\{\mathbf{P}(X^{(n)}, Y^{(n)})^{-1}\}_{n\in\mathbb{N}}$ converge weakly to $\mathbf{P}(X,\lambda I)^{-1}$ on $D[0,1]\times D[0,1]$ endowed with the product topology $\tau\otimes\tau$ generated by the sets $\{A\times B : A, B \in \tau\}$ and the Borel σ-algebra $\mathcal{B}(\tau\otimes\tau)$ generated by $\tau\otimes\tau$.*

Theorem 4 (Davidson, 1994, pg. 355) *Let $\{X^{(n)}\}_{n\in\mathbb{N}}$ and $\{Y^{(n)}\}_{n\in\mathbb{N}}$ be processes in D such that $X^{(n)} \Rightarrow X$ and $Y^{(n)} \Rightarrow a$ where a is a non-random element of D (for all $(t,\omega) \in \mathbb{R}_+\times\Omega$, $a(t,\omega) = a(t)$). Then $Y^{(n)} + X^{(n)} \Rightarrow a + X$ and $Y^{(n)}X^{(n)} \Rightarrow aX$.*

Theorem 5 (Billingsley 1968, pg. 50) *Let $h : D[0,1] \to \mathbb{R}$ be $\mathcal{B}(\tau)/\mathcal{B}(\mathbb{R})$- measurable and let $D_h = \{x \in D[0,1] : h$ is discontinuous at $x\}$. If $\{P^{(n)}\}_{n=1}^{\infty}$ and P are probability measures on $\mathcal{B}(\tau)$ such that $P^{(n)} \Rightarrow P$ and $P(D_h) = 0$ then $P^{(n)}h^{-1} \Rightarrow Ph^{-1}$ where $\{P^{(n)}h^{-1}\}_{n\in\mathbb{N}}$ and Ph^{-1} are the corresponding induced probability measures on $\mathcal{B}(\mathbb{R})$.*

Corollary 1 (Billingsley 1968, pg. 31) *If $X^{(n)} \Rightarrow X$ and $\mathbf{P}X^{-1}(D_h) = 0$ where h is as in Theorem 5 then $h(X^{(n)}) \Rightarrow h(X)$.*

Note that Theorem 5 and Corollary (1) remain true if $(D[0,1], d_1)$ is replaced with any complete, separable metric space, such as (D, d).

Theorem 6 (Billingsley 1968, pg.25, 225) *Let $\{X^{(n)}\}_{n\in\mathbb{N}}$ and $\{Y^{(n)}\}_{n\in\mathbb{N}}$ be sequences of random elements of $D[0,1]$. Since $(D[0,1], d_1)$ is a separable metric space, the map $d(X^{(n)}, Y^{(n)}) : \Omega \to \mathbb{R}$ defined by $d(X^{(n)}, Y^{(n)})(\omega) = d(X^{(n)}(\omega), Y^{(n)}(\omega))$ is a random variable. If $X^{(n)} \Rightarrow X$ and $d(X^{(n)}, Y^{(n)}) \xrightarrow{\mathbf{P}} 0$ then $Y^{(n)} \Rightarrow X$.*

Theorem 7 (Prohorov, 1956) *Let $\{\chi_k^{(n)}\}_{k=1,n\in\mathbb{N}}^{n+1}$ be random variables such that for each $n \in \mathbb{N}$, $\{\chi_k^{(n)}\}_{k=1}^{n+1}$ are i.i.d. with $\mathbf{E}[\chi_k^{(n)}] = 0$, $\mathbf{V}[\chi_k^{(n)}] = {\sigma^{(n)}}^2 \to \zeta^2 > 0$, and $\exists\delta > 0$ such that $\sup_{n\in\mathbb{N}} \mathbf{E}[|\chi_k^{(n)}|^{2+\delta}] < \infty$. Define on $(C[0,1], m_1)$ the process*

$$Y_t^{(n)} = \frac{1}{\zeta\sqrt{n}}\left[\sum_{k=1}^{\lfloor nt\rfloor}\chi_k^{(n)} + (nt - \lfloor nt\rfloor)\chi_{\lfloor nt\rfloor+1}^{(n)}\right] \quad for \quad t \in [0,1]$$

Then, $Y^{(n)} \Rightarrow W$ where W is a standard Wiener process.

The next theorem is a generalization of Billingsley's result for random time changes. As with other results taken from Furrer et al. (1996), not yet published, a detailed proof is provided here.

Theorem 8 (Furrer et al. 1996) *Let $\{X^{(n)}\}_{n\in\mathbb{N}}$, X, and $\{N^{(n)}\}_{n\in\mathbb{N}}$ be processes in D such that $X^{(n)} \Rightarrow X$ and $N^{(n)} \Rightarrow \lambda I$ for some $\lambda > 0$. Suppose also that the sample paths of $\{N^{(n)}\}_{n\in\mathbb{N}}$ are non-decreasing and $N^{(n)}(0) = 0$. Then, $X^{(n)} \circ N^{(n)} \Rightarrow X \circ \lambda I$ where $(X^{(n)} \circ N^{(n)})_t = X_{N_t^{(n)}}^{(n)}$ and $(X \circ \lambda I)_t = X_{\lambda t}$ for $t \in \mathbb{R}_+$.*

Proof. Let $D_0 = \{\phi \in D : \phi(0) = 0$ and ϕ non-decreasing $\}$. Define a composition map $\psi : D \times D_0 \to D$ by $\psi(x,\phi)(t) = x \circ \phi(t) = x(\phi(t))$. We will show that when $D \times D_0$ is suitably topologized, ψ is measurable, $(X^{(n)}, N^{(n)}) \Rightarrow (X, \lambda I)$, and $\mathbf{P}(X, \lambda I)^{-1}(D_\psi) = 0$ where D_ψ is the set of discontinuities of ψ and $\mathbf{P}(X,\lambda I)^{-1}$ is the probability measure induced by $(X, \lambda I)$. Corollary (1) is then applied to conclude that $X^{(n)} \circ N^{(n)} = \psi(X^{(n)}, N^{(n)}) \Rightarrow \psi(X, \lambda I) = X \circ \lambda I$.

Let τ_0 be the metric topology on D_0 determined by $d|_{D_0}$. τ_0 is therefore the relative topology induced on D_0 by τ and can be described by $\{U \cap D_0 : U \in \tau\}$. Let $\mathcal{B}(\tau_0)$ be the Borel σ-algebra generated by τ_0 which can be described by $\{B \cap D_0 : B \in \mathcal{B}(\tau)\}$. However, D_0 is closed with respect to d and so $D_0 \in \mathcal{B}(\tau)$. $\mathcal{B}(\tau_0)$ can therefore be described by $\{B \subset D_0 : B \in \mathcal{B}(\tau)\}$.

The metric $\tilde{d} = d \vee d|_{D_0}$ determines the product topology $\tau \otimes \tau_0$ on $D \times D_0$ which can be generated by the sets $\{U \times U_0 : U \in \tau, U_0 \in \tau_0\}$. Let $\mathcal{B}(\tau \otimes \tau_0)$

be the Borel σ-algebra on $D \times D_0$ generated by $\tau \otimes \tau_0$. Since (D, d) is separable, $\mathcal{B}(\tau \otimes \tau_0) = \mathcal{B}(\tau) \otimes \mathcal{B}(\tau_0)$ where $\mathcal{B}(\tau) \otimes \mathcal{B}(\tau_0)$ is the product σ-algebra generated by $\{B \times B_0 : B \in \mathcal{B}(\tau), B_0 \in \mathcal{B}(\tau_0)\}$. Now, for $\{t_1, \cdots, t_n\} \subset \mathbb{R}_+$ we have a projection $\pi_{t_1,\cdots,t_n} : D \to \mathbb{R}^n$ defined by $\pi_{t_1,\cdots,t_n}(x) = (x(t_1), \cdots, x(t_n))$. $\mathcal{B}(\tau)$ is in fact generated by the sets $\{\pi^{-1}_{t_1,\cdots,t_n}(B_1 \times \cdots \times B_n) : n \in \mathbb{N}, B_i \in \mathcal{B}(\mathbb{R}), \{t_1, \cdots, t_n\} \subset \mathbb{R}_+\}$ and similarly for $\mathcal{B}(\tau_0)$ (Lindvall, 1973). $\mathcal{B}(\tau) \otimes \mathcal{B}(\tau_0)$ is also generated by the pre-images of Borel rectangles under similarly defined projections of $D \times D_0$ and hence so is $\mathcal{B}(\tau \otimes \tau_0)$. Thus, $\psi : (D \times D_0, \mathcal{B}(\tau \otimes \tau_0)) \to (D, \mathcal{B}(\tau))$ is measurable since for all $t \in \mathbb{R}_+$ the projection $\pi_t \circ \psi : D \times D_0 \to \mathbb{R}$ defined by $\pi_t \circ \psi(x, \phi) = \pi_t(x \circ \phi) = x(\phi(t))$ is $\mathcal{B}(\tau \otimes \tau_0)/\mathcal{B}(\mathbb{R})$-measurable (pg. 232 Billingsley 1968). Also, $\{(X^{(n)}, N^{(n)})\}_{n \in \mathbb{N}}$ and $(X, \lambda I)$ are $\mathcal{F}/\mathcal{B}(\tau \otimes \tau_0)$-measurable and so are random elements of $D \times D_0$. By composing measurable maps, it follows that $\{\psi(X^{(n)}, N^{(n)})\}_{n \in \mathbb{N}}$ and $\psi(X, \lambda I)$ are $\mathcal{F}/\mathcal{B}(\tau)$-measurable and so are random elements of D.

By Theorem 2, $N^{(n)} \Rightarrow \lambda I$ is equivalent to $N^{(n)} \xrightarrow{\mathbf{P}} \lambda I$. By Theorem (3), $N^{(n)} \xrightarrow{\mathbf{P}} \lambda I$ and $X^{(n)} \Rightarrow X$ imply that $(X^{(n)}, N^{(n)}) \Rightarrow (X, \lambda I)$ in the space $(D \times D_0, \mathcal{B}(\tau \otimes \tau_0))$. Now, let $\tilde{P} = \mathbf{P}(X, \lambda I)^{-1}$ be the probability measure on $\mathcal{B}(\tau \otimes \tau_0)$ induced by $(X, \lambda I)$. Let $D_\psi = \{(x, \phi) \in D \times D_0 : \psi \text{ is discontinuous at } (x, \phi)\}$. If we show that $\tilde{P}(D_\psi) = 0$ then since ψ is measurable and $(X^{(n)}, N^{(n)}) \Rightarrow (X, \lambda I)$, Corollary (1) implies that $\psi(X^{(n)}, N^{(n)}) \Rightarrow \psi(X, \lambda I)$, proving the result.

Noting that $D_\psi = [D_\psi \cap (D \times \{\lambda I\})] \cup [D_\psi \cap (D \times \{\lambda I\})^c]$ and $\tilde{P} = \mathbf{P}(X, \lambda I)^{-1}$ we have that $\tilde{P}(D_\psi \cap (D \times \{\lambda I\})^c) = 0$ and so $\tilde{P}(D_\psi) = \tilde{P}(D_\psi \cap (D \times \{\lambda I\})$. Thus, we need to consider the continuity of ψ only at points in $D \times \{\lambda I\}$. Let $(x, \lambda I) \in D \times \{\lambda I\}$ be arbitrary and let $\{(x_n, \phi_n)\}_{n \in \mathbb{N}} \subset D \times D_0$ be any sequence such that $\tilde{d}((x_n, \phi_n), (x, \lambda I)) \to 0$. This is equivalent to $d(x_n, x) \to 0$ and $d(\phi_n, \lambda I) \to 0$. We want to show that $x_n \circ \phi_n = \psi(x_n, \phi_n) \to \psi(x, \lambda I) = x \circ \lambda I$ and so we use Theorem 1 characterizing convergence in D.

Let $t \in \mathbb{R}_+$ and $\{t_n\}_{n \in \mathbb{N}} \subset \mathbb{R}_+$ be such that $t_n \to t$. We want to verify that the conditions of Theorem (1) hold for the sequence of processes $\{x_n \circ \phi_n\}_{n \in \mathbb{N}}$ and $x \circ \lambda I$. Since $t_n \to t$, there is an M such that $t \in [0, M]$ and $\{t_n\}_{n \in \mathbb{N}} \subset [0, M]$. Since λI is continuous, $d(\phi_n, \lambda I) \to 0$ is equivalent to the uniform convergence of ϕ_n to λI on compacts. Hence, for all $\epsilon > 0$ there is an n_0 such that $n > n_0$ implies $|\lambda t_n - \lambda t| < \frac{\epsilon}{2}$ and $|\phi_n(s) - \lambda s| < \frac{\epsilon}{2}$ for $s \in [0, M]$, so, in particular, $|\phi_n(t_n) - \lambda t_n| < \frac{\epsilon}{2}$. Thus, $|\phi_n(t_n) - \lambda t| \leqslant |\phi_n(t_n) - \lambda t_n| + |\lambda t_n - \lambda t| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$ and hence $\phi_n(t_n) \to \lambda t$. Note that this is true for any sequence of times converging to t.

Since $d(x_n, x) \to 0$ we know the conditions of Theorem (1) hold, more specifically, we know they hold for sequences of times $\{\phi_n(t_n)\}_{n\in\mathbb{N}}$ converging to λt. Thus,

$$\begin{aligned} & \lim_{n\to\infty} |x_n \circ \phi_n(t_n) - x \circ \lambda I(t)| \wedge |x_n \circ \phi_n(t_n) - x \circ \lambda I(t-)| \\ = & \lim_{n\to\infty} |x_n(\phi_n(t_n)) - x(\lambda t)| \wedge |x_n(\phi_n(t_n)) - x((\lambda t)-)| \\ = & \ 0 \end{aligned}$$

and so since t and $\{t_n\}_{n\in\mathbb{N}}$ were arbitrary, the first condition holds for $\{x_n \circ \phi_n\}_{n\in\mathbb{N}}$ and $x \circ \lambda I$. For the second condition, suppose $|x_n \circ \phi_n(t_n) - x \circ \lambda I(t)| \to 0$, $s_n \geqslant t_n$, and $s_n \to t$. Each ϕ_n is non-decreasing so $\phi_n(s_n) \geqslant \phi_n(t_n)$ for all n and we know that $\phi_n(s_n) \to \lambda t$. The second condition holds for $\{x_n\}_{n\in\mathbb{N}}$ and x so $|x_n \circ \phi_n(s_n) - x \circ \lambda I(t)| = |x_n(\phi_n(s_n)) - x(\lambda t)| \to 0$ hence the second condition holds for $\{x_n \circ \phi_n\}_{n\in\mathbb{N}}$ and $x \circ \lambda I$. By a similar argument, the third condition is also seen to hold for $\{x_n \circ \phi_n\}_{n\in\mathbb{N}}$ and $x \circ \lambda I$. Therefore, by Theorem 1, $d(x_n \circ \phi_n, x \circ \lambda I) \to 0$, and so ψ is continuous at all points of $D \times \{\lambda I\}$. Hence $D_\psi \cap (D \times \{\lambda I\}) = \emptyset$, so $\tilde{P}(D_\psi) = \tilde{P}(D_\psi \cap (D \times \{\lambda I\}) = 0$ and the result follows.

□

3.3 The Wiener Process Approximation

We discuss in this section the result of Iglehart (1969) on the weak convergence of a sequence of reserve processes $\{R^{(n)}\}_{n\in\mathbb{N}}$ to a Wiener process $R^{(\infty)}$. Let $\{Y^{(n)}\}_{n\in\mathbb{N}}$ and $\{\chi_k^{(n)}\}_{k=1,n\in\mathbb{N}}^{n+1}$ be as in Theorem (7) and define a new sequence $\{X^{(n)}\}_{n\in\mathbb{N}}$ by $X_t^{(n)} = \frac{1}{\varsigma\sqrt{n}} \sum_{k=1}^{\lfloor nt \rfloor} \chi_k^{(n)}$ for $t \in [0,1]$. If we show $d_1(Y^{(n)}, X^{(n)}) \xrightarrow{\mathbf{P}} 0$, then, since $Y^{(n)} \Rightarrow W$, we can then apply Theorem (6) to conclude that $X^{(n)} \Rightarrow W$ where W

is a standard Wiener process. Iglehart uses a more elaborate result of Liggett and Rosén (1968) to conclude that $X^{(n)} \Rightarrow W$ although Theorem (6) suffices. In addition, Iglehart does not explicitly show that $d_1(Y^{(n)}, X^{(n)}) \xrightarrow{\mathbf{P}} 0$ and so it is shown here for completeness.

Proposition 1 *With* $\{Y^{(n)}\}_{n\in\mathbb{N}}$ *and* $\{X^{(n)}\}_{n\in\mathbb{N}}$ *as above,* $d_1(Y^{(n)}, X^{(n)}) \xrightarrow{\mathbf{P}} 0$.

Proof: Let $\omega \in \Omega$ and $n \in \mathbb{N}$. Since the choice $\lambda(t) = t$ yields $\|\lambda\|_{\Lambda_1} = 0$ we have that the set $\{\epsilon > 0 : \sup_{t\in[0,1]} \left|Y_t^{(n)}(\omega) - X_t^{(n)}(\omega)\right| < \epsilon\}$ is contained in the set $\{\epsilon > 0 : \exists \lambda \in \Lambda_1, \|\lambda\|_{\Lambda_1} \vee \sup_{t\in[0,1]} \left|Y_t^{(n)}(\omega) - X_{\lambda(t)}^{(n)}(\omega)\right| < \epsilon\}$ and so

$$\begin{aligned}
d_1(Y^{(n)}, X^{(n)})(\omega) &= d_1(Y^{(n)}(\omega), X^{(n)}(\omega)) \\
&= \inf\left\{\epsilon > 0 : \exists\lambda \in \Lambda_1, \|\lambda\|_{\Lambda_1} \vee \sup_{t\in[0,1]} \left|Y_t^{(n)}(\omega) - X_{\lambda(t)}^{(n)}(\omega)\right| < \epsilon\right\} \\
&\leqslant \inf\left\{\epsilon > 0 : \sup_{t\in[0,1]} \left|Y_t^{(n)}(\omega) - X_t^{(n)}(\omega)\right| < \epsilon\right\} \\
&= \sup_{t\in[0,1]} \left|Y_t^{(n)}(\omega) - X_t^{(n)}(\omega)\right| = \sup_{t\in[0,1]} \frac{nt - \lfloor nt\rfloor}{\zeta\sqrt{n}} \left|\chi_{\lfloor nt\rfloor+1}^{(n)}(\omega)\right| \\
&= \sup_{t\in(0,1)} \frac{1}{\zeta\sqrt{n}} \left|\chi_{\lfloor nt\rfloor+1}^{(n)}(\omega)\right| = \max_{1\leqslant k\leqslant n} \frac{1}{\zeta\sqrt{n}} \left|\chi_k^{(n)}(\omega)\right|
\end{aligned}$$

Thus, for any $a > 0$,

$$\begin{aligned}
&\mathbf{P}\{\omega : d(Y^{(n)}(\omega), X^{(n)}(\omega)) > a\} \leqslant \mathbf{P}\left\{\omega : \max_{1\leqslant k\leqslant n} \frac{1}{\zeta\sqrt{n}} \left|\chi_k^{(n)}(\omega)\right| > a\right\} \\
&= \mathbf{P}\left\{\omega : \max_{1\leqslant k\leqslant n} \left|\chi_k^{(n)}(\omega)\right| > \zeta\sqrt{n}a\right\} \leqslant \frac{\mathbf{E}\left[(\max_{1\leqslant k\leqslant n} |\chi_k^{(n)}|)^{2+\delta}\right]}{(\zeta\sqrt{n}a)^{2+\delta}} \\
&= \frac{\mathbf{E}\left[\max_{1\leqslant k\leqslant n} |\chi_k^{(n)}|^{2+\delta}\right]}{(\zeta\sqrt{n}a)^{2+\delta}} \leqslant \frac{\mathbf{E}\left[\sum_{k=1}^{n} |\chi_k^{(n)}|^{2+\delta}\right]}{(\zeta\sqrt{n}a)^{2+\delta}} = \frac{n\mathbf{E}\left[\left|\chi_k^{(n)}\right|^{2+\delta}\right]}{n^{1+\delta/2}(\zeta a)^{2+\delta}} = \frac{\mathbf{E}\left[\left|\chi_k^{(n)}\right|^{2+\delta}\right]}{n^{\delta/2}(\zeta a)^{2+\delta}}
\end{aligned}$$

where Chebychev's inequality is used with power $2 + \delta > 0$ to pass to the expectation. Now, choose $\delta > 0$ so that $M = \sup_{n\in\mathbb{N}} \mathbf{E}\left[\left|\chi_k^{(n)}\right|^{2+\delta}\right] < \infty$. Thus, $\forall n \in \mathbb{N}$,

$\mathbf{P}\{\omega : d(Y^{(n)}, X^{(n)})(\omega) > a\} \leqslant \frac{M}{n^{\delta/2}(\zeta a)^{2+\delta}}$ and so $d_1(Y^{(n)}, X^{(n)}) \xrightarrow{\mathbf{P}} 0.$

$\square$

Thus, $X^{(n)} \Rightarrow W$. If we set $N_t^{(n)} = \frac{N^\chi_{nt}}{n}$ then from Theorem (17.3) in Billingsley (1968) we have that $N_t^{(n)} \Rightarrow \lambda t$. Now, applying Theorem (8), we conclude that

$$\frac{1}{\zeta\sqrt{n}}\sum_{k=1}^{N^\chi_{nt}} \chi_k^{(n)} = \frac{1}{\zeta\sqrt{n}}\sum_{k=1}^{\lfloor nN_t^{(n)}\rfloor} \chi_k^{(n)} = (X^{(n)} \circ N^{(n)})_t \Rightarrow (W \circ \lambda I)_t = W_{\lambda t}$$

so, using the self-similarity of Wiener processes, $\frac{1}{\sqrt{n}}\sum_{k=1}^{N^\chi_{nt}} \chi_k^{(n)} \Rightarrow \zeta W_{\lambda t} \overset{\mathcal{D}}{=} \zeta\sqrt{\lambda}W_t$.

We can now present Iglehart's main convergence theorem and a slight simplification of his proof.

Theorem 9 (Iglehart, 1969) *Let N^χ be a renewal counting process with finite mean inter-claim time $\lambda^{-1} > 0$. Define a sequence $\{R^{(n)}\}_{n\in\mathbb{N}}$ of risk processes by*

$$R_t^{(n)} = \frac{1}{\sqrt{n}}\left(u^{(n)} + \pi^{(n)}nt - \sum_{k=1}^{N^\chi_{nt}} \chi_k^{(n)}\right) \quad for \quad t \in [0,1]$$

where for each $n \in \mathbb{N}$, $\{\chi_k^{(n)}\}_{k=1}^{n+1}$ are i.i.d. with $\mathbf{E}[\chi_k^{(n)}] = \mu^{(n)} > 0$, $\mathbf{V}[\chi_k^{(n)}] = {\sigma^{(n)}}^2 > 0$, and, $\exists\delta > 0$ such that $\sup_{n\in\mathbb{N}} \mathbf{E}[|\chi_k^{(n)}|^{2+\delta}] < \infty$. Suppose also that $u^{(n)} = u\sqrt{n} + o(\sqrt{n})$, $\pi^{(n)} = \frac{\pi}{\sqrt{n}} + o(\frac{1}{\sqrt{n}})$, $\mu^{(n)} = \frac{\mu}{\sqrt{n}} + o(\frac{1}{\sqrt{n}})$, and ${\sigma^{(n)}}^2 \to \zeta^2 > 0$ for constants $u, \pi, \mu, \zeta > 0$. Define $R_t^{(\infty)} = u + (\pi - \mu\lambda)t - \zeta\sqrt{\lambda}W_t$ for $t \in [0,1]$. Then, $R^{(n)} \Rightarrow R^{(\infty)}$.

Proof: First, write $R_t^{(n)}$ in the following form:

$$\begin{aligned}
R_t^{(n)} &= \frac{u^{(n)}}{\sqrt{n}} + \pi^{(n)}\sqrt{n}t - \frac{1}{\sqrt{n}}\sum_{k=1}^{N^\chi_{nt}}(\chi_k^{(n)} - \mu^{(n)}) - \frac{N^\chi_{nt}\mu^{(n)}}{\sqrt{n}} \\
&= \left(u + \frac{o(\sqrt{n})}{\sqrt{n}}\right) + \left(\pi + \frac{o(1/\sqrt{n})}{1/\sqrt{n}}\right)t - \frac{1}{\sqrt{n}}\sum_{k=1}^{N^\chi_{nt}}(\chi_k^{(n)} - \mu^{(n)}) \\
&\quad - \frac{N^\chi_{nt}}{n}\left(\mu + \frac{o(1/\sqrt{n})}{1/\sqrt{n}}\right)
\end{aligned}$$

By Theorem (17.3) in Billingsley (1968), $\frac{N^\chi_{nt}}{n} \Rightarrow \lambda t$ and clearly $\mu + \frac{o(1/\sqrt{n})}{1/\sqrt{n}} \Rightarrow \mu$ so by Theorem (4), $\frac{N^\chi_{nt}}{n}\left(\mu + \frac{o(1/\sqrt{n})}{1/\sqrt{n}}\right) \Rightarrow \mu\lambda t$, a non-random function. Also, it is clear that

$\left(u + \frac{o(\sqrt{n})}{\sqrt{n}}\right) + \left(\pi + \frac{o(1/\sqrt{n})}{1/\sqrt{n}}\right) t \Rightarrow u + \pi t$ and so again by Theorem (4) we have that $\left(u + \frac{o(\sqrt{n})}{\sqrt{n}}\right) + \left(\pi + \frac{o(1/\sqrt{n})}{1/\sqrt{n}}\right) t - \frac{N^X_{nt}}{n} \left(\mu + \frac{o(1/\sqrt{n})}{1/\sqrt{n}}\right) \Rightarrow u + \pi t - \mu\lambda t = u + (\pi - \mu\lambda)t$. The random variables $\chi_k^{(n)} - \mu^{(n)}$ satisfy the conditions of Theorem (7) so by the previous argument, $\frac{1}{\sqrt{n}} \sum_{k=1}^{N^X_{nt}} (\chi_k^{(n)} - \mu^{(n)}) \Rightarrow \zeta\sqrt{\lambda} W_t$. Applying Theorem (4) once again yields $R_t^{(n)} \Rightarrow u + (\pi - \mu\lambda)t - \zeta\sqrt{\lambda}W_t \stackrel{\mathcal{D}}{=} u + (\pi - \mu\lambda)t + \zeta\sqrt{\lambda}W_t$.

□

Iglehart extends these convergence results to $D[0,\infty)$ using a result of Stone (1963). However, Lindvall (1973) has corrected a small error in Stone's theorem and we therefore use Lindvall's version. Given a probability measure P on $D[0,\infty)$ define $T_P = \{t : P\{x : x(t) = x(t-)\} = 1\}$. For $T \in \mathbb{R}_+$, define the time index restriction map $r_T : D[0,\infty) \to D[0,T]$ by $(r_T(x))(t) = x(t)$ for $t \in [0,T]$.

Theorem 10 (Lindvall, 1973) *If P and $\{P^{(n)}\}_{n\in\mathbb{N}}$ are probability measures on $D[0,\infty)$ then $P^{(n)} \Rightarrow P$ if and only if $P^{(n)}r_T^{-1} \Rightarrow Pr_T^{-1}$ for all $T \in T_P$.*

Now, extend $\{R^{(n)}\}_{n\in\mathbb{N}}$ and $R^{(\infty)}$ to $D[0,T]$ and $D[0,\infty)$ by simply enlarging the array of random variables $\{\chi_k^{(n)}\}_{k=1,n\in\mathbb{N}}^{n+1}$ to $\{\chi_k^{(n)}\}_{k,n\in\mathbb{N}}$ retaining the i.i.d. properties as before and allowing $t \in [0,T]$ or $t \in \mathbb{R}_+$. All of Billingsley's results carry over to $(D[0,T], d_T)$ for all $T \in \mathbb{R}_+$ and hence Iglehart's convergence results holds as well on $(D[0,T], d_T)$ for all $T \in \mathbb{R}_+$. Thus, $\mathbf{P}(r_T(R^{(n)}))^{-1} = \mathbf{P}R^{(n)^{-1}}r_T^{-1} \Rightarrow \mathbf{P}R^{(\infty)^{-1}}r_T^{-1} = \mathbf{P}(r_T(R^{(\infty)}))^{-1}$ for all $T \in \mathbb{R}_+$ which, by Theorem (10), is more than sufficient to ensure that $\mathbf{P}R^{(n)^{-1}} \Rightarrow \mathbf{P}R^{(\infty)^{-1}}$, ie, $R^{(n)} \Rightarrow R^{(\infty)}$ on (D,d).

The projection functional $\pi_t : D \to \mathbb{R}$ defined by $\pi_t(x) = x(t)$ is $\mathcal{B}(\tau)/\mathcal{B}(\mathbb{R})$-measurable and almost surely continuous with respect to the measure $\mathbf{P}R^{(\infty)^{-1}}$ (Lindvall 1973). Since $\mathbf{P}R^{(n)^{-1}} \Rightarrow \mathbf{P}R^{(\infty)^{-1}}$ on D we can apply Theorem (5), which applies to any complete, separable metric space, to conclude that $\mathbf{P}R^{(n)^{-1}}\pi_t^{-1} \Rightarrow \mathbf{P}R^{(\infty)^{-1}}\pi_t^{-1}$. Thus,

$$\mathbf{P}\{R_t^{(n)} \leqslant x\} = \mathbf{P}R^{(n)^{-1}}\pi_t^{-1}(-\infty, x] \to \mathbf{P}R^{(\infty)^{-1}}\pi_t^{-1}(-\infty, x] = \mathbf{P}\{R_t^{(\infty)} \leqslant x\}$$

using the fact that $\partial(-\infty, x] = \{x\}$ has zero measure since $R_t^{(\infty)}$ is a continuous random variable (see Billingsley (1968)). Similar to the derivation of (3.1), we can obtain a tractable expression for $\mathbf{P}\{R_t^{(\infty)} \leqslant x\}$.

The functional $T^r : D \to \mathbb{R}$ defined by $T^r(x) = \inf\{t > 0 : x(t) < 0\}$ is also $\mathcal{B}(\tau)/\mathcal{B}(\mathbb{R})$-measurable and continuous almost surely with respect to $\mathbf{P}R^{(\infty)^{-1}}$ on $D[0,\infty)$ since $x \in D$ is measurable and the Wiener measure $\mathbf{P}R^{(\infty)^{-1}}$ corresponds to a process with $\mathbf{P}$-a.s. continuous sample paths (Stroock 1993). Applying Theorem (5) again we have that $\mathbf{P}R^{(n)^{-1}}(T^r)^{-1} \Rightarrow \mathbf{P}R^{(\infty)^{-1}}(T^r)^{-1}$ and thus

$$\begin{aligned}
\mathbf{P}\{T^r(R^{(n)}) \leqslant t\} &= \mathbf{P}R^{(n)^{-1}}T^{r^{-1}}(-\infty, t] \\
&\to \mathbf{P}R^{(\infty)^{-1}}T^{r^{-1}}(-\infty, t] = \mathbf{P}\{T^r(R^{(\infty)}) \leq t\}
\end{aligned}$$

whose distribution is as in (3.2). It should be noted that these convergence results apply to any model having a renewal counting process N^X, without explosion, for claim arrivals which need not be independent of the claim severities. However, it is seen that if the claims are not distributed in a reasonably symmetric manner this approximation doesn't perform well (Gluckman 1979). Furthermore, it is often the case that claim severities are highly skewed with infinite second moments, ruling out altogether the use of the Wiener process approximation (Embrechts & Veraverbeeke 1982). For example, Pareto, LogGamma, Weibull distributions and others are commonly used to model claims distributions (Hogg & Klugman 1984). As a way of dealing with this problem, Furrer et al. (1996) have considered a larger class of approximating processes based on α-stable distributions, of which the Wiener process approximation is a close neighbor.

3.4 α-Stable Lévy Processes

Definition 1 *A random variable X has a stable distribution if it has a non-empty domain of attraction: there is an i.i.d. sequence random variables $\{X_n\}_{n\in\mathbb{N}}$, with common distribution function F, and sequences of real numbers $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$*

with $a_n > 0$ such that $\frac{1}{a_n}\sum_{k=1}^{n}(X_k - b_n) \Rightarrow X$ in which case we say that F is in the domain of attraction of X.

Thus, stable distributions are the only possible distributions that can arise as weak limits of sequences of normalized partial sums of i.i.d. random variables, which, by the Central Limit Theorem, includes the family $N(\mu, \sigma^2)$ of normal distributions. The stable distributions also possess another interesting property : infinite divisibility.

Definition 2 *A random variable X with law P and characteristic function $\widehat{P}(\theta)$ is infinitely divisible (ID) if one of the following equivalent conditions holds:*
(1) $\forall n \in \mathbb{N}\ \exists\ n$ i.i.d. random variables $\{X_k^{(n)}\}_{k=1}^{n}$ such that $X \stackrel{\mathcal{D}}{=} X_1^{(n)} + \cdots + X_n^{(n)}$
(2) $\forall n \in \mathbb{N}\ \exists\ P_n \in \mathcal{P}_{\mathbb{R}}$ such that $P = P_n^{*n}$
(3) $\forall n \in \mathbb{N}\ \exists\ \widehat{P_n}(\theta)$ such that $\widehat{P}(\theta) = \left[\widehat{P_n}(\theta)\right]^n$ and $\widehat{P_n}(0) = 1$

P_n and $\widehat{P_n}(\theta)$ are the law and characteristic function of the $\{X_k^{(n)}\}_{k=1}^{n}$. A beautiful classical result completely characterizes ID laws in terms of their characteristic functions:

Theorem 11 (Lévy-Khintchine) *A random variable X is infinitely divisible if and only if its characteristic function is of the form $\widehat{P}(\theta) = e^{\phi(\theta)}$ where*

$$\phi(\theta) = i\theta\mu - \frac{1}{2}\theta^2\tilde{\sigma}^2 + \int_{\mathbb{R}\setminus\{0\}} \left(e^{i\theta x} - 1 - i\theta\psi(x)\right) dM(x)$$

for some $\mu \in \mathbb{R}$, $\tilde{\sigma} \in \mathbb{R}_+$, $M \in \mathcal{M}_{\mathbb{R}\setminus\{0\}}^{1\wedge x^2}$, and ψ is a bounded measurable function $\psi : \mathbb{R}\setminus\{0\} \to \mathbb{R}\setminus\{0\}$ satisfying $\sup_{\mathbb{R}\setminus\{0\}}\left|\frac{\psi(x)-x}{x^2}\right| < \infty$.

Proof: See Stroock (1993).

The above representation is not unique since there is considerable freedom in the choice of ψ. For a fixed choice of ψ the above representation is unique, and, for this reason, $(\mu, \tilde{\sigma}^2, M)$ are called the Lévy characteristics of X and M is the corresponding Lévy measure of X. The original choice $\psi(x) = \frac{x}{1+x^2}$ of Lévy and Khintchine together with the definition of a stable distribution leads to the Lévy characteristics for stable random variables $(\mu, \tilde{\sigma}^2, 0)$ and $(\mu, 0, M)$ where $\mu \in \mathbb{R}, \tilde{\sigma} \in \mathbb{R}_+$, and $M \in \mathcal{M}^{1\wedge x^2}_{\mathbb{R}\setminus\{0\}}$ is given by

$$dM(x) = \frac{Q}{|x|^{\alpha+1}} 1_{(-\infty,0)}(x)dx + \frac{P}{x^{\alpha+1}} 1_{(0,\infty)}(x)dx$$

for some $\alpha \in (0,2)$ and $P, Q \in \mathbb{R}_+$ such that $P + Q > 0$. P and Q give the relative weighting of positive and negative values of X (Kolmogorov & Gnedenko 1968).

The first set of characteristics $(\mu, \tilde{\sigma}^2, 0)$ when $\tilde{\sigma} = 0$ yields $\widehat{P}(\theta) = e^{i\theta\mu}$ corresponding to the unit point mass at μ. The first set $(\mu, \tilde{\sigma}^2, 0)$ when $\tilde{\sigma} > 0$ yields $\widehat{P}(\theta) = e^{i\theta\mu - \frac{1}{2}\theta^2\tilde{\sigma}^2}$ which corresponds to a $N(\mu, \tilde{\sigma}^2)$ law. The second set $(\mu, 0, M)$ is of primary interest here. Setting $\beta = \frac{P-Q}{P+Q} \in [-1, 1]$ to represent skew and performing an integration as described in Feller (1966), yields the explicit representation

$$\ln \widehat{P}(\theta) = \begin{cases} -\sigma^\alpha |\theta|^\alpha \left[1 - i\beta \operatorname{sign}(\theta) \tan \frac{\pi\alpha}{2}\right] + i\mu\theta & \text{if } \alpha \in (0,1) \cup (1,2) \\ -\sigma|\theta| \left[1 + \frac{2i\beta}{\pi} \operatorname{sign}(\theta) \ln|\theta|\right] + i\mu\theta & \text{if } \alpha = 1 \end{cases}$$

where $\operatorname{sign}(\theta) = 1_{(0,\infty)}(\theta) - 1_{(-\infty,0)}(\theta)$ and $\sigma > 0$ is a constant (different to the $\tilde{\sigma}$ appearing in the Lévy characteristics). The parameters α, β, σ and μ are unique. Setting $\sigma = 0$ in the above characteristic function yields the law δ_μ, the unit mass at μ. In this case, the parameters α and β are irrelevant. Setting $\alpha = 2$ in the above yields a $N(\mu, 2\sigma^2)$ distribution, in which case β is irrelevant. Thus, the above family of characteristic functions includes all stable characteristics functions if we allow the cases $\sigma = 0$ and $\alpha = 2$ and accept non-uniqueness of the irrelevant parameters. Denote the entire class of α-stable distributions by $S_\alpha(\sigma, \beta, \mu)$ where $\alpha \in (0, 2]$ is the index of stability, $\sigma \in \mathbb{R}_+$ is the dispersion, $\beta \in [-1, 1]$ is the skewness, and $\mu \in \mathbb{R}$ is the location. Since the Dirac mass δ_μ has exceptional properties, we exclude it from further consideration by the restriction $\sigma > 0$, which will be implicitly assumed from here on. The following arithmetic properties of stable random variables can be deduced directly from the form of the characteristic function above.

Proposition 2 *If* $X \sim S_\alpha(\sigma, \beta, \mu)$ *then for any* $a \in \mathbb{R}$, $X + a \sim S_{\alpha,}(\sigma, \beta, \mu + a,)$

Proposition 3 *If* $X \sim S_\alpha(\sigma, \beta, \mu)$ *then for any* $a \in \mathbb{R} \setminus \{0\}$,

$$aX \sim \begin{cases} S_\alpha(|a|\sigma, \beta\text{sign}(a), a\mu) & \text{if } \alpha \neq 1 \\ S_\alpha(|a|\sigma, \beta\text{sign}(a), a\mu - \frac{2a\sigma\beta}{\pi}\ln|a|) & \text{if } \alpha = 1 \end{cases}$$

Proposition 4 *If* $Z \sim S_\alpha(1, \beta, 0)$, $\sigma > 0$, *and* $\mu \in \mathbb{R}$ *then*

$$\begin{array}{ll} X = \sigma Z + \mu \sim S_\alpha(\sigma, \beta, \mu) & \text{if } \alpha \neq 1 \\ X = \sigma Z + (\mu + \frac{2\sigma\beta}{\pi}\ln(\sigma)) \sim S_\alpha(\sigma, \beta, \mu) & \text{if } \alpha = 1 \end{array}$$

Thus, with a simple scale-location transformation we can move from $S_\alpha(1, \beta, 0)$ to $S_\alpha(\sigma, \beta, \mu)$ and so for simulation purposes we only need to consider distributions from $S_\alpha(1, \beta, 0)$.

Proposition 5 *If* $X_i \sim S_\alpha(\sigma_i, \beta_i, \mu_i)$, $i = 1, 2$ *are two independent random variables then* $X_1 + X_2 \sim S_\alpha(\sigma, \beta, \mu)$ *where*

$$\beta = \frac{\beta_1\sigma_1^\alpha + \beta_2\sigma_2^\alpha}{\sigma_1^\alpha + \sigma_2^\alpha}, \quad \mu = \mu_1 + \mu_2, \quad \sigma = (\sigma_1^\alpha + \sigma_2^\alpha)^{1/\alpha}$$

Proposition 6 *Let* $\{X_k\}_{k \in \mathbb{N}}$ *be i.i.d. random variables distributed as* $S_\alpha(\sigma, \beta, \mu)$. *Then,*

$$\begin{array}{ll} \frac{1}{n^{1/\alpha}}\sum_{k=1}^n (X_k - \mu) \sim S_\alpha(\sigma, \beta, 0) & \text{if } \alpha \neq 1 \\ \frac{1}{n^{1/\alpha}}\sum_{k=1}^n (X_k - \mu - \frac{2\sigma\beta}{\pi}\ln n) \sim S_\alpha(\sigma, \beta, 0) & \text{if } \alpha = 1 \end{array}$$

Surprisingly, the inversion of the characteristic function of a stable distribution is known explicitly in only four cases, despite the fact that it is known that they are continuous, unimodal distributions (Zolotarev 1986).

1.) The Normal distribution $S_2(\sigma, \beta, \mu) = N(\mu, 2\sigma^2)$ with density

$$f(x) = \frac{1}{\sqrt{4\pi\sigma^2}} \exp\left\{\frac{-(x-\mu)^2}{4\sigma^2}\right\}$$

2.) The Cauchy distribution $S_1(\sigma, 0, \mu)$ with density

$$f(x) = \frac{\sigma}{\pi((x-\mu)^2 + \sigma^2)}$$

3.) The Lévy distribution $S_{\frac{1}{2}}(\sigma, 1, \mu)$ with density

$$f(x) = \left(\frac{\sigma}{2\pi}\right)^{\frac{1}{2}} \frac{1}{(x-\mu)^{3/2}} \exp\left\{-\frac{\sigma}{2(x-\mu)}\right\} 1_{(\mu,\infty)}(x)$$

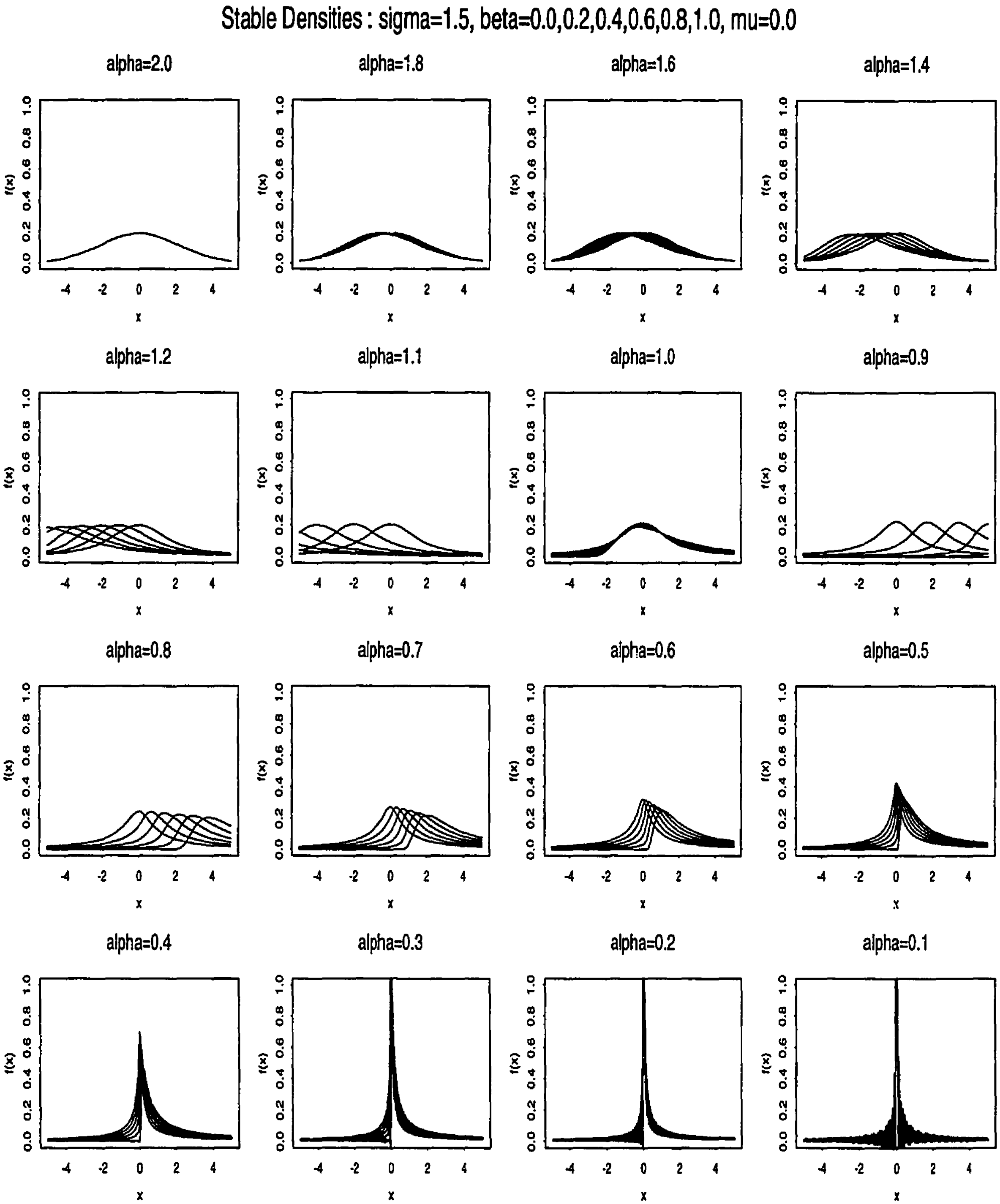
and its reflection $f(-x)$.

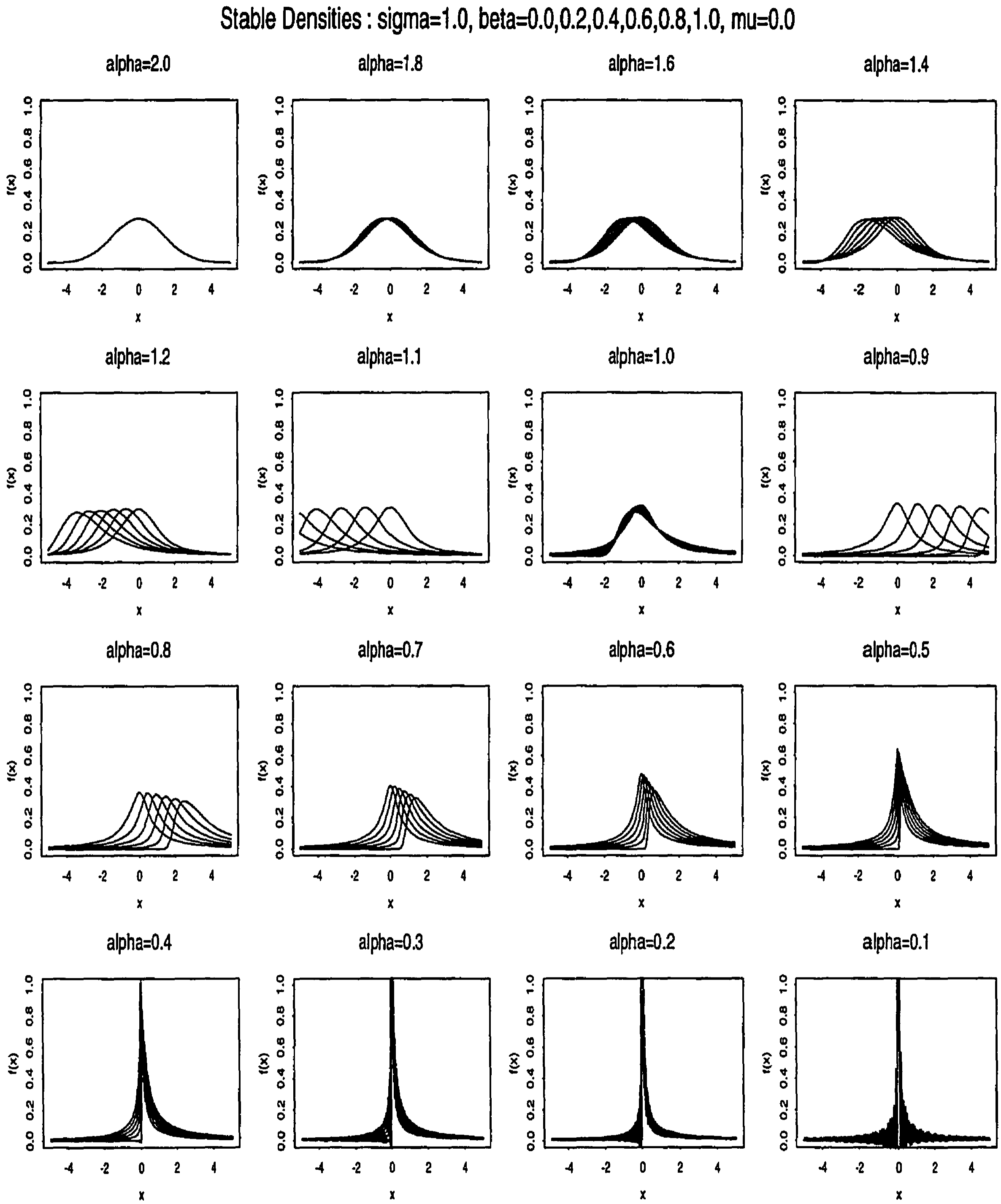
4.) The Dirac delta function $S_\alpha(0, 0, \mu)$, $\delta_\mu(x)$.

The next three figures are densities from $S_\alpha(\sigma, \beta, 0)$ for various values of α, β and σ and were obtained by numerically computing

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\theta x} \widehat{P}(\theta) d\theta.$$

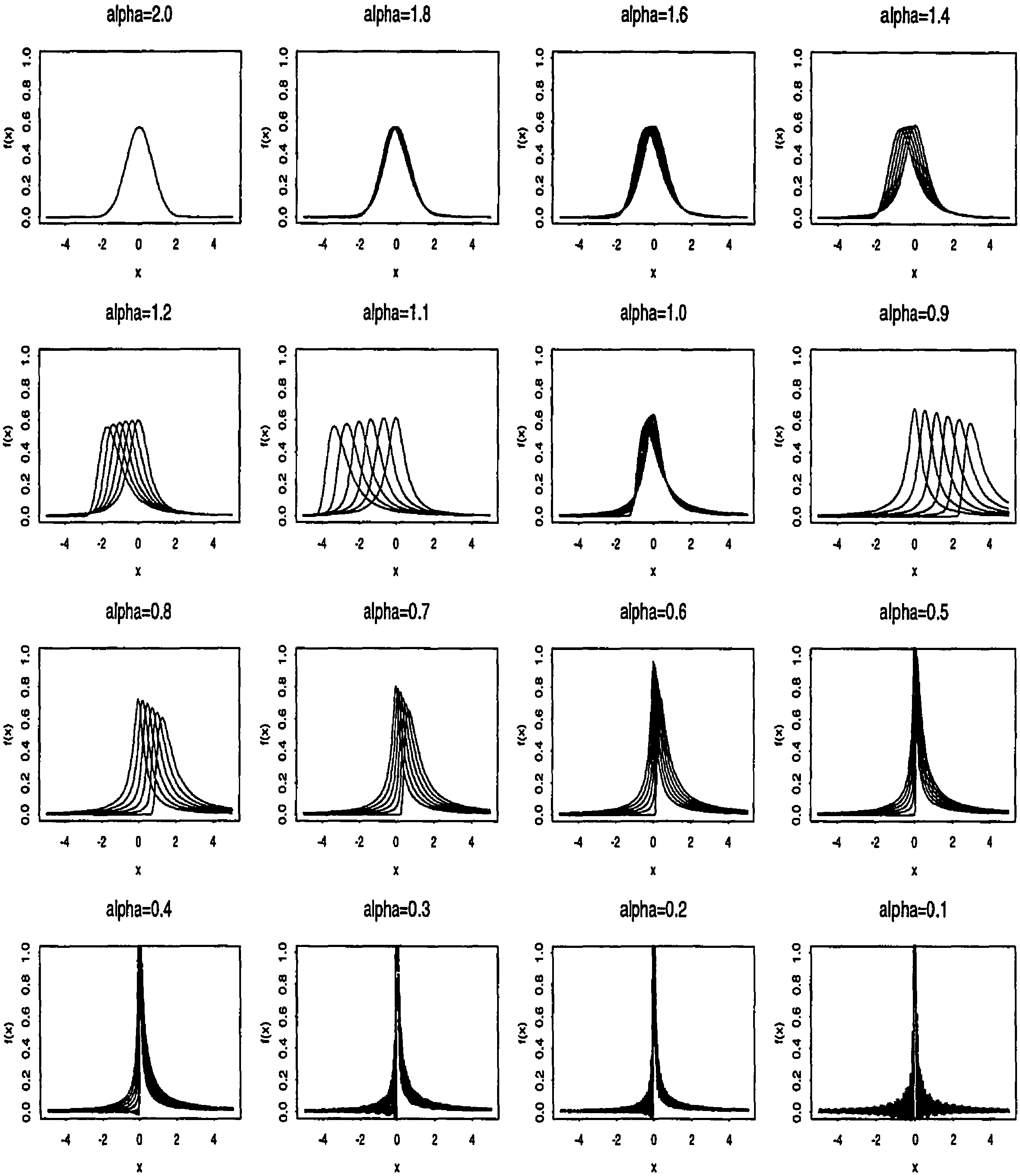
$\mu = 0$ is chosen for convenience since it is merely a location parameter. It is clear from the graphs that $S_\alpha(\sigma, \beta, \mu)$ is a very rich family.


Stable Densities : sigma=1.5, beta=0.0,0.2,0.4,0.6,0.8,1.0, mu=0.0
alpha=2.0
alpha=1.8
alpha=1.6
alpha=1.4
alpha=1.2
alpha=1.1
alpha=1.0
alpha=0.9
alpha=0.8
alpha=0.7
alpha=0.6
alpha=0.5
alpha=0.4
alpha=0.3
alpha=0.2
alpha=0.1
f(x)
x



Stable Densities : sigma=1.0, beta=0.0,0.2,0.4,0.6,0.8,1.0, mu=0.0
alpha=2.0
alpha=1.8
alpha=1.6
alpha=1.4
alpha=1.2
alpha=1.1
alpha=1.0
alpha=0.9
alpha=0.8
alpha=0.7
alpha=0.6
alpha=0.5
alpha=0.4
alpha=0.3
alpha=0.2
alpha=0.1
f(x)
x

Stable Densities : sigma=0.5, beta=0.0,0.2,0.4,0.6,0.8,1.0, mu=0.0



Note that as $\alpha \to 0$, our numerical inversion of the characteristic function becomes inaccurate due to the sensitivity of the inversion to the truncation of the integration range. Note also that as $\alpha \to 1$ from above or below we see that increases in β have the effect of spreading the densities along the axis. This singular behaviour at $\alpha = 1$ is merely a consequence of the particular way the family $S_\alpha(\sigma, \beta, \mu)$ has been parametrized via the choice of ψ. Reparametrizing by a different choice of ψ would have the effect of altering μ, P, and Q only (and hence β and σ) but does not alter the value of α. For instance, Feller (1966) and Zolotarev (1986) work with the functions

$$\psi(x) = \sin x \qquad \psi(x) = 1_{(1,\infty)}(x) + x1_{[-1,1]}(x) - 1_{(-\infty,1)}(x) \qquad \psi(x) = x1_{[-1,1]}(x)$$

Another method to reparametrize $S_\alpha(\sigma, \beta, \mu)$, employed extensively by Zolotarev (1986), Chambers, Mallows & Stuck (1976), and Kanter (1975), other than a different choice of ψ, is to reparametrize the characteristic function for the case $\alpha \neq 1$ by the following scheme. Consider the characteristic function for $\alpha \neq 1$, $\ln \widehat{P}(\theta) = -\sigma^\alpha |\theta|^\alpha \left[1 - i\beta \text{sign}(\theta) \tan \frac{\pi\alpha}{2}\right] + i\mu\theta$, and represent the first term in complex polar form:

$$\begin{aligned}
& -\sigma^\alpha |\theta|^\alpha \left[1 - i\beta \text{sign}(\theta) \tan\left(\frac{\pi\alpha}{2}\right)\right] \\
= \;& -\sigma_2^\alpha |\theta|^\alpha \exp\left[-i\beta_2 \text{sign}(\theta) \frac{\pi}{2} K(\alpha)\right] \\
= \;& -\sigma_2^\alpha |\theta|^\alpha \cos\left(\beta_2 \frac{\pi}{2} K(\alpha)\right) \left[1 - i\text{sign}(\theta) \tan\left(\beta_2 \frac{\pi}{2} K(\alpha)\right)\right]
\end{aligned}$$

where, for a suitable choice of $K(\alpha)$, we can find the new parameters β_2 and σ_2 from

$$\beta \tan\left(\frac{\pi\alpha}{2}\right) = \tan\left(\beta_2 \frac{\pi}{2} K(\alpha)\right) \qquad \textit{and} \qquad \sigma^\alpha = \sigma_2^\alpha \cos\left(\beta_2 \frac{\pi}{2} K(\alpha)\right)$$

With this parametrization the characteristic function takes the form

$$\ln \widehat{P}(\theta) = \begin{cases} -\sigma_2^\alpha |\theta|^\alpha \exp(-i\beta_2 \text{sign}(\theta) \frac{\pi}{2} K(\alpha)) + i\mu\theta & \text{if } \alpha \neq 1 \\ -\sigma|\theta|(1 + \frac{2i\beta}{\pi} \text{sign}(\theta) \ln|\theta|) + i\mu\theta & \text{if } \alpha = 1 \end{cases}$$

Zolotarev (1986) has used both $K(\alpha) = 1 - |1-\alpha|$ and $K(\alpha) = \alpha - 1 + \text{sign}(\alpha - 1)$, which both ensure that $\beta_2 \in [-1, 1]$, in his investigations into integral representations of stable density functions. Chambers et al. (1976) have used one of Zolotarev's integral representations involving $K(\alpha) = 1 - |1-\alpha|$ to derive a representation of densities in $S_\alpha(1, \beta_2, 0)$ in terms of a pair of independent uniform and exponential random variables. Let $U \sim Uniform(-\frac{\pi}{2}, \frac{\pi}{2})$ $E \sim Exponential(1)$ be independent and let $c = \frac{\pi\beta_2}{2\alpha}K(\alpha)$. Then

$$S_\alpha(1,\beta_2,0) = \begin{cases} \frac{\sin[\alpha(U+c)]}{(\cos U)^{1/\alpha}} \left(\frac{\cos[U-\alpha(U+c)]}{E}\right)^{\frac{1-\alpha}{\alpha}} & \text{if } \alpha \neq 1 \\ \frac{2}{\pi}\left((\frac{\pi}{2}+\beta_2 U)\tan U - \beta_2 \ln\left(\frac{\frac{\pi}{2}E\cos U}{\frac{\pi}{2}+\beta_2 U}\right)\right) & \text{if } \alpha = 1 \end{cases}$$

Chambers et al. (1976) define yet another parametrization that yields a characteristic function continuous at $\alpha = 1$ as well as give an efficient and numerically accurate algorithm for simulating these densities. As in Janicki & Weron (1994), we choose to work with the parametrization arising from the choice $\psi(x) = \frac{x}{1+x^2}$ and so a modification of the above representation must be made. This can be done by noticing that

$$c = \frac{\pi\beta_2}{2\alpha}K(\alpha) = \frac{1}{\alpha}\tan^{-1}\left[\beta\tan\left(\frac{\pi\alpha}{2}\right)\right]$$

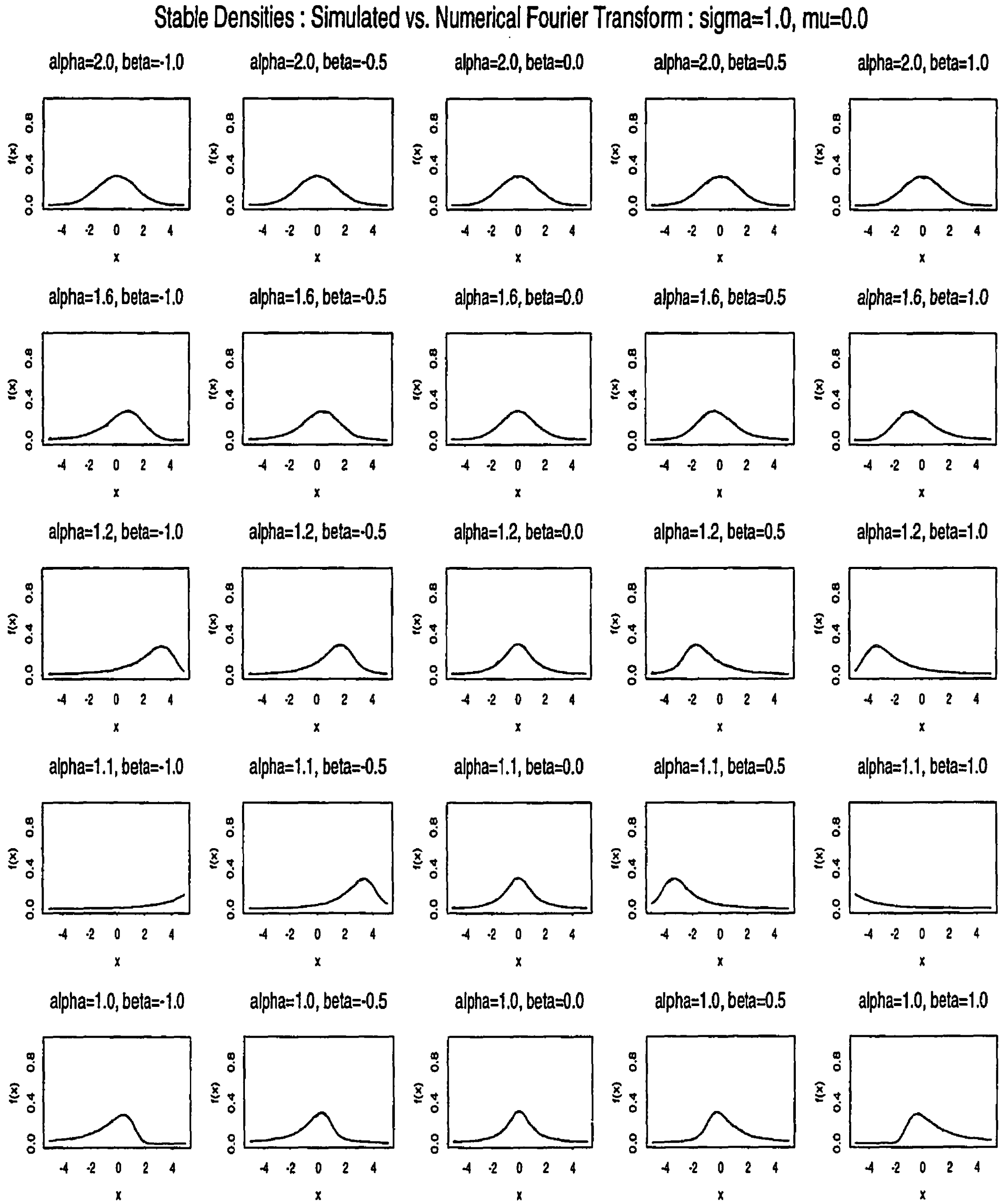
and for $\sigma_2 = 1$ we have that $\sigma = \cos[\tan^{-1}(\beta\tan(\frac{\pi\alpha}{2}))]^{\frac{1}{\alpha}}$. Thus, defining

$$C_{\alpha,\beta} = \frac{1}{\alpha}\tan^{-1}\left[\beta\tan\left(\frac{\pi\alpha}{2}\right)\right] \quad \textit{and} \quad D_{\alpha,\beta} = \cos\left(\tan^{-1}\left[\beta\tan\left(\frac{\pi\alpha}{2}\right)\right]\right)^{\frac{-1}{\alpha}}$$

we then get that

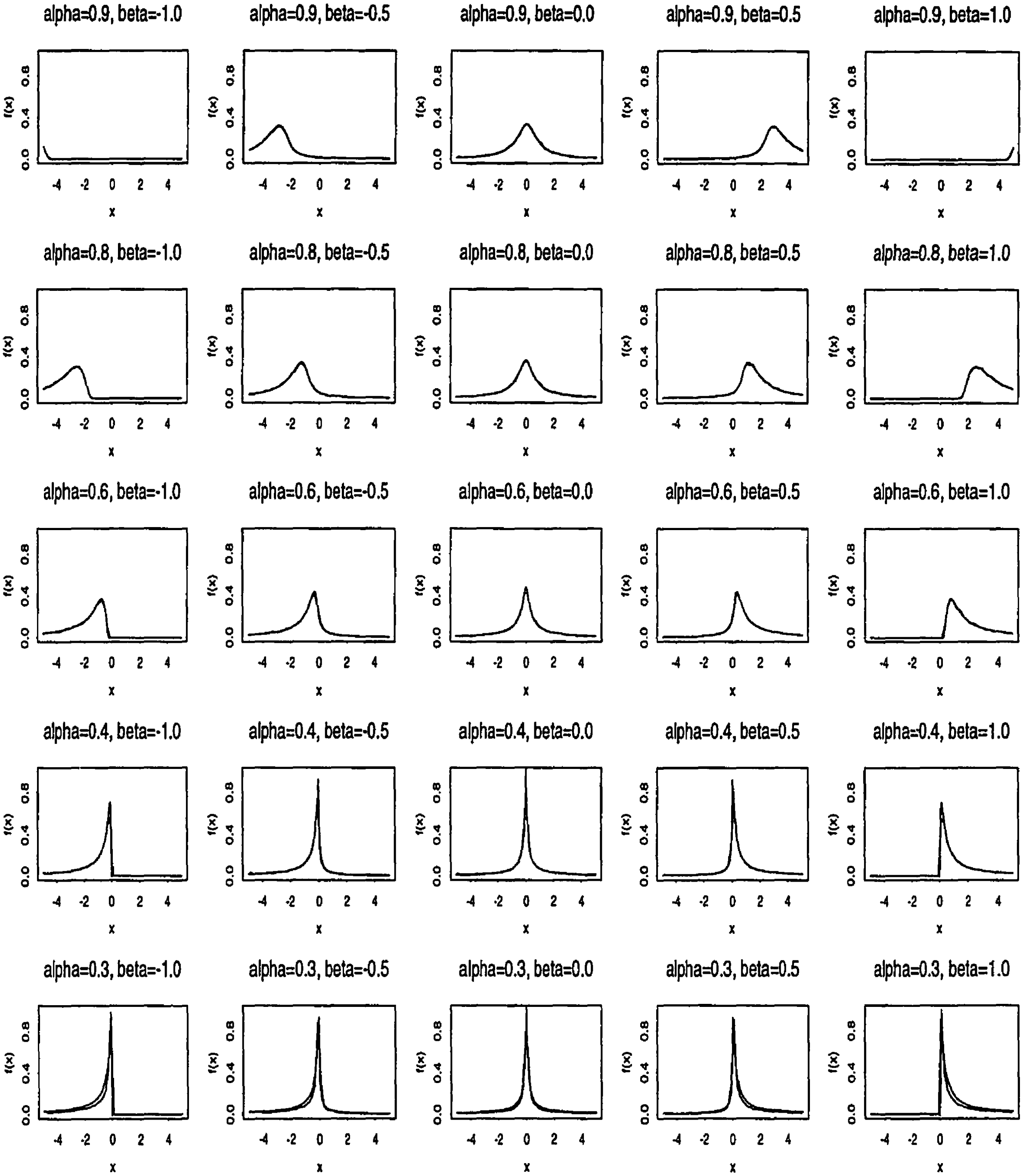
$$S_\alpha(1,\beta,0) = \begin{cases} D_{\alpha,\beta}\frac{\sin[\alpha(U+C_{\alpha,\beta})]}{(\cos U)^{1/\alpha}} \left(\frac{\cos[U-\alpha(U+C_{\alpha,\beta})]}{E}\right)^{\frac{1-\alpha}{\alpha}} & \text{if } \alpha \neq 1 \\ \frac{2}{\pi}\left((\frac{\pi}{2}+\beta U)\tan U - \beta \ln\left(\frac{\frac{\pi}{2}E\cos U}{\frac{\pi}{2}+\beta U}\right)\right) & \text{if } \alpha = 1 \end{cases}$$

Note that the case $\alpha = 1$ is unchanged. Also note that this expression was given incorrectly in Janicki & Weron (1994). The following figures numerically demonstrate its correctness by comparing the densities computed numerically using the inverse Fourier transform and the densities of 30,000 deviates simulated with the above representation. The discrepancies for the smaller values of α are due primarily to the increased sensitivity of the Fourier integral to the range truncation involved in its numerical computation.



Stable Densities : Simulated vs. Numerical Fourier Transform : sigma=1.0, mu=0.0
alpha=2.0, beta=-1.0
alpha=2.0, beta=-0.5
alpha=2.0, beta=0.0
alpha=2.0, beta=0.5
alpha=2.0, beta=1.0
alpha=1.6, beta=-1.0
alpha=1.6, beta=-0.5
alpha=1.6, beta=0.0
alpha=1.6, beta=0.5
alpha=1.6, beta=1.0
alpha=1.2, beta=-1.0
alpha=1.2, beta=-0.5
alpha=1.2, beta=0.0
alpha=1.2, beta=0.5
alpha=1.2, beta=1.0
alpha=1.1, beta=-1.0
alpha=1.1, beta=-0.5
alpha=1.1, beta=0.0
alpha=1.1, beta=0.5
alpha=1.1, beta=1.0
alpha=1.0, beta=-1.0
alpha=1.0, beta=-0.5
alpha=1.0, beta=0.0
alpha=1.0, beta=0.5
alpha=1.0, beta=1.0
f(x)
x

Stable Densities : Simulated vs. Numerical Fourier Transform : sigma=1.0, mu=0.0



We now collect some results on domains of attraction and series representations of α-stable random variables and α-stable Lévy processes.

Definition 3 *A function $L(x)$ is said to be slowly varying at infinity if for all $a > 0$,*

$$\lim_{x\to\infty} \frac{L(ax)}{L(x)} = 1$$

Theorem 12 (Mijnheer 1975) *A distribution function F is in the domain of attraction of $S_\alpha(1,\beta,0)$ for some $\alpha \in (0,2)$ and $\beta \in [-1,1]$ if and only if*
(1) $L_F(x) \stackrel{def}{=} x^\alpha[1 - F(x) + F(-x)]$ is slowly varying at infinity
(2) $\lim_{x\to\infty} \frac{F(-x)}{1-F(x)+F(-x)} = \frac{1-\beta}{2}$
Furthermore, if F is in the domain of attraction and $\{X_n\}_{n\in\mathbb{N}}$ are i.i.d. random variables distributed by F then the sequences $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$, $a_n > 0$, such that $\frac{1}{a_n}\sum_{k=1}^n (X_k - b_n) \Rightarrow S_\alpha(1,\beta,0)$ satisfy

$$\lim_{n\to\infty} \frac{nL_F(n)}{a_n^\alpha} = \begin{cases} \Gamma(1-\alpha)\cos\frac{\alpha\pi}{2} & \text{if } \alpha \in (0,1) \\ \frac{2}{\pi} & \text{if } \alpha = 1 \\ \frac{\Gamma(2-\alpha)}{\alpha-1}|\cos\frac{\alpha\pi}{2}| & \text{if } \alpha \in (1,2) \end{cases}$$

which implies that $a_n = n^{1/\alpha}L(n)$ for some L slowly varying at infinity and b_n can be chosen according to

$$b_n = \begin{cases} 0 & \text{if } \alpha \in (0,1) \\ a_n \int_{\mathbb{R}} \sin\frac{x}{a_n} dF(x) & \text{if } \alpha = 1 \\ \int_{\mathbb{R}} x dF(x) & \text{if } \alpha \in (1,2) \end{cases}$$

Note that the crucial parameters here are α and β. The scale-location parameters μ and σ can be incorporated into the sequences $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$ for the general case of $S_\alpha(\sigma,\beta,\mu)$ domains of attraction. The case $\alpha = 2$ is taken care of by the central limit theorem: F is in the domain of attraction of $N(0,1)$ if and only if F has finite mean and variance.

Definition 4 *An* $\mathbf{F}$*-adapted process* L *is a Lévy process if*

(1) $L_0 \overset{a.s.}{=} 0$

(2) L *has independent increments:* $\forall\ s,t \in \mathbb{R}_+,\ L_{t+s} - L_s \amalg \mathcal{F}_s$.

(3) L *has stationary increments:* $\forall\ s,t \in \mathbb{R}_+,\ L_{t+s} - L_s \overset{\mathcal{D}}{=} L_t$

(4) L *is stochastically continuous:* $\forall\ t \in \mathbb{R}_+,\ \mathbf{P}\lim_{s\to t} L_s^{\alpha,\beta} = L_t^{\alpha,\beta}$

Theorem 13 *Every Lévy process has a unique càdlàg modification which is also a Lévy process (*X *is a modification of* Y *if* $\forall t \in \mathbb{R}_+,\ \mathbf{P}\{X_t = Y_t\} = 1$*).*

Proof: See Protter (1990) We can therefore always choose to work with càdlàg modifications.

Definition 5 *An* $\mathbf{F}$*-adapted process* $L^{\alpha,\beta}$ *is a standard* α*-stable Lévy process if*

(1) $L^{\alpha,\beta}$ *is a Lévy process*

(2) $L^{\alpha,\beta}$ *has* α*-stable-stable increments:* $\forall\ t \in \mathbb{R}_+,\ L_t^{\alpha,\beta} \sim S_\alpha(t^{1/\alpha}, \beta, 0)$

When $\alpha = 2$, β is irrelevant and $\frac{1}{\sqrt{2}}L^{2,\beta}$ is a standard Wiener process.

Theorem 14 (Samorodnitsky & Taqqu) *If* $L^{\alpha,\beta}$ *is a standard* α*-stable Lévy process on* $[0,1]$ *for some* $\alpha \in (0,2)$ *and* $\beta \in [-1,1]$ *then*

$$L^{\alpha,\beta} \overset{\mathcal{D}}{=} \begin{cases} \left\{ C_\alpha^{1/\alpha} \sum_{n\in\mathbf{N}} \gamma_n \Gamma_n^{-1/\alpha} 1_{\{U_n \leqslant t\}} \right\}_{t\in[0,1]} & \text{if } \alpha \in (0,1) \\ \left\{ C_\alpha^{1/\alpha} \sum_{n\in\mathbf{N}} \left(\gamma_n \Gamma_n^{-1/\alpha} 1_{\{U_n \leqslant t\}} - \beta t b_n^{(\alpha)} \right) - \beta t \frac{2}{\pi} \ln \frac{2}{\pi} \right\}_{t\in[0,1]} & \text{if } \alpha = 1 \\ \left\{ C_\alpha^{1/\alpha} \sum_{n\in\mathbf{N}} \left(\gamma_n \Gamma_n^{-1/\alpha} 1_{\{U_n \leqslant t\}} - \beta t b_n^{(\alpha)} \right) \right\}_{t\in[0,1]} & \text{if } \alpha \in (1,2) \end{cases}$$

where $\{\gamma_n\}_{n\in\mathbf{N}}$ *are i.i.d. with* $\mathbf{P}\{\gamma_n = 1\} = \frac{1+\beta}{2}$ *and* $\mathbf{P}\{\gamma_n = -1\} = \frac{1-\beta}{2}$*,* $\{\Gamma_n\}_{n\in\mathbf{N}}$ *are the jump times of a Poisson process with unit arrival rate, and,* $\{U_n\}_{n\in\mathbf{N}}$ *are i.i.d. and uniformly distributed over* $[0,1]$*. Furthermore,* $\{\gamma_n\}_{n\in\mathbf{N}}$*,* $\{\Gamma_n\}_{n\in\mathbf{N}}$*, and* $\{U_n\}_{n\in\mathbf{N}}$ *are mutually independent. The constants* C_α *and* $\{b_n^{(\alpha)}\}_{n\in\mathbf{N}}$ *are given by*

$$C_\alpha = \begin{cases} \frac{1-\alpha}{\Gamma(2-\alpha)\cos\frac{\pi\alpha}{2}} & \text{if } \alpha \in (0,1) \cup (1,2) \\ \frac{2}{\pi} & \text{if } \alpha = 1 \end{cases}$$

$$b_n^{(\alpha)} = \begin{cases} \int_{[\frac{1}{n},\frac{1}{n-1}]} x^{-2}\sin x dx & \text{if } \alpha = 1 \\ \frac{\alpha}{\alpha-1}\left(n^{\frac{\alpha-1}{\alpha}} - (n-1)^{\frac{\alpha-1}{\alpha}}\right) & \text{if } \alpha \in (1,2) \end{cases}$$

The essential elements in this series representation are the jump times $\{U_n\}_{n\in\mathbb{N}}$, the jump directions $\{\gamma_n\}_{n\in\mathbb{N}}$ (1 for up and -1 for down), and the decreasing jump heights $\{\Gamma_n^{-1/\alpha}\}_{n\in\mathbb{N}}$, which are all independent of each other. The $b_n^{(\alpha)}$ and C_α are only constants.

3.5 The α-Stable Lévy Process Approximation

Theorem 15 (Furrer et al, 1996) *Let $\{\chi_k\}_{k\in\mathbb{N}}$ be a sequence of i.i.d. random variables having mean μ and common distribution function F in the domain of attraction of $S_\alpha(1,\beta,0)$ for some $\alpha \in (1,2)$ and $\beta \in [-1,1]$. Let $\varphi(n) = n^{1/\alpha}L(n)$ where $L(n)$ is the function slowly varying at infinity such that*

$$\frac{1}{\varphi(n)}\sum_{k=1}^{n}(\chi_k - \mu) \Rightarrow S_\alpha(1,\beta,0)$$

whose existence is given by Theorem (12). Let $\{N^{(n)}\}_{n\in\mathbb{N}} \subset D$ be a sequence of point processes such that for some constant $\lambda > 0$, $\frac{N^{(n)}-n\lambda I}{\varphi(n)} \Rightarrow 0$. Define $R_t^{(n)} = u^{(n)} + \pi^{(n)}t - \frac{1}{\varphi(n)}\sum_{k=1}^{N_t^{(n)}}\chi_k$ and $R_t^{(\infty)} = u + \pi t - \lambda^{1/\alpha}L_t^{\alpha,\beta}$ where $\pi^{(n)} - \lambda\mu\frac{n}{\varphi(n)} \to \pi$ and $u^{(n)} \to u$. Then, $R^{(n)} \Rightarrow R^{(\infty)}$.

Proof: First, write $R_t^{(n)}$ in the following form:

$$\begin{aligned} R_t^{(n)} &= u^{(n)} + \pi^{(n)}t - \frac{1}{\varphi(n)}\sum_{k=1}^{N_t^{(n)}}\chi_k \\ &= u^{(n)} + t\left(\pi^{(n)} - \frac{n\lambda\mu}{\varphi(n)}\right) - \mu\left(\frac{N_t^{(n)} - n\lambda t}{\varphi(n)}\right) - \frac{1}{\varphi(n)}\sum_{k=1}^{N_t^{(n)}}(\chi_k - \mu) \end{aligned}$$

Since

$$\frac{1}{\varphi(n)}\sum_{k=1}^{n}(\chi_k - \mu) \Rightarrow L_1^{\alpha,\beta}$$

we have that

$$\frac{1}{\varphi(\lfloor nt \rfloor)} \sum_{k=1}^{\lfloor nt \rfloor} (\chi_k - \mu) \Rightarrow L_1^{\alpha,\beta}$$

and so

$$\frac{\varphi(n)}{\varphi(\lfloor nt \rfloor)} \frac{1}{\varphi(n)} \sum_{k=1}^{\lfloor nt \rfloor} (\chi_k - \mu) \Rightarrow L_1^{\alpha,\beta}$$

but

$$\frac{\varphi(n)}{\varphi(\lfloor nt \rfloor)} = \frac{n^{1/\alpha} L(n)}{(\lfloor nt \rfloor)^{1/\alpha} L(\lfloor nt \rfloor)} = \left(\frac{n}{\lfloor nt \rfloor}\right)^{1/\alpha} \frac{L(n)}{L(\lfloor nt \rfloor)}$$

and since

$$\frac{L(n)}{L(\lfloor nt \rfloor)} \to 1 \quad and \quad \frac{\lfloor nt \rfloor}{n} \to t$$

we have that

$$L_t^{(n)} \stackrel{def}{=} \frac{1}{\varphi(n)} \sum_{k=1}^{\lfloor nt \rfloor} (\chi_k - \mu) \;\Rightarrow\; t^{1/\alpha} L_1^{\alpha,\beta} \stackrel{\mathcal{D}}{=} L_t^{\alpha,\beta}$$

Since $\frac{N^{(n)} - n\lambda I}{\varphi(n)} = \frac{\frac{N^{(n)}}{n} - \lambda I}{n^{1/\alpha - 1} L(n)} \xrightarrow{\mathbf{P}} 0$ and $n^{1/\alpha-1} L(n) \to 0$ for $\alpha > 1$ we have that $\frac{N^{(n)}}{n} - \lambda I \xrightarrow{\mathbf{P}} 0$ and so $\frac{N^{(n)}}{n} \Rightarrow \lambda I$. Applying Theorem (8) we get that

$$\begin{aligned} \frac{1}{\varphi(n)} \sum_{k=1}^{N_t^{(n)}} (\chi_k - \mu) &= \frac{1}{\varphi(n)} \sum_{k=1}^{\left\lfloor n \frac{N_t^{(n)}}{n} \right\rfloor} (\chi_k - \mu) = \left(L^{(n)} \circ \frac{N^{(n)}}{n}\right)_t \\ &\Rightarrow (L^{\alpha,\beta} \circ \lambda I)_t = L_{\lambda t}^{\alpha,\beta} \stackrel{\mathcal{D}}{=} \lambda^{1/\alpha} L_t^{\alpha,\beta} \end{aligned}$$

by the self-similarity of stable processes and so the result is proved.

□

The generality gained by this result is that the second moment of the claims distribution need not exist and the claim arrivals need not form a renewal process. However, α is restricted to $(1,2)$. The case $\alpha = 2$ is handled by the Wiener diffusion approximation. The range $\alpha \in (0,1]$ has been excluded. Fortunately, many of the applicable heavy tailed distributions, such as the Pareto or LogGamma distributions, are in a domain of attraction for some $\alpha \in (1,2)$. Also, for $\alpha \in (1,2)$ we have finite means whereas for $\alpha \in (0,1]$ we do not.

Theorem 16 (Furrer et al. 1996) *If $\{R^{(n)}\}_{n\in\mathbb{N}}$ and $R^{(\infty)}$ are the risk processes defined in Theorem (15) for which $R^{(n)} \Rightarrow R^{(\infty)}$ then $T^r(R^{(n)}) \Rightarrow T^r(R^{(\infty)})$ where $T^r(x) = \inf\{t > 0 : x(t) < 0\}$ is the ruin functional on D.*

In order to prove this theorem we first need three lemmas.

Lemma 1 (Furrer et al. 1996) *Let $X_t = \pi t - \lambda^{\frac{1}{\alpha}} L_t^{\alpha,\beta}$ for $t \in \mathbb{R}_+$. Then, $\forall \epsilon > 0$, $\mathbf{P}$-a.e. trajectory of X crosses 0 for infinitely many times in $[0, \epsilon]$.*

Proof: Let $\{t_n\}_{n\in\mathbb{N}} \subset \mathbb{R}_+$ be a sequence of times such that $t_n \to 0$. We will show that for infinitely many n, $X_{t_n} < 0$ and $X_{t_n} > 0$ $\mathbf{P}$-a.s.

$$\begin{aligned}
\mathbf{P}\{\underline{\lim}_n\{X_{t_n} \geqslant 0\}\} &\leqslant \underline{\lim}_n \mathbf{P}\{X_{t_n} \geqslant 0\} \\
&= \underline{\lim}_n \mathbf{P}\{\lambda^{\frac{1}{\alpha}} L_{t_n}^{\alpha,\beta} \leqslant \pi t_n\} \\
&= \underline{\lim}_n \mathbf{P}\{(\lambda t_n)^{\frac{1}{\alpha}} L_1^{\alpha,\beta} \leqslant \pi t_n\} \\
&= \underline{\lim}_n \mathbf{P}\{L_1^{\alpha,\beta} \leqslant \pi \lambda^{-\frac{1}{\alpha}} t_n^{1-\frac{1}{\alpha}}\} \\
&\leqslant \underline{\lim}_n \mathbf{P}\{L_1^{\alpha,\beta} \leqslant \pi \lambda^{-\frac{1}{\alpha}}\} \\
&= \mathbf{P}\{L_1^{\alpha,\beta} \leqslant \pi \lambda^{-\frac{1}{\alpha}}\}
\end{aligned}$$

since $\alpha > 1$ and $t_n \leqslant 1$ implies that $t_n^{1-\frac{1}{\alpha}} \leqslant 1$. The support of the density of $L_1^{\alpha,\beta}$ for $\alpha > 1$ is all of $\mathbb{R}$ so it follows that $\mathbf{P}\{L_1^{\alpha,\beta} \leqslant \pi\lambda^{-\frac{1}{\alpha}}\} = 1 - p < 1$ for some $p > 0$. Thus, $\mathbf{P}\{\underline{\lim}_n\{X_{t_n} \geqslant 0\}\} < 1$. Also, by right continuity,

$$\begin{aligned}
X_{t_n} &= X_{t_n} - X_0 = X_{t_n} - \lim_{N\to\infty} X_{t_{n+N+1}} = \lim_{N\to\infty}[X_{t_n} - X_{t_{n+N+1}}] \\
&= \lim_{N\to\infty} \sum_{k=0}^{N} [X_{t_{n+k}} - X_{t_{n+k+1}}] = \sum_{k=0}^{\infty} [X_{t_{n+k}} - X_{t_{n+k+1}}]
\end{aligned}$$

Let $\mathcal{G}_n = \sigma(X_{t_{n+k}} - X_{t_{n+k+1}} : k \in \mathbb{N}_0)$. Thus, $X_{t_n} \in \mathcal{G}_n$ for all n and therefore $\underline{\lim}_n\{X_{t_n} \geqslant 0\} \in \bigcap_{n\in\mathbb{N}} \mathcal{G}_n$. From the Kolmogorov $0-1$ law, $\mathbf{P}\{\underline{\lim}_n\{X_{t_n} \geqslant 0\}\} = 0$ or 1 but we have that $\mathbf{P}\{\underline{\lim}_n\{X_{t_n} \geqslant 0\}\} < 1$ and so $\mathbf{P}\{\underline{\lim}_n\{X_{t_n} \geqslant 0\}\} = 0$. Thus,

$\mathbf{P}\{\overline{\lim}_n\{X_{t_n} < 0\}\} = 1$, and so $X_{t_n} < 0$ for infinitely many n $\mathbf{P}$-a.s. So, for any $\epsilon > 0$, $X < 0$ $\mathbf{P}$-a.s. for infinitely many time points in $[0, \epsilon]$. By a similar argument, $X > 0$ $\mathbf{P}$-a.s. for infinitely many times in $[0, \epsilon]$ and therefore, $\mathbf{P}$-a.s., X crosses 0 infinitely often in $[0, \epsilon]$.

$\square$

Lemma 2 (Furrer et al. 1996) *Let* $T_0 = \inf\{t > 0 : R^{(\infty)} = 0\}$*. Then, for all* $\epsilon > 0$*,* $\mathbf{P}$*-a.e. trajectory of* $R^{(\infty)}$ *crosses* 0 *for infinitely many times in the stochastic interval* $[T_0, T_0 + \epsilon]$.

Proof: Define a process $X_t = R^{(\infty)}_{T_0+t} - R^{(\infty)}_{T_0}$. By the strong Markov property of Lévy processes (Protter 1990), X is again an α-stable Lévy process starting from $X_0 = 0$ $\mathbf{P}$-a.s. Furthermore,

$$\begin{aligned} X_t &= u + \pi(T_0 + t) - \lambda^{1/\alpha} L^{\alpha,\beta}_{T_0+t} - \left[u + \pi T_0 - \lambda^{1/\alpha} L^{\alpha,\beta}_{T_0}\right] \\ &= \pi t - \left[L^{\alpha,\beta}_{T_0+t} - L^{\alpha,\beta}_{T_0}\right] \stackrel{\mathcal{D}}{=} \pi t - \lambda^{1/\alpha} L^{\alpha,\beta}_t \end{aligned}$$

Thus, from Lemma 1 we have that for $\mathbf{P}$-a.e. $\omega \in \Omega$, $X(\omega)$ crosses zero infinitely often in every right neighborhood of zero. Let $\omega \in \Omega$ be such that $X(\omega)$ crosses zero infinitely often in every right neighborhood of zero. Since $X(\omega)$ is simply $R^{(\infty)}(\omega)$ started at $T_0(\omega)$ we have that $R^{(\infty)}(\omega)$ crosses zero infinitely often in every right neighborhood of $T_0(\omega)$ and so the result follows. $\square$

Lemma 3 (Furrer et al. 1996) *Let* $\{\tau_n\}_{n\in\mathbb{N}}$ *be the jump times of the process* $R^{(\infty)}$*. Then*

$$\mathbf{P}\left\{\bigcup_{n\in\mathbb{N}}\{R^{(\infty)}_{\tau_n-} \neq 0\}\right\} = 1$$

Proof: First, since

$$\bigcup_{n\in\mathbb{N}}\{R^{(\infty)}_{\tau_n-} \neq 0\} = \bigcup_{n\in\mathbb{N}}\bigcup_{k\in\mathbb{N}}\{R^{(\infty)}_{\tau_n-} \neq 0, \tau_n \leqslant k\}$$

we need only show that $\mathbf{P}\{R^{(\infty)}_{\tau_n-} \neq 0, \tau_n \leqslant k\} = 1$ for any n and k in $\mathbb{N}$. However, Since $\{R^{(\infty)}_{\tau_n-} \neq 0, \tau_n \leqslant 1\} \subset \{R^{(\infty)}_{\tau_n-} \neq 0, \tau_n \leqslant k\}$ for all $k, n \in \mathbb{N}$ it suffices to show that $\mathbf{P}\{R^{(\infty)}_{\tau_n-} \neq 0, \tau_n \leqslant 1\} = 1$ and so we can consider $R^{(\infty)}_t$ with $t, \{\tau_n\}_{n\in\mathbf{N}} \subset [0,1]$, in which case $\mathbf{P}\{\tau_n \leqslant 1\} = 1$ and the jump times are uniformly distributed on $[0,1]$.

First, note that $\mathbf{P}\{R^{(\infty)}_{\tau_j-} \neq 0\} = \mathbf{P}\{R^{(\infty)}_{\tau_j} - \Delta R^{(\infty)}_{\tau_j} \neq 0\}$. Now, using the series representation for α-stable Lévy processes on $[0,1]$,

$$
\begin{aligned}
& \Delta R^{(\infty)}_{\tau_j} = R^{(\infty)}_{\tau_j} - R^{(\infty)}_{\tau_j-} \\
= & \left\{u + \pi(\tau_j) - \lambda^{1/\alpha} L^{\alpha,\beta}_{\tau_j}\right\} - \left\{u + \pi(\tau_j-) - \lambda^{1/\alpha} L^{\alpha,\beta}_{\tau_j-}\right\} \\
= & \lambda^{1/\alpha} L^{\alpha,\beta}_{\tau_j} - \lambda^{1/\alpha} L^{\alpha,\beta}_{\tau_j-} \\
= & \lambda^{1/\alpha} C_\alpha^{1/\alpha} \sum_{n\in\mathbf{N}} \left\{\gamma_n \Gamma_n^{-1/\alpha} 1_{\{\tau_n \leqslant \tau_j\}} - \beta \tau_j b_n^{(\alpha)}\right\} \\
& -\lambda^{1/\alpha} C_\alpha^{1/\alpha} \sum_{n\in\mathbf{N}} \left\{\gamma_n \Gamma_n^{-1/\alpha} 1_{\{\tau_n \leqslant (\tau_j-)\}} - \beta (\tau_j-) b_n^{(\alpha)}\right\} \\
= & \lambda^{1/\alpha} C_\alpha^{1/\alpha} \gamma_j \Gamma_j^{-1/\alpha}
\end{aligned}
$$

and so

$$
\begin{aligned}
& \mathbf{P}\left\{R^{(\infty)}_{\tau_j-} \neq 0\right\} \\
= & \mathbf{P}\left\{R^{(\infty)}_{\tau_j} - \Delta R^{(\infty)}_{\tau_j} \neq 0\right\} \\
= & \mathbf{P}\left\{u + \pi\tau_j + (\lambda C_\alpha)^{1/\alpha}\left(\beta\tau_j b_j^{(\alpha)} - \sum_{n\in\mathbb{N}\setminus\{j\}} \left\{\gamma_n \Gamma_n^{-1/\alpha} 1_{\{\tau_n \leqslant \tau_j\}} - \beta\tau_j b_n^{(\alpha)}\right\}\right) \neq 0\right\}
\end{aligned}
$$

Now,

$$
\begin{aligned}
& \mathbf{P}\left\{u + \pi t + (\lambda C_\alpha)^{1/\alpha}\left(\beta t b_j^{(\alpha)} - \sum_{n\in\mathbb{N}\setminus\{j\}} \left\{\gamma_n \Gamma_n^{-1/\alpha} 1_{\{\tau_n \leqslant t\}} - \beta t b_n^{(\alpha)}\right\}\right) \neq 0\right\} \\
= & \mathbf{P}\left\{u + \pi t - \lambda^{1/\alpha} L^{\alpha,\beta}_t \neq 0 \;\middle|\; t < \tau_j\right\} \\
= & 1
\end{aligned}
$$

since it is known that stable densities are continuous (Zolotarev 1986). Using the law of total probability, the independence of the random variables in the series representation, and the fact that the jump times are uniformly distributed over [0, 1] we have that

$$
\begin{aligned}
&\mathbf{P}\left\{R^{(\infty)}_{\tau_j-} \neq 0, \tau_n \leqslant 1\right\} \\
= \;& \mathbf{P}\left\{R^{(\infty)}_{\tau_j-} \neq 0\right\} \\
= \;& \mathbf{P}\left\{u + \pi\tau_j + (\lambda C_\alpha)^{\frac{1}{\alpha}}\left(\sum_{n\in\mathbb{N}\setminus\{j\}}\left\{\gamma_n\Gamma_n^{\frac{-1}{\alpha}}1_{\{\tau_n\leqslant\tau_j\}} - \beta\tau_j b_n^{(\alpha)}\right\} - \beta\tau_j b_j^{(\alpha)}-\right) \neq 0\right\} \\
= \;& \int_0^1 \mathbf{P}\left\{u + \pi\tau_j + (\lambda C_\alpha)^{\frac{1}{\alpha}}\left(\sum_{n\in\mathbb{N}\setminus\{j\}}\left\{\gamma_n\Gamma_n^{\frac{-1}{\alpha}}1_{\{\tau_n\leqslant\tau_j\}} - \beta\tau_j b_n^{(\alpha)}\right\} - \beta\tau_j b_j^{(\alpha)}\right) \neq 0 \,\middle|\, \tau_j = t\right\} dt \\
= \;& \int_0^1 \mathbf{P}\left\{u + \pi t + (\lambda C_\alpha)^{\frac{1}{\alpha}}\left(\sum_{n\in\mathbb{N}\setminus\{j\}}\left\{\gamma_n\Gamma_n^{\frac{-1}{\alpha}}1_{\{\tau_n\leqslant t\}} - \beta t b_n^{(\alpha)}\right\} - \beta t b_j^{(\alpha)}\right) \neq 0\right\} dt \\
= \;& \int_0^1 1 dt = 1
\end{aligned}
$$

□

Proof of Theorem 16:. If we can show $T^r(x) = \inf\{t > 0 : x(t) < 0\}$ is continuous for $\mathbf{P}R^{(\infty)^{-1}}$-a.e. trajectory in D then we can apply Corollary (1) to conclude that since $R^{(n)} \Rightarrow R^{(\infty)}$ we have that $T^r(R^{(n)}) \Rightarrow T^r(R^{(\infty)})$, proving the result.

For $R^{(\infty)}_t = u + \pi t - \lambda^{1/\alpha}L^{\alpha,\beta}_t$, let $A \subset \Omega$ be the set such that for $\omega \in A$, $R^{(\infty)}(\omega)$ crosses zero infinitely often in every right neighborhood of $S(\omega) = \inf\{t > 0 : R^{(\infty)}_t(\omega) = 0\}$. By Lemma (2), $\mathbf{P}(A) = 1$. Let $B = \cap_{n\in\mathbb{N}}\{R^{(\infty)}_{\tau_n-} \neq 0\}$ where $\{\tau_n\}_{n\in\mathbb{N}}$ are the sequence of jump times of $R^{(\infty)}$. By Lemma (3), $\mathbf{P}(B) = 1$. Thus, $\mathbf{P}(A \cap B) = 1$.

Now, let $\omega \in A \cap B$ and set $x(t) = R^{(\infty)}_t(\omega)$. Let $\{x_n\}_{n\in\mathbb{N}} \subset D$ be any sequence such that $x_n \to x$ in (D, d). Assume that $T^r(x_n)$ does not converge to $T^r(x)$. Thus, either $\underline{\lim}_n T^r(x_n) \neq T^r(x)$ or $\overline{\lim}_n T^r(x_n) \neq T^r(x)$. In either case we can find a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ such that $T^r(x_{n_k}) \to T \neq T^r(x)$. There are two cases.

In the first case, $0 \leqslant T^r(x) < T \leqslant \infty$. Now, x has countably many jumps so we can find an abitrarily small $\delta_1 > 0$ such that x is continuous at $T^r(x) + \delta_1$, ie, $\Delta x(T^r(x) + \delta_1) = 0$ where $\Delta x(t) = x(t) - x(t-)$. Since $T^r(x) < T$ we can find an arbitrarily small $\delta_2 > 0$ such that $T^r(x) + \delta_2 < T$. Finally, $T^r(x)$ is defined as an infimum so we can find an arbitrarily small $\delta_3 > 0$ such that $x(T^r(x) + \delta_3) < -\epsilon < 0$ for some $\epsilon > 0$. By the right continuity of x we can find a single $\delta > 0$ such that for some $\epsilon > 0$

$$\Delta x(T^r(x) + \delta) = 0, \qquad T^r(x) + \delta < T, \qquad x(T^r(x) + \delta) < -\epsilon < 0$$

Since x is continuous at $T^r(x) + \delta$, $x_{n_k}(T^r(x) + \delta) \to x(T^r(x) + \delta)$ and so $\exists K_1$ such that $k > K_1 \Rightarrow x_{n_k}(T^r(x) + \delta) < \epsilon + x(T^r(x) + \delta) < 0$ and so for $k > K_1$, $T^r(x_{n_k}) \leqslant T^r(x) + \delta < T$, or, in other words, $k > K_1 \Rightarrow |T - T^r(x_{n_k})| > \delta > 0$ which contradicts that $T^r(x_{n_k}) \to T$.

In the second case, $0 \leqslant T < T^r(x) \leqslant \infty$. Since $x_{n_k} \to x$ and $T^r(x_{n_k}) \to T$ we have from Theorem (1) that $x_{n_k}(T^r(x_{n_k})) \to x(T)$ or $x(T-)$. If the limit is $x(T)$ then since $T < T^r(x)$, $x(T) \geqslant 0$. But, $x_{n_k}(T^r(x_{n_k})) \leqslant 0$ for all $k \in \mathbb{N}$ so $x(T) \leqslant 0$. Thus, $x(T) = 0$. Similarly, if $x(T-)$ is the limit then $x(T-) = 0$. Suppose that $x(T) = 0$ is the limit. Since $S(\omega) = \inf\{t > 0 : R_t^{(\infty)}(\omega) = 0\} = \inf\{t > 0 : x(t) = 0\}$ we have that $S(\omega) \leqslant T$. By Lemma (2), x crosses zero infinitely often in every right neighborhood of $S(\omega)$. Choose $\rho > 0$ such that $T + \rho < T^r(x)$. Thus, $\exists t^* \in [S(\omega), T + \rho]$ such that $x(t^*) < 0$ but since $t^* < T^r(x)$ we have a contradiction. Now suppose $x(T-) = 0$ is the limit. Since $\omega \in B$, $x(\tau_n(\omega)-) \neq 0$ for all jump times τ_n and so T is not a jump time of x. Thus, $x(T) = x(T-) = 0$ and as in the case $x(T) = 0$ we obtain a contradiction. Thus, $T^r(x_n)$ must converge to $T^r(x)$ and so T^r is continuous at $R^{(\infty)}$ for any $\omega \in A \cap B$, proving the result.

$\square$

Unfortunately, the lack of path-wise continuity and explicit forms for stable densities makes it difficult to obtain exact expressions for the ruin probabilities. However, an upper bound has been obtained which is easily computed numerically:

Proposition 7 (Furrer et al. 1996) *For approximations with standard α-stable Lévy processes $L^{\alpha,\beta}$ for $\alpha \in (1,2)$, $\beta \in [-1,1]$ and $u, \tilde{\pi}, \lambda > 0$,*

$$\mathbf{P}\{T^r(u + \tilde{\pi}t - \lambda^{1/\alpha} L_t^{\alpha,\beta}) < T\} \leqslant \frac{1}{\rho}\mathbf{P}\{L_T^{\alpha,\beta} > u\lambda^{-1/\alpha}\}$$

where $\rho = \frac{1}{2} + \frac{1}{\pi\alpha}\arctan(\beta\tan(\frac{\pi\alpha}{2}))$.

Now, we apply these results to reserve models with a large class of renewal claim arrivals. Note that a small error in the second case of the proof below has been corrected.

Proposition 8 (Furrer et al. 1996) *Let N^χ be a renewal counting process with inter-claim times $\{\tau_k\}_{k\in\mathbf{N}}$. If there is a $\lambda > 0$ and a function L slowly varying at infinity such that $\frac{1}{\varphi(n)}\sum_{k=1}^{\lfloor nt\rfloor}(\tau_k - \lambda^{-1}) \Rightarrow W_t$ where $\varphi(n) = n^{1/2}L(n)$ then for $\alpha \in (1,2)$ we have that*

$$\frac{N^\chi_{nt} - \lambda nt}{n^{1/\alpha}} \xrightarrow{\mathbf{P}} 0$$

Proof: Since $\sup_{s\in[0,t]} n^{-1/\alpha}|N^\chi_{ns} - \lambda ns| \geqslant n^{-1/\alpha}|N^\chi_{nt} - \lambda nt|$ it sufficient to show that $\sup_{s\in[0,t]} n^{-1/\alpha}|N^\chi_{ns} - \lambda ns| \xrightarrow{\mathbf{P}} 0$. Now, $\sup_{s\in[0,t]} n^{-1/\alpha}|N^\chi_{ns} - \lambda ns| > \epsilon$ if and only if $\exists s^* \in [0,t]$ such that $|N^\chi_{ns^*} - \lambda ns^*| > \epsilon n^{1/\alpha}$ which occurs if and only if $N^\chi_{ns^*} > \epsilon n^{1/\alpha} + \lambda ns^*$ or $N^\chi_{ns^*} < -\epsilon n^{1/\alpha} + \lambda ns^*$. Case 1: $N^\chi_{ns^*} > \epsilon n^{1/\alpha} + \lambda ns^*$.
Since

$$ns^* \geqslant \sum_{k=1}^{N^\chi_{ns^*}} \tau_k > \sum_{k=1}^{\lfloor \epsilon n^{1/\alpha} + \lambda ns^*\rfloor} \tau_k$$

we have, setting $nu_1 = \epsilon n^{1/\alpha} + \lambda ns^*$,

$$ns^* > \sum_{k=1}^{\lfloor nu_1\rfloor} \tau_k$$

$$\iff \quad \frac{nu_1 - \lfloor nu_1\rfloor}{\lambda n^{1/\alpha}} - \frac{\epsilon}{\lambda} > \frac{1}{n^{1/\alpha}}\sum_{k=1}^{\lfloor nu_1\rfloor}(\tau_k - \lambda^{-1})$$

which follows from

$$-\frac{\epsilon}{\lambda} > \frac{1}{n^{1/\alpha}}\sum_{k=1}^{\lfloor nu_1\rfloor}(\tau_k - \lambda^{-1})$$

Case 2: $N^\chi_{ns^*} < -\epsilon n^{1/\alpha} + \lambda ns^*$.
Since

$$N^\chi_{ns^*} \leqslant \lfloor -\epsilon n^{1/\alpha} + \lambda ns^*\rfloor < \lfloor -\epsilon n^{1/\alpha} + \lambda ns^*\rfloor + 1$$

we have, setting $nu_2 = -\epsilon n^{1/\alpha} + \lambda n s^*$,

$$
\begin{aligned}
& ns^* < \sum_{k=1}^{\lfloor nu_2 \rfloor + 1} \tau_k \\
\iff \quad & \frac{nu_2 - \lfloor nu_2 \rfloor - 1}{\lambda n^{1/\alpha}} + \frac{\epsilon}{\lambda} < \frac{1}{n^{1/\alpha}} \sum_{k=1}^{\lfloor nu_2 + 1 \rfloor} (\tau_k - \lambda^{-1})
\end{aligned}
$$

which follows from

$$
\frac{\epsilon}{\lambda} < \frac{1}{n^{1/\alpha}} \sum_{k=1}^{\lfloor n(u_2 + 1/n) \rfloor} (\tau_k - \lambda^{-1})
$$

Since we have that the two variables $u_1 \in U_1(n) \stackrel{def}{=} [\epsilon n^{1/\alpha - 1}, \epsilon n^{1/\alpha - 1} + \lambda t]$ and $\tilde{u}_2 = u_2 + 1/n \in U_2(n) \stackrel{def}{=} [1/n - \epsilon n^{1/\alpha - 1}, 1/n - \epsilon n^{1/\alpha - 1} + \lambda t]$,

$$
\sup_{u_1 \in U_1(n)} \left| n^{-1/\alpha} \sum_{k=1}^{\lfloor nu_1 \rfloor} (\tau_k - \lambda^{-1}) \right| \xrightarrow{\mathbf{P}} 0 \;\Rightarrow\; \left| n^{-1/\alpha} \sum_{k=1}^{\lfloor nu_1 \rfloor} (\tau_k - \lambda^{-1}) \right| \xrightarrow{\mathbf{P}} 0
$$

and

$$
\sup_{\tilde{u}_2 \in U_2(n)} \left| n^{-1/\alpha} \sum_{k=1}^{\lfloor n\tilde{u}_2 \rfloor} (\tau_k - \lambda^{-1}) \right| \xrightarrow{\mathbf{P}} 0 \;\Rightarrow\; \left| n^{-1/\alpha} \sum_{k=1}^{\lfloor n\tilde{u}_2 \rfloor} (\tau_k - \lambda^{-1}) \right| \xrightarrow{\mathbf{P}} 0
$$

Thus, if either of these supremums converges to zero in probability then it follows that $\sup_{s \in [0,t]} n^{-1/\alpha} |N^{\chi}_{ns} - \lambda n s| \xrightarrow{\mathbf{P}} 0$, proving the result.

From the assumption that

$$
\frac{1}{n^{1/2} L(n)} \sum_{k=1}^{\lfloor nt \rfloor} (\tau_k - \lambda^{-1}) \Rightarrow W_t
$$

we have that

$$
\frac{1}{n^{1/2 - 1/\alpha} L(n)} \frac{1}{n^{1/\alpha}} \sum_{k=1}^{\lfloor nt \rfloor} (\tau_k - \lambda^{-1}) \Rightarrow W_t
$$

But, for $\alpha < 2$, $1/2 - 1/\alpha < 0$ so $n^{1/2 - 1/\alpha} L(n) \to 0$, and so for each t we must have that

$$
\frac{1}{n^{1/\alpha}} \sum_{k=1}^{\lfloor nt \rfloor} (\tau_k - \lambda^{-1}) \Rightarrow 0
$$

which is equivalent to

$$\frac{1}{n^{1/\alpha}}\sum_{k=1}^{\lfloor nt\rfloor}(\tau_k-\lambda^{-1}) \xrightarrow{\mathbf{P}} 0$$

Finally, for any bounded interval $[a,b]$, this implies that

$$\sup_{t\in[a,b]}\left|\frac{1}{n^{1/\alpha}}\sum_{k=1}^{\lfloor nt\rfloor}(\tau_k-\lambda^{-1})\right| \xrightarrow{\mathbf{P}} 0$$

Since we know the intervals in the supremums $U_1(n)=[\epsilon n^{1/\alpha-1},\epsilon n^{1/\alpha-1}+\lambda t]$ and $U_2(n)=[1/n-\epsilon n^{1/\alpha-1},1/n-\epsilon n^{1/\alpha-1}+\lambda t]$ are bounded in n if $\alpha\in(1,2)$, the supremums do converge to zero in probability in either case and so the result follows.

□

Now, given a risk process $R_t=u+\pi t-\sum_{k=1}^{N_t^\chi}\chi_k$ where N^χ is a Poisson process with rate λ, $\{\chi_k\}_{k\in\mathbb{N}}$ are i.i.d. with common distribution function F, mean μ, and F is in the domain of attraction of $L_1^{\alpha,\beta}$ for some $\alpha\in(1,2)$ and $\beta\in[-1,1]$, Furrer suggests the following "weak approximation".

Since N^χ is Poisson process, $\frac{N_{nt}^\chi-\lambda nt}{\sqrt{n}}\Rightarrow\gamma\lambda^{3/2}W_t$ where γ is the variance of the inter-claim times and W is a standard Wiener process (Billingsley 1968). Thus, for $\alpha\in(1,2)$,

$$\frac{N_{nt}^\chi-\lambda nt}{n^{1/\alpha}}=\frac{\sqrt{n}}{n^{1/\alpha}}\frac{N_{nt}^\chi-\lambda nt}{\sqrt{n}}\Rightarrow 0$$

since $\frac{\sqrt{n}}{n^{1/\alpha}}\to 0$. Now, for each fixed $n\in\mathbb{N}$ we have

$$\begin{aligned}
\Psi(u,T) &= \mathbf{P}\{T^r(R)\leqslant T\}\\
&= \mathbf{P}\left\{\inf_{t\in[0,T]}\left(u+\pi t-\sum_{k=1}^{N_t^\chi}\chi_k\right)<0\right\}\\
&= \mathbf{P}\left\{\inf_{t\in[0,T]}\left(\frac{u}{\varphi(n)}+\frac{\pi}{\varphi(n)}t-\frac{1}{\varphi(n)}\sum_{k=1}^{N_t^\chi}\chi_k\right)<0\right\}\\
&= \mathbf{P}\left\{\inf_{t\in[0,\frac{T}{n}]}\left(\frac{u}{\varphi(n)}+\frac{\pi}{\varphi(n)}nt-\frac{1}{\varphi(n)}\sum_{k=1}^{N_{nt}^\chi}\chi_k\right)<0\right\}
\end{aligned}$$

Considering this last process

$$
\begin{aligned}
&\frac{u}{\varphi(n)}+\frac{\pi n t}{\varphi(n)}-\frac{1}{\varphi(n)}\sum_{k=1}^{N_{nt}^{\chi}}\chi_k \\
=\ &\frac{u}{\varphi(n)}+(\pi-\mu\lambda)\frac{nt}{\varphi(n)}-\frac{1}{\varphi(n)}\sum_{k=1}^{N_{nt}^{\chi}}(\chi_k-\mu)-\mu\left(\frac{N_{nt}^{\chi}-n\lambda t}{\varphi(n)}\right)
\end{aligned}
$$

we see that for n large we have that

$$\frac{1}{\varphi(n)}\sum_{k=1}^{N_{nt}^{\chi}}(\chi_k-\mu)\approx\lambda^{1/\alpha}L_t^{\alpha,\beta}$$

and

$$\frac{N_{nt}^{\chi}-n\lambda t}{\varphi(n)}\approx 0$$

and so this process is "close" to

$$\frac{u}{\varphi(n)}+(\pi-\mu\lambda)\frac{n}{\varphi(n)}t-\lambda^{1/\alpha}L_t^{\alpha,\beta}$$

Using $\pi=(\theta+1)\mu\lambda$ we have that the process is close to

$$\frac{u}{\varphi(n)}+\theta\mu\lambda\frac{n}{\varphi(n)}t-\lambda^{1/\alpha}L_t^{\alpha,\beta}$$

which then leads to an approximation for the ruin probability:

$$
\begin{aligned}
\Psi(u,T) &\approx \mathbf{P}\left\{\inf_{t\in[0,\frac{T}{n}]}\left(\frac{u}{\varphi(n)}+\theta\mu\lambda\frac{n}{\varphi(n)}t-\lambda^{1/\alpha}L_t^{\alpha,\beta}\right)<0\right\} \\
&= \mathbf{P}\left\{T^r\left(\left[\frac{u}{\varphi(n)}+\theta\mu\lambda\frac{n}{\varphi(n)}t-\lambda^{1/\alpha}L_t^{\alpha,\beta}\right]\right)<\frac{T}{n}\right\} \\
&\leqslant \frac{1}{\rho}\mathbf{P}\left\{L_{\frac{T}{n}}^{\alpha,\beta}>\frac{u\lambda^{-1/\alpha}}{\varphi(n)}\right\}
\end{aligned}
$$

where $\rho=\frac{1}{2}+\frac{1}{\pi\alpha}\arctan\left(\beta\tan\left(\frac{\pi\alpha}{2}\right)\right)$ and which can be evaluated by a feasible numerical procedure (see Proposition (7)).

Chapter 4

Conclusions & Further Research

4.1 An Assessement of Weak Approximations

The Wiener process approximation is very tractable but its poor performance in the case of highly skewed claims distributions prompted the generalization to α-stable Lévy process approximations, somewhat less tractable but still feasible. Thus, the work in Chapter 3 points to two issues absolutely crucial in any application of weak approximations: (i) accuracy, or fit, and (ii) tractability. In this chapter we focus on fit, not a numerical evaluation of weak approximations, easily found in the Risk Theory literature (Asmussen (1984), Furrer et al. (1996)), but a theoretically based assessment/interpretation which appears to be conspicuously absent in from the literature. If an understanding of how weak limits fit the original process can be obtained, one could then make informed choices in sacrificing fit for tractability. Ideally, one hopes to be able to select from a very tractable class of models one that fits the original process very well, avoiding the adopting of unrealistic simplifying assumptions.

Let $R_t = u + \pi t - \sum_{k=1}^{N_t^\chi} \chi_k$ be an ordinary renewal reserve model with finite mean inter-claim time $\lambda^{-1} > 0$ and i.i.d. claims $\{\chi_k\}_{k \in \mathbf{N}}$ having finite mean and variance $\mu = \mathbf{E}[\chi_k]$ and $\sigma^2 = \mathbf{V}[\chi_k]$. Let $M(t) = \mathbf{E}[N_t^\chi]$ and $V(t) = \mathbf{V}[N_t^\chi]$ be the renewal and variance functions, respectively, for the process N^{χ}. We wish to assess the fit of the weak limit approximation $R^{(\infty)}$ arising from Iglehart's construction of $\{R^{(n)}\}_{n\in\mathbf{N}}$ as well as provide a satisfying interpretation of the procedure. Iglehart (1969) remarks only that the convergence $R^{(n)} \Rightarrow R^{(\infty)}$ enables one to approximate the distributions and functionals of $R^{(n)}$ for n large by those of $R^{(\infty)}$, the so-called heavy traffic approximation, but there is no mention of how well $R^{(\infty)}$ fits R. In order to use the heavy traffic approximation, R must be "close" to $R^{(n)}$ for some large n so that R is then "close" to $R^{(\infty)}$ via the proximity of $R^{(n)}$ to $R^{(\infty)}$. However, there is a great deal of freedom in choosing $u^{(n)}, \pi^{(n)}, \mu^{(n)}$, and ${\sigma^{(n)}}^2$, and hence even more freedom in choosing $\{\chi_k^{(n)}\}_{k,n\in\mathbf{N}}$ since two moments do not uniquely specify a distribution. It is therefore possible to construct a sequence $\{R^{(n)}\}_{n\in\mathbf{N}}$ that is never "near" R or that "diverges" significantly from R, rendering $R^{(\infty)}$ a poor approximation of R. To ensure that $\{R^{(n)}\}_{n\in\mathbf{N}}$ does not "miss" R we insist that $R^{(1)} = R$ and to avoid "divergence" from R we match moments. In addition, we show that following this procedure leads to intuitively appealing choices for $u^{(n)}, \pi^{(n)}$, and $\{\chi_k^{(n)}\}_{k,n\in\mathbf{N}}$, clarifies the role of the normalization factor $\frac{1}{\sqrt{n}}$, and indicates why the common choice of $R_t^{(\infty)} = u + (\pi - \mu\lambda)t - \sigma\sqrt{\lambda}W_t$ to approximate R is an inappropriate application of Iglehart's result (see Grandell (1977,1991) or Asmussen (1984) for example).

First, define a sequence of reserve processes by $R_t^{(n)} = u^{(n)} + \pi^{(n)}t - \sum_{k=1}^{N_t^{(n)}} \chi_k^{(n)}$, for $t \in \mathbb{R}_+, n \in \mathbb{N}$, and note that this differs from Iglehart's prescription in that the normalization factor $\frac{1}{\sqrt{n}}$ has been absorbed into $u^{(n)}, \pi^{(n)}$, and $\{\chi_k^{(n)}\}_{k\in\mathbf{N}}$, and, in addition, $\pi^{(n)}$ has further absorbed the time compression factor n. Setting $u^{(1)} = u, \pi^{(1)} = \pi$, $\chi_k^{(1)} = \chi_k$ for $k \in \mathbb{N}$, and $N_t^{(1)} = N_t^{\chi}$ ensures that $R^{(1)} = R$. Now, we choose $u^{(n)}, \pi^{(n)}$, and $\{\chi_k^{(n)}\}_{k\in\mathbf{N}}$ for $n \geqslant 2$ according to a satisfying approximation principle that keeps $\{R^{(n)}\}_{n\in\mathbf{N}}$ "close" to $R^{(1)} = R$. The procedure we use is common in the physical sciences: approximate the behaviour of a finite number of particles, each contributing a finite amount of some property to the ensemble's macroscopic behaviour, by an infinite number of particles, each contributing infinitesimally. Such an approximation is constructed so that all relevant macroscopic behaviours are held constant as the number of particles is allowed to increase to infinity and their contributions become infinitesimal. For reserve models, we think of the number of claims per unit time as the number of particles, the claim sizes as the amount of the property contributed to the ensemble, and the aggregate claims made per unit time (or the aggregate net income per unit time) as the macroscopic property of interest. To construct a sequence $\{R^{(n)}\}_{n\in\mathbf{N}}$ whose macroscopic behaviour closely resembles R, we incorporate the probabilistic structure of R into the approximating sequence by insisting that for each $n \in \mathbb{N}$, $N^{(n)}$ is again an ordinary renewal counting process independent of the i.i.d. claims $\{\chi_k^{(n)}\}_{k\in\mathbf{N}}$. Thus, for each $n \in \mathbb{N}$, $R^{(n)}$ is an ordinary renewal model. We want the number of claims per unit time to increase in n at some reference rate, say $O(n)$. A natural way to achieve this is to compress the time scale by a factor of $\frac{1}{n}$ by setting $N_t^{(n)} = N_{nt}^{\chi}$, as Iglehart and Furrer et. al. have done. Since $\frac{M(t)}{t} \to \lambda$ and $\frac{M(nt)}{nt} \to \lambda$ as $t \to \infty$ we have that $M(nt) \approx nM(t)$ for large t, giving the desired $O(n)$ increase in average arrival rate. In order to keep the laws of $\{R^{(n)}\}_{n\in\mathbf{N}}$ "constant" and "close" to that of $R^{(1)} = R$ we match the first two moments for all $t \in \mathbb{R}_+$ and $n \in \mathbb{N}$. Let $\mu^{(n)} = \mathbf{E}[\chi_k^{(n)}]$ and ${\sigma^{(n)}}^2 = \mathbf{V}[\chi_k^{(n)}]$. We have $\mathbf{E}[R_t] = u + \pi t - \mu M(t)$ and $\mathbf{E}[R_t^{(n)}] = u^{(n)} + \pi^{(n)}t - \mu^{(n)}M(nt)$. Performing a conditioning argument similar to the one in the classical Normal approximation, we obtain $\mathbf{V}[R_t] = \sigma^2 M(t) + \mu^2 V(t)$ and $\mathbf{V}[R_t^{(n)}] = {\sigma^{(n)}}^2 M(nt) + {\mu^{(n)}}^2 V(nt)$. Equating means and variances and noting that $M(0) = 0$ yields

$$u^{(n)} = u, \quad \forall n \in \mathbb{N} \tag{4.1}$$

$$\mu^{(n)} = \frac{\mu M(t) + (\pi^{(n)} - \pi)t}{M(nt)}, \quad \forall n \in \mathbb{N} \tag{4.2}$$

$${\sigma^{(n)}}^2 = \sigma^2 \frac{M(t)}{M(nt)} + \mu^2 \frac{V(t)}{M(nt)} - {\mu^{(n)}}^2 \frac{V(nt)}{M(nt)}, \quad \forall n \in \mathbb{N} \tag{4.3}$$

Thus, for given choices of $\pi^{(n)}$, we must choose $u^{(n)}$ and $\{\chi_k^{(n)}\}_{k\in\mathbf{N}}$ to satisfy (4.1), (4.2), and (4.3) in order to match the first two moments of R and $R^{(n)}$. It is important to note that if $\mu^{(n)}$ or ${\sigma^{(n)}}^2$ are time dependent then the assumed form of

$R^{(n)}$ is incompatible with moment matching since the claims distributions are time dependent and so are not i.i.d.

For the remainder of this section, we restrict ourselves to the case in which N^{χ} is a Poisson process of rate λ. Then, $M(t) = V(t) = \lambda t$ and from (4.2) and (4.3) we obtain

$$\mu^{(n)} = \frac{\mu}{n} + \frac{\pi^{(n)} - \pi}{n\lambda}, \quad \forall n \in \mathbb{N} \tag{4.4}$$

$$\sigma^{(n)2} = \frac{\sigma^2 + \mu^2}{n} - \left[\frac{\mu}{n} + \frac{(\pi^{(n)} - \pi)}{n\lambda}\right]^2 \quad \forall n \in \mathbb{N} \tag{4.5}$$

which are both independent of time. If we further consider the deterministic parts of the processes separately, not involving them in the weak approximation of the stochastic components, we would set $u^{(n)} + \pi^{(n)}t = u + \pi t$ which, with (4.1), gives

$$\pi^{(n)} = \pi, \quad \forall n \in \mathbb{N} \tag{4.6}$$

and so (4.4) and (4.5) become

$$\mu^{(n)} = \frac{\mu}{n}, \quad \forall n \in \mathbb{N} \tag{4.7}$$

$$\sigma^{(n)2} = \frac{\sigma^2 + \mu^2}{n} - \frac{\mu^2}{n^2}, \quad \forall n \in \mathbb{N} \tag{4.8}$$

That $\mu^{(n)} = \frac{\mu}{n}$ is intuitively appealing since our approximation principle would require the n-fold increase in claim arrivals to be balanced by a $\frac{1}{n}$ reduction in claim sizes to keep the aggregate claims constant. Similarly, the reduction in variance is asymptotically like $\frac{1}{n}$ which is desirable since the variance of the sum of n i.i.d. random variables is n times the variance of one of them, again requiring a n-$\frac{1}{n}$ balance. Note that the claims size reduction is precisely like $\frac{1}{n}$ whereas the reduction in variance is only asymptotically so.

Adopting (4.6) and choosing $\{\chi_k^{(n)}\}_{k\in\mathbb{N}}$ to satisfy (4.7) and (4.8), the approximating sequence takes the form

$$
\begin{aligned}
R_t^{(n)} &= u^{(n)} + \pi^{(n)}t - \sum_{k=1}^{N_t^{(n)}} \chi_k^{(n)} \\
&= u + \pi t - \sum_{k=1}^{N_{nt}^{\chi}} \chi_k^{(n)} \\
&= \frac{1}{\sqrt{n}}\left(u\sqrt{n} + \frac{\pi}{\sqrt{n}}nt - \sum_{k=1}^{N_{nt}^{\chi}} \sqrt{n}\chi_k^{(n)}\right) \\
&= \frac{1}{\sqrt{n}}\left(\tilde{u}^{(n)} + \tilde{\pi}^{(n)}nt - \sum_{k=1}^{N_{nt}^{\chi}} \tilde{\chi}_k^{(n)}\right)
\end{aligned}
$$

where we have set $\tilde{u}^{(n)} = u\sqrt{n}, \tilde{\pi}^{(n)} = \frac{\pi}{\sqrt{n}}$, and $\tilde{\chi}_k^{(n)} = \sqrt{n}\chi_k^{(n)}$ for $k,n \in \mathbb{N}$. Let $\tilde{\mu}^{(n)} = \mathbf{E}[\tilde{\chi}_k^{(n)}] = \frac{\mu}{\sqrt{n}}$ and $\tilde{\sigma}^{(n)^2} = \mathbf{V}[\tilde{\chi}_k^{(n)}] = \sigma^2 + \mu^2 - \frac{\mu^2}{n}$ and, for technical purposes only, add the condition that there $\exists\ \delta > 0$ such that $\sup_n \mathbf{E}[|\tilde{\chi}_k^{(n)}|^{2+\delta}] < \infty$, which is a mild requirement that the tails of $\chi_k^{(n)}$ decrease sufficiently quickly with n. Now, noticing that $\tilde{\sigma}^{(n)^2} = \sigma^2 + \mu^2 - \frac{\mu^2}{n} \to \zeta^2 = \sigma^2 + \mu^2$ and $\tilde{u}^{(n)}, \tilde{\pi}^{(n)}$ and $\{\tilde{\chi}_k^{(n)}\}_{k,n\in\mathbb{N}}$ satisfy the conditions of Iglehart's theorem (Theorem (9)), we can apply his result with $\zeta^2 = \sigma^2 + \mu^2$ to conclude that

$$
R^{(n)} \Rightarrow u + (\pi - \lambda\mu)t - \sqrt{\lambda(\sigma^2 + \mu^2)}W_t \stackrel{\mathcal{D}}{=} u + (\pi - \lambda\mu)t + \sqrt{\lambda(\sigma^2 + \mu^2)}W_t
$$

which is precisely the classical Normal approximation.

This argument provides additional justification for the classical Normal approximation in the form of a rigorous convergence argument constructed according to a satisfying approximation principle and whose limit is fitted to R by the matching of moments. At the same time, it demonstrates a consistency between our approximation procedure and Iglehart's construction; the presence of the crucial balance between the n-fold increase in arrival rate and the $\frac{1}{n}$ decrease in average claim size. Iglehart's choices for $u^{(n)}, \pi^{(n)}$ and $\{\chi_k^{(n)}\}_{k\in\mathbb{N}}$, and his use of the normalization factor $\frac{1}{\sqrt{n}}$ appear to be motivated by the presence of the normalization factor $\frac{1}{\sqrt{n}}$ in

Prohorov's theorem (Theorem 7), the engine behind his proof. It is now clear that his choices merely disguise the n-$\frac{1}{n}$ balance between claim arrivals and claim size. Finally, we give a satisfying interpretation of the term $\lambda\zeta^2 = \lambda(\sigma^2 + \mu^2)$. Recall the formula $\mathbf{V}[R_t] = \mathbf{E}[\mathbf{V}[R_t|N_t^\chi]] + \mathbf{V}[\mathbf{E}[R_t|N_t^\chi]]$. $\mathbf{V}[R_t^{(\infty)}] = \lambda t(\sigma^2 + \mu^2)$ represents the variance accumulated by $R^{(\infty)}$ up to time t. Consider now the accumulated variance $\mathbf{V}[R_t]$ of R up to time t, which has two sources: the variance in jump heights and the variance in the number of jumps over $[0,t]$. During $[0,t]$, R experiences an average of λt jumps, each jump contributing a variance of σ^2 to R. Thus, the "average" accumulated variance in R due only to jump height variance is $\lambda t \times \sigma^2$. This variance corresponds to the term $\mathbf{E}[\mathbf{V}[R_t|N_t^\chi]]$. During $[0,t]$, the Poisson counting process has an accumulated variance of $\lambda t\ jumps^2$, or an accumulated deviation of $\sqrt{\lambda t}\ jumps$, and the average jump height is μ. Thus, the "average" accumulated deviation in R due only to the variance in the number of jumps is $\sqrt{\lambda t} \times \mu$, or an "average" accumulated variance of $\lambda t \times \mu^2$. This variance corresponds to the term $\mathbf{V}[\mathbf{E}[R_t|N_t^\chi]]$. Since jump heights and inter-claim times are independent, we expect to be able to just add these two independent sources of variance to get a total "average" accumulated variance of $\lambda t(\sigma^2 + \mu^2)$ in R over $[0,t]$. This interpretation shows that important information about the probabilistic structure of R (ie: jump height and jump time variances) has been correctly carried into the limit process $R^{(\infty)}$. It also makes clear why the commonly used approximation $u + (\pi - \mu\lambda)t - \sigma\sqrt{\lambda}W_t$ cannot be considered as the appropriate approximation to R: the total process variance of R is dependent on μ and so any reasonable approximation should reflect this dependence. The variance term $\sigma\sqrt{\lambda}$ of the approximation $u + (\pi - \mu\lambda)t - \sigma\sqrt{\lambda}W_t$ is completely insensitive to changes in μ and so has not been correctly fitted to R.

Now, consider R as before except that the i.i.d. claims $\{\chi_k^{(n)}\}_{k\in\mathbb{N}}$ have finite mean μ but infinite variance. The infinite variance rules out the use of the Wiener diffusion approximation but an approximation by an α-stable Lévy process is still possible. Recall that N^χ is Poisson with rate λ. The approximating sequence

$R^{(n)} = u^{(n)} + \pi^{(n)}t - \sum_{k=1}^{N_t^{(n)}} \chi_k^{(n)}$ was constructed with the choices

$$u^{(n)} \quad \text{with} \quad u^{(n)} \to u \tag{4.9}$$

$$\varphi(n) \quad \text{with} \quad \frac{1}{\varphi(n)} \sum_{k=1}^{n} (\chi_k - \mu) \Rightarrow S_\alpha(1, \beta, 0) \tag{4.10}$$

$$\pi^{(n)} \quad \text{with} \quad \left(\pi^{(n)} - \mu\lambda \frac{n}{\varphi(n)}\right) \to c > 0 \tag{4.11}$$

$$N_t^{(n)} = N_{nt}^{\chi} \tag{4.12}$$

$$\chi_k^{(n)} = \frac{1}{\varphi(n)} \chi_k \tag{4.13}$$

which was shown to converge to $R_t^{(\infty)} = u + ct - \lambda^{1/\alpha} L_t^{\alpha,\beta}$. Since $R^{(n)}$ and $R^{(\infty)}$ have finite first moments and N^χ is Poisson with rate λ, (4.1) and (4.4) apply yielding $u^{(n)} = u$ and $\mu^{(n)} = \frac{\mu}{n} + \frac{\pi^{(n)} - \pi}{n\lambda}$, $\forall n \in \mathbb{N}$. Since $\chi_k^{(n)} = \frac{1}{\varphi(n)}\chi_k$ we have that

$$\mu^{(n)} = \mathbf{E}[\chi_k^{(n)}] = \mathbf{E}\left[\frac{\chi_k}{\varphi(n)}\right] = \frac{\mu}{\varphi(n)} \quad \forall n \in \mathbb{N} \tag{4.14}$$

and so combining (4.4) and (4.14) we obtain

$$\pi^{(n)} - \mu\lambda \frac{n}{\varphi(n)} = \pi - \mu\lambda \quad \forall n \in \mathbb{N} \tag{4.15}$$

Thus, matching first moments determines $u^{(n)}$, $\pi^{(n)}$ and that $c = \pi - \mu\lambda$ yielding the "fitted" approximation

$$R_t^{(\infty)} = u + (\pi - \mu\lambda)t - \lambda^{1/\alpha} L_t^{\alpha,\beta}$$

There are two problems with this approximation. First, the dispersion of $R^{(\infty)}$ (measured by the α-stable parameter σ) is the same for all mean claim sizes μ in the original model. As argued in the Wiener approximation, this cannot be considered correct. Second, the $O(n)$ rate of increase in arrivals is not precisely balanced by the $\frac{1}{\varphi(n)}$ decrease in mean claim sizes (recall that $\varphi(n) = n^{1/\alpha} L(n)$ for some slowly varying L) and represents an arrival rate/claims size rescaling that is at odds with our approximation principle. Furthermore, this imbalance necessitates an adjustment of the premium rates according to (4.15) which is also undesirable.

Finally, it should be noted that the sequence of processes defined in Furrer et al. (1996) in the application of their weak approximation to the ruin functional is not weakly convergent. The said process can be written in the form

$$\frac{u}{\varphi(n)} + (\pi - \mu\lambda)\frac{n}{\varphi(n)}t - \frac{1}{\varphi(n)}\sum_{k=1}^{N_{nt}^{\chi}}(\chi_k - \mu) - \mu\left(\frac{N_{nt}^{\chi} - n\lambda t}{\varphi(n)}\right)$$

and has as a drift term $\pi^{(n)} - \lambda\mu\frac{n}{\varphi(n)} = (\pi - \lambda\mu)\frac{n}{\varphi(n)}$ which does not converge to some $c > 0$ if $\pi - \mu\lambda \neq 0$ (see Theorem (15) and the discussion following Proposition (8)). Thus, their application does not make direct use of their theorem on the convergence of functionals (16); the ruin functional is considered in an ad hoc manner and in isolation.

As a way of recovering the balance between arrival rate and claim sizes in the α-stable Lévy process approximation, as well as to include a dependence of the limiting dispersion on the mean claim size, we provide the following convergence argument which properly includes Iglehart's result but is not included in the result of Furrer et al.

Theorem 17 *Let $R_t = u + \pi t - \sum_{k=1}^{N_t^{\chi}} \chi_k$ be a reserve model where N^{χ} is an ordinary renewal counting process with mean inter-claim time λ^{-1} and χ_k are i.i.d. $S_\alpha(\sigma, \beta, \mu)$ random variables for some $\alpha \in (1, 2]$, $\sigma \in \mathbb{R}_+ \setminus \{0\}$, $\beta \in [-1, 1]$, and $\mu \in \mathbb{R}$. Define a sequence of processes by $R_t^{(n)} = u + \pi t - \sum_{k=1}^{N_{nt}^{\chi}} \chi_k^{(n)}$ for $\{\chi_k^{(n)}\}_{k\in\mathbb{N}}$ i.i.d. $S_\alpha(\frac{\sigma^{(n)}}{n^{-1/\alpha}}, \beta, \frac{\mu}{n})$ random variables and where $\sigma^{(n)} \to \zeta$ for some $\zeta > 0$ with $\sigma^{(1)} = \sigma$. Then,*

$$R_t^{(n)} \Rightarrow R^{(\infty)} \stackrel{def}{=} u + (\pi - \mu\lambda)t - \zeta\lambda^{1/\alpha}L_t^{\alpha,\beta}$$

Proof:

$$\begin{aligned}
R_t^{(n)} &= u + \pi t - \sum_{k=1}^{N_{nt}^{\chi}} \chi_k^{(n)} \\
&= u + \pi t - \mu\frac{N_{nt}^{\chi}}{n} - \sum_{k=1}^{N_{nt}^{\chi}}\left(\chi_k^{(n)} - \frac{\mu}{n}\right) \\
&= u + \pi t - \mu\frac{N_{nt}^{\chi}}{n} - \frac{\sigma^{(n)}}{n^{1/\alpha}}\sum_{k=1}^{N_{nt}^{\chi}}\left(\tilde{\chi}_k^{(n)} - \frac{\mu}{\sigma^{(n)}n^{1-1/\alpha}}\right)
\end{aligned}$$

where we have set $\tilde{\chi}_k^{(n)} = \frac{n^{1/\alpha}}{\sigma^{(n)}}\chi_k^{(n)}$ so that

$$\tilde{\chi}_k^{(n)} - \frac{\mu}{\sigma^{(n)} n^{1-1/\alpha}} = \frac{n^{1/\alpha}}{\sigma^{(n)}}\left(\chi_k^{(n)} - \frac{\mu}{n}\right) \sim S_\alpha(1, \beta, 0)$$

From the arithmetic properties of α-stable-stable distributions,

$$\frac{\sigma^{(n)}}{n^{1/\alpha}} \sum_{k=1}^{n} \left(\tilde{\chi}_k^{(n)} - \frac{\mu}{\sigma^{(n)} n^{1-1/\alpha}}\right) \sim S_\alpha(\sigma^{(n)}, \beta, 0)$$

and, since $\sigma^{(n)} \to \zeta$, we have the trivial convergence

$$\frac{\sigma^{(n)}}{n^{1/\alpha}} \sum_{k=1}^{n} \left(\tilde{\chi}_k^{(n)} - \frac{\mu}{\sigma^{(n)} n^{1-1/\alpha}}\right) \Rightarrow S_\alpha(\zeta, \beta, 0)$$

Since $\frac{N^{\chi}_{nt}}{n} \Rightarrow \lambda t$ we then have the convergence of the composition of the processes

$$\frac{\sigma^{(n)}}{n^{1/\alpha}} \sum_{k=1}^{N^{\chi}_{nt}} \left(\tilde{\chi}_k^{(n)} - \frac{\mu}{\sigma^{(n)} n^{1-1/\alpha}}\right) \Rightarrow \zeta \lambda^{1/\alpha} L_t^{\alpha,\beta}$$

and the result follows.

□

This formulation has the following desirable properties: $R^{(1)} = R$, $\mathbf{E}[\chi_k^{(n)}] = \frac{\mu}{n}$ which exactly offsets the $O(n)$ arrival rate increase, the premium rate is not adjusted to obtain convergence, the extra parameter ζ allows a fitting of the limiting dispersion, and the crucial shape parameters α and β remain constant during convergence: the trivial convergence $S_\alpha(\sigma^{(n)}, \beta, 0) \Rightarrow S_\alpha(\zeta, \beta, 0)$ is "clean" in the sense that the claims distribution is not reshaped by the convergence procedure and so no information about the claims distribution of R is distorted or lost in the limit.

The two disadvantages are that ζ cannot be specified by matching second moments since they don't exist. However, $\mathbf{E}[|\chi_k|^p] < \infty$ for $0 \leqslant p < \alpha$ and all logarithmic moments exist (Zolotarev 1986) so it may be possible to determine an appropriate value for ζ through the matching of fractional or logarithmic moments. Another drawback is that the claims distribution of R may not be α-stable. The family $S_\alpha(\sigma, \beta, \mu)$ is extremely rich, possessing the right qualitative features for claims distributions and so one may be able to find values of α, β, μ, and σ that provide a good fit to the claims distribution (see Samorodnitsky & Taqqu (1994)).

4.2 Weak Approximation and Infinite Divisibility

The three cases discussed in the previous section essentially involved the weak approximation of compound Poisson sums, re-centered about zero by subtracting off their means. We are interested in studying the nature of weak limits of sequences of compound Poisson sum laws. A corollary to the Lévy-Khintchine Theorem is that the set of all ID laws is the weak closure of the set of all compound Poisson sum laws, or, equivalently, the point-wise closure of the set of all characteristic functions of compound Poisson sums, which we now examine closely.

Let $Z_t = \sum_{k=1}^{N_t} X_k$ where N is Poisson with finite rate λ and $\{X_k\}_{k\in\mathbb{N}}$ are i.i.d. jumps with law $P \in \mathcal{P}_{\mathbb{R}}$. Since N_t has law

$$\mathbf{P}\{N_t = n\} = \frac{(\lambda t)^n e^{-\lambda t}}{n!}, \quad n \in \mathbb{N}_0$$

the law of Z_t is given by

$$\begin{aligned}
\mathbf{P}Z_t^{-1}(-\infty, x] = \mathbf{P}\{Z_t \leqslant x\} &= \sum_{n\in\mathbb{N}_0} \mathbf{P}\{Z_t \leqslant x | N_t = n\}\mathbf{P}\{N_t = n\} \\
&= \sum_{n\in\mathbb{N}_0} \mathbf{P}\{X_1 + \cdots + X_n \leqslant x | N_t = n\} \frac{(\lambda t)^n e^{-\lambda t}}{n!} \\
&= \sum_{n\in\mathbb{N}_0} P^{*n}(-\infty, x] \frac{(\lambda t)^n e^{-\lambda t}}{n!}
\end{aligned}$$

where we take $P^{*0} = \delta_0$, the point mass at 0. Therefore $\mathbf{P}Z_t^{-1} = e^{-\lambda t}\sum_{n\in\mathbb{N}_0} \frac{(\lambda t)^n}{n!} P^{*n}$. Using dominated convergence, the characteristic function is given by

$$\begin{aligned}
\widehat{\mathbf{P}Z_t^{-1}}(\theta) &= e^{-\lambda t}\int e^{i\theta x}\sum_{n\in\mathbb{N}_0}\frac{(\lambda t)^n}{n!} dP^{*n}(x) = e^{-\lambda t}\sum_{n\in\mathbb{N}_0}\frac{(\lambda t)^n}{n!}\int e^{i\theta x} dP^{*n}(x) \\
&= e^{-\lambda t}\sum_{n\in\mathbb{N}_0}\frac{(\lambda t)^n}{n!}\widehat{P^{*n}}(\theta) = e^{-\lambda t}\sum_{n\in\mathbb{N}_0}\frac{(\lambda t)^n}{n!}[\widehat{P}(\theta)]^n \\
&= e^{-\lambda t}e^{\lambda t \widehat{P}(\theta)} = e^{\lambda t(\widehat{P}(\theta)-1)} \\
&= e^{\lambda t \int (e^{i\theta x}-1) dP(x)}
\end{aligned}$$

and thus Z_t is ID. Denote the set of all such characteristic functions by

$$\widehat{\mathcal{P}_0} = \left\{ \exp\left[t\lambda \int_{\mathbb{R}} \left(e^{i\theta x} - 1\right) dP(x) \right] : P \in \mathcal{P}_{\mathbb{R}} \right\}$$

All possible weak limits of compound Poisson sum laws correspond to the point-wise closure of $\widehat{\mathcal{P}}_0$. We examine the structure of these weak limits by considering how the point-wise closure of $\widehat{\mathcal{P}}_0$ is formed. The free parameters of $\widehat{\mathcal{P}}_0$ are the finite arrival rate λ and the jump law P. Combine these paramteres by defining the measure $M = \lambda P \in \mathcal{M}^1_{\mathbb{R}}$ and consider all sequences $\{M^{(n)}\}_{n \in \mathbb{N}}$ such that $\lim_{n\to\infty} \int_{\mathbb{R}} (e^{i\theta x} - 1) dM^{(n)}(x)$ exists for each θ.

First, M is a finite measure on $\mathbb{R}$ and $e^{i\theta x} - 1$ is zero for $x = 0$ so we can eliminate zero from the range of integration:

$$\widehat{\mathcal{P}_0} = \left\{ \exp\left[t \int_{\mathbb{R}\setminus\{0\}} \left(e^{i\theta x} - 1\right) dM(x) \right] : M \in \mathcal{M}^1_{\mathbb{R}} \right\}$$

For $M \in \mathcal{M}^1_{\mathbb{R}}$, the unit mass at zero δ_0, and $f : \mathbb{R} \to \mathbb{R}$ define

$$M^{(n)} = M + f(n)\delta_0 \in \mathcal{M}^1_{\mathbb{R}}$$

Now, $\lim_{n\to\infty} \int_{\mathbb{R}\setminus\{0\}} \left(e^{i\theta x} - 1\right) dM^{(n)}(x) = \int_{\mathbb{R}\setminus\{0\}} (e^{i\theta x} - 1) dM(x)$ for each θ. Thus we can drop the restriction that $M(\{0\})$ is finite; we only require that $M \in \mathcal{M}^1_{\mathbb{R}\setminus\{0\}}$:

$$\widehat{\mathcal{P}_0} = \left\{ \exp\left[t \int_{\mathbb{R}\setminus\{0\}} \left(e^{i\theta x} - 1\right) dM(x) \right] : M \in \mathcal{M}^1_{\mathbb{R}\setminus\{0\}} \right\}$$

For $M \in \mathcal{M}^1_{\mathbb{R}\setminus\{0\}}$ and $a \in \mathbb{R}$ define another sequence by

$$M^{(n)} = M + n\delta_{\frac{a}{n}} \in \mathcal{M}^1_{\mathbb{R}\setminus\{0\}}$$

It is easy to show that $\lim_{n\to\infty} \int_{\mathbb{R}\setminus\{0\}} \left(e^{i\theta x} - 1\right) dM^{(n)}(x) = i\theta a + \int_{\mathbb{R}\setminus\{0\}} (e^{i\theta x} - 1) dM(x)$. Since $a \in \mathbb{R}$ is arbitrary we have enlarged $\widehat{\mathcal{P}_0}$ to

$$\widehat{\mathcal{P}_1} = \left\{ \exp\left[i\theta a t + t \int_{\mathbb{R}\setminus\{0\}} \left(e^{i\theta x} - 1\right) dM(x) \right] : M \in \mathcal{M}^1_{\mathbb{R}\setminus\{0\}} \right\}$$

For $M \in \mathcal{M}^1_{\mathbb{R}\setminus\{0\}}$ and $\sigma \in \mathbb{R}_+$ define a sequence by

$$M^{(n)} = M + \frac{n^2}{2}\left(\delta_{\frac{\sigma}{n}} + \delta_{\frac{\sigma}{n}}\right) \in \mathcal{M}^1_{\mathbb{R}\setminus\{0\}}$$

for which $\lim_{n\to\infty} \int_{\mathbb{R}\setminus\{0\}} \left(e^{i\theta x} - 1\right) dM^{(n)}(x) = -\frac{1}{2}\theta^2\sigma^2 + \int_{\mathbb{R}\setminus\{0\}}(e^{i\theta x} - 1)dM(x)$ and so $\widehat{\mathcal{P}_1}$ has been enlarged to

$$\widehat{\mathcal{P}_2} = \left\{\exp\left[i\theta a t - \frac{\theta^2\sigma^2}{2}t + t\int_{\mathbb{R}\setminus\{0\}}\left(e^{i\theta x} - 1\right) dM(x)\right] : M \in \mathcal{M}^1_{\mathbb{R}\setminus\{0\}}\right\}$$

where $a \in \mathbb{R}$ and $\sigma \in \mathbb{R}_+$ are arbitrary.

The condition that $M \in \mathcal{M}^1_{\mathbb{R}\setminus\{0\}}$ can be weakened further to $M \in \mathcal{M}^{1\wedge|x|}_{\mathbb{R}\setminus\{0\}}$ since it can be seen that

$$\left|\int_{\mathbb{R}\setminus\{0\}}(e^{i\theta x} - 1)dM(x)\right| \leqslant 2e^{|\theta|}\int_{\mathbb{R}\setminus\{0\}}(1 \wedge |x|)\,|dM(x)|$$

This weakening can be achieved, for example, by the sequence of measures

$$M^{(n)} = \left(\frac{\lambda}{1 \wedge |x|^{\frac{n-1}{n}}}\right) P \Rightarrow \frac{\lambda}{1 \wedge |x|} P$$

and results in the enlargement

$$\widehat{\mathcal{P}_3} = \left\{\exp\left[i\theta a t - \frac{\theta^2\sigma^2}{2}t + t\int_{\mathbb{R}\setminus\{0\}}(e^{i\theta x} - 1)dM(x)\right] : M \in \mathcal{M}^{1\wedge|x|}_{\mathbb{R}\setminus\{0\}}\right\}$$

A further weakening to $M \in \mathcal{M}^{1\wedge x^2}_{\mathbb{R}\setminus\{0\}}$ completes the closure of $\widehat{\mathcal{P}}_0$. However, there is a slight complication; the integral $\int_{\mathbb{R}\setminus\{0\}}(e^{i\theta x} - 1)dM(x)$ does not exists for $M \in \mathcal{M}^{1\wedge x^2}_{\mathbb{R}\setminus\{0\}} \setminus \mathcal{M}^{1\wedge|x|}_{\mathbb{R}\setminus\{0\}}$. The Taylor series expansion of $e^{i\theta x} - 1$ about $x = 0$ yields $e^{i\theta x} - 1 \approx i\theta x$ for x near zero which does not go to zero quickly enough to dominate the $\frac{1}{x^2}$ singularity of M near $x = 0$. If we subtract off the term $i\theta x 1_{[-\epsilon,\epsilon]}(x)$ for any $\epsilon > 0$ then near $x = 0$, $e^{i\theta x} - 1 - i\theta 1_{[-\epsilon,\epsilon]}(x) \approx -\frac{\theta^2\sigma^2}{2}$ which is enough to dominate the $\frac{1}{x^2}$ singularity of M. Setting $\epsilon = 1$ for convenience and noting that for $M \in \mathcal{M}^{1\wedge|x|}_{\mathbb{R}\setminus\{0\}}$ both integrals

$$\int_{\mathbb{R}\setminus\{0\}}(e^{i\theta x} - 1 - i\theta x 1_{[-1,1]}(x))dM(x)$$

$$i\theta \int_{\mathbb{R}\setminus\{0\}} x 1_{[-1,1]} dM(x)$$

exist, we can extract the second integral and rewrite $\widehat{\mathcal{P}_3}$ in the form

$$\widehat{\mathcal{P}_3} = \left\{ \exp\left[i\theta\mu t - \frac{\theta^2\sigma^2}{2} t + t \int_{\mathbb{R}\setminus\{0\}} (e^{i\theta x} - 1 - i\theta x 1_{[-1,1]}(x)) dM(x) \right] : M \in \mathcal{M}^{1\wedge|x|}_{\mathbb{R}\setminus\{0\}} \right\}$$

where we have set $\mu = a + \int_{\mathbb{R}\setminus\{0\}} x 1_{[-1,1]}(x) dM(x)$. Now it can be seen that

$$\left| \int_{\mathbb{R}\setminus\{0\}} (e^{i\theta x} - i\theta 1_{[-1,1]}(x)) dM(x) \right| \leqslant 2e^{|\theta|} \int_{\mathbb{R}\setminus\{0\}} (1 \wedge x^2) \, |dM(x)|$$

and so allowing a weakening to $M \in \mathcal{M}^{1\wedge x^2}_{\mathbb{R}\setminus\{0\}}$, for example, by the sequence

$$M^{(n)} = \left(\frac{\lambda}{1 \wedge x^{\frac{2n-2}{n}}} \right) P \Rightarrow \frac{\lambda}{1 \wedge x^2} P$$

maintains the integrability of $e^{i\theta x} - 1 - i\theta x 1_{[-1,1]}$ and so we obtain the enlargement

$$\widehat{\mathcal{P}_4} = \left\{ \exp\left[i\theta\mu t - \frac{\theta^2\sigma^2}{2} t + t \int_{\mathbb{R}\setminus\{0\}} \left(e^{i\theta x} - 1 - i\theta x 1_{[-1,1]}(x)\right) dM(x) \right] : M \in \mathcal{M}^{1\wedge x^2}_{\mathbb{R}\setminus\{0\}} \right\}$$

where $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}_+$ are arbitrary. By the Lévy-Khintchine Theorem, $\overline{\overline{\widehat{\mathcal{P}_0}}} = \widehat{\mathcal{P}_4}$ and we have obtained the closure.

Each step in the forming the closure has a probabilistic interpretation. For the compound Poisson process $Z_t = \sum_{k=1}^{N_t} X_k$ where N is Poisson with rate λ and jump law P, we have the Lévy measure $M = \lambda P \in \mathcal{M}^1_{\mathbb{R}}$: all jumps arrive at the rate λ. The sequence $M^{(n)} = M + f(n)\delta_0$ allows M to be singular at $x = 0$ which corresponds to allowing the jumps of zero height to arrive at an arbitrary rate, even infinitely quickly. This weakening does not result in new characteristic functions since jumps of zero height contribute nothing to the process.

The sequence $M^{(n)} = M + n\delta_{\frac{a}{n}}$ corresponds to a perturbation of the compound Poisson process by an external agent that causes additional jumps of height $\frac{a}{n}$ to arrive at rate n. The infinitesimal perturbation "$\lim_{n\to\infty} n\delta_{\frac{a}{n}}$" is a uniform embedding of infinitesimally small jumps arriving infinitely quickly and results in the appearance of a new drift term $i\theta a$ in the characteristic function.

The sequence $M^{(n)} = M + \frac{n^2}{2}\left(\delta_{-\frac{\sigma}{n}} + \delta_{\frac{\sigma}{n}}\right)$ is again due to an externally supplied perturbation that causes additional symmetric jumps of heights $-\frac{\sigma}{n}$ and $\frac{\sigma}{n}$ to arrive at the rate $\frac{n^2}{2}$. The infinitesimal perturbation "$\lim_{n\to\infty} \frac{n^2}{2}\left(\delta_{-\frac{\sigma}{n}} + \delta_{\frac{\sigma}{n}}\right)$" is a symmetric, uniform embedding of infinitesimally small jumps arriving infinitely quickly. However, the rate $\frac{n^2}{2}$ is much faster than the rate n of the first perturbation and so these symmetric jumps are more densely embedded. The first order effects of the equal and opposite jumps cancel, resulting in no contribution to the drift but the density of these jumps is such that a second order effect appears: variance. Ie, the symmetric jumps are sufficiently dense that at any time scale, there are an equal number of up jumps and down jumps and hence no net drift. However, they are not so dense that all randomness associated with the arrival of these symmetric jumps is lost. Thus, for $M \in \mathcal{M}_{\mathbb{R}\setminus\{0\}}$, the compound Poisson process Z subject to these two perturbations has a characteristic function of the form

$$\exp\left[i\theta a t - \frac{\theta^2\sigma^2}{2}t + t\int_{\mathbb{R}\setminus\{0\}} (e^{i\theta x} - 1)dM(x)\right]$$

which corresponds to a process of the form $at + \sigma W_t + Z_t$ where W is a standard Wiener process independent of Z.

The successive weakenings $M \in \mathcal{M}_{\mathbb{R}\setminus\{0\}}^{1\wedge|x|}$ and $M \in \mathcal{M}_{\mathbb{R}\setminus\{0\}}^{1\wedge x^2}$ allow for smaller jumps to arrive at different rates. The examples to keep in mind are

$$M_1 = \frac{\lambda}{1 \wedge |x|}P$$

$$M_2 = \frac{\lambda}{1 \wedge x^2}P$$

For jumps outside of $[-1, 1]$, $M_1 = M_2 = \lambda P$ and so these jumps all arrive at rate λ. For jumps in $[-1, 1] \setminus \{0\}$ we have

$$dM_1(x) = \frac{\lambda}{|x|}dP(x)$$

$$dM_2(x) = \frac{\lambda}{x^2}dP(x)$$

with non-uniform arrival rates $\frac{\lambda}{|x|}$ and $\frac{\lambda}{x^2}$, respectively. These rates are non-uniform and are obtained as the original uniform rate λ is made to vary in the limit using the sequences

$$M^{(n)} = \left(\frac{\lambda}{1 \wedge |x|^{\frac{n-1}{n}}}\right) P \Rightarrow \frac{\lambda}{1 \wedge |x|} P$$

$$M^{(n)} = \left(\frac{\lambda}{1 \wedge x^{\frac{2n-2}{n}}}\right) P \Rightarrow \frac{\lambda}{1 \wedge x^2} P$$

As mentioned, the weakening to $M \in \mathcal{M}^{1\wedge x^2}_{\mathbb{R}\setminus\{0\}}$ has an interesting complication which we now describe. Since $\mathbf{P}\{N_t < \infty\} = \sum_{n\in\mathbb{N}_0} \mathbf{P}\{N_t = n\} = \sum_{n\in\mathbb{N}_0} \frac{(\lambda t)^n e^{-\lambda t}}{n!} = 1$ and $P \in \mathcal{P}_{\mathbb{R}}$, so that $\mathbf{P}\{X_k < \infty\} = 1$, Z is without explosion. The total variation of Z on $[0,t]$, given by $V_0^t(\omega) = \sum_{k=1}^{N_t(\omega)} |X_k(\omega)|$, is $\mathbf{P}$-a.s. finite. However, the average total variation $\mathbf{E}[V_0^t] = \mathbf{E}[N_t]\mathbf{E}[|X_k|] = \lambda t \int_{\mathbb{R}} |x| dP(x)$ may be infinite. If $\mathbf{E}[V_0^t] < \infty$, then $\mathbf{E}[Z_t] = \mathbf{E}[N_t]\mathbf{E}[X_k] = \lambda t \int_{\mathbb{R}} x dP(x) =$ (mean # of jumps in [0,t])$\times$(mean jump height) is finite and represents the average net variation of Z during the time interval $[0,t]$; $\lambda \int_{\mathbb{R}\setminus\{0\}} x dP(x)$ is therefore the average net variation over the unit interval $[0,1]$, or, the drift rate of Z. If $\mathbf{E}[V_0^t] = \infty$ then we cannot identify the drift rate as the average net variation over $[0,1]$. We are interested in weak limits involving claim size rescalings so we focus our attention on the average net variation due to those jumps with sizes in $[-1,1]$ by using the truncation function $\psi(x) = x 1_{[-1,1]}(x)$. Now, $\int_{\mathbb{R}\setminus\{0\}} x 1_{[-1,1]}(x) dP(x)$ exists for any $P \in \mathcal{P}_{\mathbb{R}}$, $\frac{1}{P([-1,1])} \int_{\mathbb{R}} x 1_{[-1,1]}(x) dP(x)$ is the average jump size of those jumps with magnitudes in $[-1,1]$, and so $\int_{\mathbb{R}} x 1_{[-1,1]}(x) dP(x)$ is a weighted contribution of the average net effect of the smaller jumps. As stated in the Lévy-Khintchine Theorem, any bounded, measurable function $\psi(x)$ satisfying $\sup_{\mathbb{R}\setminus\{0\}} \left|\frac{\psi(x)-x}{x^2}\right| < \infty$ would emphasize the smaller jumps; ψ being bounded rescales larger jumps to within its bounds so that $\int_{\mathbb{R}\setminus\{0\}} \psi(x) dP(x)$ is finite and the supremum condition forces $\psi(x)$ to behave like x near zero, leaving the smaller jumps at their original size. The choice $\psi(x) = \frac{x}{1+x^2}$ is differentiable and is useful in calculations and limit theorems; this was the function used in calculating the characteristic function of α-stable densities.

$\psi(x) = x1_{[-1,1]}(x)$ is more intuitive for understanding ID distributions since it completely disregards the larger jumps and leaves the jumps in $[-1, 1]$ scaled exactly as they were. Now define $b = \lambda \int_{\mathbb{R}} x1_{[-1,1]}(x)dP(x) = \int_{\mathbb{R}\setminus\{0\}} x1_{[-1,1]}(x)dM(x)$, which is finite, to represent the weighted contribution of the smaller jumps to the average net variation over $[0, 1]$. Note that b is neither arbitrary nor unique but depends on the choice of $\psi(x)$. Other choices of $\psi(x)$ result only in a reparametrization of the family of characteristic functions by $b \mapsto b + \int_{\mathbb{R}\setminus\{0\}} (\psi(x) - x1_{[-1,1]}(x))\, dM(x)$.

Now, up to and including the weakening $M \in \mathcal{M}^{1\wedge|x|}_{\mathbb{R}\setminus\{0\}}$, b is finite and so we can extract or reabsorb this drift in the characteristic functions:

$$\begin{aligned}
\widehat{\mathcal{P}_3} &= \left\{\exp\left[i\theta a t - \frac{\theta^2\sigma^2}{2}t + t\int_{\mathbb{R}\setminus\{0\}}(e^{i\theta x} - 1)dM(x)\right] : M \in \mathcal{M}^{1\wedge|x|}_{\mathbb{R}\setminus\{0\}}\right\} \\
&= \left\{\exp\left[i\theta\mu t - \frac{\theta^2\sigma^2}{2}t + t\int_{\mathbb{R}\setminus\{0\}}(e^{i\theta x} - 1 - i\theta x1_{[-1,1]}(x))dM(x)\right] : M \in \mathcal{M}^{1\wedge|x|}_{\mathbb{R}\setminus\{0\}}\right\}
\end{aligned}$$

where we have set $\mu = a+b$. Thus, we can decompose these processes as $at+\sigma W_t+Z_t$ or $\mu t + \sigma W_t + J_t$ where J is a "partially compensated" process that results from extracting from Z the average net effect of the jumps in $[-1, 1]$.

However, in the weakening $M \in \mathcal{M}^{1\wedge x^2}_{\mathbb{R}\setminus\{0\}}$, the internally arising drift b must first be extracted before the limit is taken. This drift cannot be reabsorbed into the limiting integral and so we no longer have two decompositions of the process for $M \in \mathcal{M}^{1\wedge x^2}_{\mathbb{R}\setminus\{0\}} \setminus \mathcal{M}^{1\wedge|x|}_{\mathbb{R}\setminus\{0\}}$, we only have the decomposition $\mu t + \sigma W_t + J_t$.

For $M \in \mathcal{M}^{1\wedge|x|}_{\mathbb{R}\setminus\{0\}}$ it is always possible to distinguish between the internal drift of the compound Poisson process and the external drift due to the infinitesimal perturbation "$\lim_{n\to\infty} n\delta_{\frac{a}{n}}$" since

$$\begin{aligned}
a = \text{external drift} &= \mu - \int_{\mathbb{R}\setminus\{0\}} x1_{[-1,1]}(x)dM(x) \\
b = \text{internal drift} &= \int_{\mathbb{R}\setminus\{0\}} x1_{[-1,1]}(x)dM(x)
\end{aligned}$$

are both finite quantities. As the limit is taken, the integrals $\int_{\mathbb{R}\setminus\{0\}}(e^{i\theta x} - 1)dM^{(n)}(x)$ and $\int_{\mathbb{R}\setminus\{0\}} x1_{[-1,1]}(x)dM^{(n)}(x)$ diverge even though $\int_{\mathbb{R}\setminus\{0\}}(e^{i\theta x} - 1 - i\theta\psi(x))dM^{(n)}(x)$

converges. Also, there is no other rescaling $\psi(x)$ that will make all three integrals exist in the limit simultaneously (Stroock 1993). The embedding of the smaller jumps is just too dense to isolate their effect from any externally applied perturbation; the drifts are bound together in μ in such a way that results in the renormalization of infinite quantities in the limit: μ is finite but

$$\begin{aligned} a = \text{external drift} &= \mu - \int_{\mathbb{R}\setminus\{0\}} x1_{[-1,1]}(x)dM(x) = \mp\infty \\ b = \text{internal drift} &= \int_{\mathbb{R}\setminus\{0\}} x1_{[-1,1]}(x)dM(x) = \pm\infty \end{aligned}$$

This marks a fundamental shift in the path structure.

We now examine the limiting procedures that give rise to the Wiener and α-stable Lévy process approximations in terms of nonuniform changes in arrival rates, changes in jump laws, and external perturbations. We formed the closure of $\widehat{\mathcal{P}}_0$ by sequences of the form

$$dM^{(n)}(x) = \lambda(n,x)dP(x)$$

We set $\lambda(1,x) = \lambda$ so that $M^{(1)} = \lambda P$ corresponds to the original compound Poisson process of rate λ, jump law P, and characteristic function $\exp\left[t\int(e^{i\theta x}-1)dM(x)\right]$.

Case 1: $\lambda(n,x) = \lambda^{(n)}$ such that $\lambda^{(1)} = \lambda$ and $\lambda^{(n)} \to \lambda^{(\infty)}$. Thus $M^{(n)} = \lambda^{(n)}P$ has a constant jump law, all jumps arriving with rate $\lambda^{(n)}$ and which changes uniformly to $\lambda^{(\infty)}$; all jumps experience the same uniform change in arrival rate and so only the time of the original process has been rescaled. The limit is another compound Poisson sum with finite rate $\lambda^{(\infty)}$ and jump law P and if $\lambda \neq \lambda^{(\infty)}$ then

$$\lambda\int_{\mathbb{R}\setminus\{0\}}(e^{i\theta x}-1)dP(x) \neq \lambda(\infty)\int_{\mathbb{R}\setminus\{0\}}(e^{i\theta x}-1)dP(x)$$

Case 2: $\lambda(n,x)$ is bounded in x and $\lambda(n,x) \to \lambda(\infty,x)$. For example, $\lambda(n,x) = \frac{1}{2}\lambda^{(n)}(1+|x|^{1-n})$ where $\lambda^{(1)} = \lambda$ and $\lambda^{(n)} \to \lambda^{(\infty)}$. In this example, jumps of different sizes can arrive at different rates but, asymptotically, all jumps arrive at the rate $\frac{1}{2}\lambda^{(\infty)}$ and have jump law P. Another example is $\lambda(n,x) = \lambda(n)[(1-\frac{1}{n})f(x)+\frac{1}{n}]$ for some function $f(x)$ bounded in x. In this case the limiting jump law Q and rate are

$$dQ(x) = \frac{f(x)dP(x)}{\int f(x)dP(x)} \qquad \text{and} \qquad \lambda^{(\infty)} \int f(x)dP(x)$$

In either case, we obtain compound Poisson laws. However, in this case it would be interesting to investigate if it is possible to find a $\lambda(n,x)$ such that

$$\lambda \int_{\mathbb{R}\setminus\{0\}} (e^{i\theta x} - 1)dP(x) = \int_{\mathbb{R}\setminus\{0\}} (e^{i\theta x} - 1)\lambda(n,x)dP(x) = \lambda^{(\infty)} \int_{\mathbb{R}\setminus\{0\}} (e^{i\theta x} - 1)dQ(x)$$

for all n; perhaps it is possible to "steer" via $\lambda(n,x)$ the original process of rate λ and jump law P to another compound Poisson process with rate $\lambda^{(\infty)}$ and jump law Q that is particularly easy to work with, for example, where Q is exponential, Normal, or Poisson.

Case 3: Suppose $\lambda(n,x)$ converges to $\lambda(\infty,x)$ for each x and is non-constant in x; it may be unbounded in x for each n or in the limit as $n \to \infty$. This case is seen in the last two steps in forming the closure where the conditions on M were weakened twice. For any n we have that $\lambda(n,x)$ can at most become singular near $x = 0$ like $\frac{1}{|x|}$ or $\frac{1}{x^2}$. Such non-uniform changes in the jump arrival rates is a significant perversion of the original process and is the nature of the α-stable Lévy process approximation. Recall that the Lévy measure M is given by

$$dM(x) = \frac{Q}{|x|^{\alpha+1}} 1_{(-\infty,0)}(x)dx + \frac{P}{x^{\alpha+1}} 1_{(0,\infty)}(x)dx$$

where $P, Q \in \mathbb{R}_+$, $P+Q > 0$, and $\alpha \in (0,2)$. For convenience we take $P = Q = 1$ and write the measure as $dM(x) = |x|^{-\alpha-1} 1_{x\neq 0}(x)dx$. For any $\alpha \in (0,2)$, $\int_{\mathbb{R}\setminus\{0\}} dM(x) = \infty$ and so we are not in the regime of compound Poisson processes. It is easy to verify that $\int_{\mathbb{R}\setminus\{0\}} (1 \wedge x^2)dM(x)$ is finite for $\alpha \in (0,2)$. However, $\int_{\mathbb{R}\setminus\{0\}} (1 \wedge |x|)dM(x)$ is finite for $\alpha \in (0,1)$ and is infinite for $\alpha \in [1,2)$. Thus, the α-stable Lévy processes are neatly split into two classes by their Lévy measures. From the previous discussion, we know that for $\alpha \in (0,1)$ we can extract the internal drift and so can distinguish it from any external perturbation resulting in drift. However, for $\alpha \in [1,2)$, we cannot. Probabilistically, this means that the α-stable Lévy processes for $\alpha \in (0,1)$

are "closer" in structure to compound Poisson processes than those with $\alpha \in [1, 2)$. This is intuitively apparent in the simulated paths of α-stable Lévy processes as presented in Janicki & Weron (1994) and Samorodnitsky & Taqqu (1994). This singularly non-uniform distortion of the original arrival rate, the indistinguishability of the drifts for $\alpha \in [1, 2)$, and the ability to only match first moments raises doubt about the suitability of α-stable Lévy process approximations and may explain the large relative errors observed in simulation comparisons (Furrer 1996).

Case 4: The first order perturbation $n\delta_{\frac{a}{n}}$ is an insertion of a jump of height $\frac{a}{n}$ with rate n. This is a very localized disturbance of the compound Poisson process in the sense that the rate of only one jump is increased. The limit results in the addition of a linear drift rate, which represents only a translation of the compound Poisson law. The second order perturbation $\frac{n^2}{2}(\delta_{-\frac{\sigma}{n}} + \delta_{\frac{\sigma}{n}})$ is again a localized insertion of jumps but results in something qualitatively new. As mentioned, the first order drift effect of the symmetric jumps cancel. The second order effect arises from the much faster arrival rate and the limit amounts to an insertion of an independent Wiener process with variance σ^2, and significantly changes the character of the process. In this light, the convergence of sequences of Poisson reserve processes to a Wiener process is extremely contrived. For example, consider the measure

$$M^{(n)} = \frac{n-1}{n}\left(n\delta_{\frac{\mu}{n}} + \frac{n^2}{2}(\delta_{\frac{-\sigma\epsilon}{n}} + n\delta_{\frac{\sigma}{n}})\right) + \frac{\lambda}{n}P$$

where $M = \lambda P$ corresponds to the compound Poisson process to be approximated. It is easy to see that

$$\lim_{n\to\infty}\int_{\mathbb{R}\setminus\{0\}}(e^{i\theta x} - 1)dM^{(n)}(x) = i\theta\mu - \frac{1}{2}\theta^2\sigma^2$$

There is no careful balancing of rates and jump sizes in this example; μ and σ^2 are completely arbitrary and are not connected to the original process in any way. Even if they were, this convergence argument is simply the progressive phasing in of one ID law and phasing out of another and therefore seems quite arbitrary.

For our approximation procedure, a more precise formulation of the fitting problem is this. Given a compound Poisson process of rate λ and jump law P and a subdivision of the process that results in an n-fold increase in arrival rate, ie: $n\lambda$, can we find a sequence of probability measures $\{P^{(n)}\}_{n\in\mathbb{N}}$ such that

$$\lambda \int_{\mathbb{R}\setminus\{0\}} (e^{i\theta x} - 1)dP(x) = n\lambda \int_{\mathbb{R}\setminus\{0\}} (e^{i\theta x} - 1)dP^{(n)}(x)$$

which would ensure that the laws of the processes remain exact. If not, then perhaps

$$\lim_{n\to\infty} \left| \lambda \int_{\mathbb{R}\setminus\{0\}} (e^{i\theta x} - 1)dP(x) - n\lambda \int_{\mathbb{R}\setminus\{0\}} (e^{i\theta x} - 1)dP^{(n)}(x) \right|$$

might serve as a useful measure as to the accuracy of the weak limit as an approximation.

One final point on both the Wiener and α-stable Lévy process approximations. In both convergence arguments, the mean was extracted from the sum. In the Wiener process case

$$\sum_{k=1}^{N_{nt}^{X}} \chi_k^{(n)} = \sum_{k=1}^{N_{nt}^{X}} \left(\chi_k^{(n)} - \frac{\mu}{n}\right) + \mu \frac{N_{nt}^{X}}{n}$$

In the α-stable Lévy process case

$$\sum_{k=1}^{N_{nt}^{X}} \chi_k^{(n)} = \sum_{k=1}^{N_{nt}^{X}} \left(\chi_k^{(n)} - \frac{\mu}{n}\right) + \mu \left(\frac{N_{nt}^{X} - n\lambda t}{n}\right) + \mu\lambda \frac{n}{\varphi(n)}$$

For any n, the components of these decompositions involving the claims and counting process are stochastically dependent. In the limit, $\mu\frac{N_{nt}^{X}}{n} \xrightarrow{\mathbf{P}} \mu\lambda t$ and $\frac{N_{nt}^{X} - n\lambda t}{\varphi(n)} \to 0$ so all stochastic information is "squeezed out", in particular, the stochastic dependence is lost.

4.3 Lévy-Grigelionis-Jacod Characteristics

In the previous section we considered only time homogeneous Poisson models and possible reasons for the poor fit of the corresponding weak approximations were identified. Ultimately we are interested in non-Poisson arrivals as well as non-stationary arrivals and claims distributions.

We saw that for non-Poisson claim arrivals, $M(t)$ is not linear and from (4.2), (4.3), and (4.6) we have

$$\mu^{(n)}(t) = \mu \frac{M(t)}{M(nt)}$$

$$\sigma^{(n)2} = \sigma^2 \frac{M(t)}{M(nt)} + \mu^2 \frac{V(t)}{M(nt)} - \mu^2 \frac{M^2(t)V(nt)}{M^3(nt)}$$

which are time dependent and so moment matching is incompatible with the modelling assumptions. However, since $\frac{M(nt)}{nt} \to \lambda$, $M(nt)$ is asymptotically like n for any t and so $\mu^{(n)}(t) = \mu \frac{M(t)}{M(nt)}$ is asymptotically like $\frac{1}{n}$, as desired. If it were the case that $\mu^{(n)}(t)$ was only mildly time dependent (for example, for each $n \in \mathbb{N}$, $\sup_{t \in \mathbb{R}_+} \mu^{(n)}(t) - \inf_{t \in \mathbb{R}_+} \mu^{(n)}(t)$ or $\sup_{t \in \mathbb{R}_+} \left|\frac{d}{dt}\mu^{(n)}(t)\right|$ are acceptably small) one could set $\bar{\mu}^{(n)} = \lim_{t \to \infty} \frac{1}{t} \int_0^t \mu^{(n)}(s)ds$ or some other time average of $\mu^{(n)}(t)$. Using this time average $\bar{\mu}^{(n)}$ for $\mu^{(n)}(t)$ we can approximate the ratio $\frac{M(t)}{M(nt)}$ by $\frac{\bar{\mu}^{(n)}}{\mu}$ and from (4.3) obtain

$$\sigma^{(n)2} \approx \sigma^2 \frac{\bar{\mu}^{(n)}}{\mu} + \mu^2 \frac{\mathbf{E}[(N_t^\chi)^2]}{M(nt)} - (\bar{\mu}^{(n)})^2 \frac{\mathbf{E}[(N_{nt}^\chi)^2]}{M(nt)}$$

If N^χ has a finite second moment and for each $t \in \mathbb{R}_+$, $\lim_{n \to \infty} \frac{\mathbf{E}[(N_{nt}^\chi)^2]}{nM(nt)}$ exists then $\sigma^{(n)2}$ is asymptotically like $\frac{1}{n}$. Even if $\mu^{(n)} = \mu \frac{M(t)}{M(nt)}$ were free of time, $\sigma^{(n)2}(t)$ may still be time dependent. Requiring that $\frac{M(t)}{M(nt)}$, $\frac{\mathbf{E}[(N_t^\chi)^2]}{M(nt)}$, and $\frac{\mathbf{E}[(N_{nt}^\chi)^2]}{M(nt)}$ be time independent and that $\lim_{n \to \infty} \frac{\mathbf{E}[(N_{nt}^\chi)^2]}{nM(nt)}$ exists may pose significant restrictions on N^χ. For instance, merely requiring that $M(t)$ be linear forces N^χ to be Poisson. These requirements might therefore rule out renewal counting processes which are not well approximated by a Wiener or α-stable Lévy process. One could attempt to match higher moments, resulting in further conditions involving higher moments of N^χ as well as the claims distribution, perhaps providing a finer criterion for the suitability of these approximations. It is also possible that as higher moments are matched, these increasingly complex requirements rule out all processes, indicating that such weak approximations are never appropriate. This time dependence suggests that we may have to weaken our modelling assumptions; rather than insisting that the limit be a SIIP, we drop the requirement of stationarity and look at IIP's. As long as we can construct a sequence $\{R^{(n)}\}_{n\in\mathbf{N}}$ by an infinite subdivision of the original process R which converges to some process $R^{(\infty)}$ then it may be possible to fit $R^{(\infty)}$ to R by matching their Lévy-Grigelionis-Jacod characteristics. Also, dropping stationarity introduces the possibility of weakly approximating non-stationary models as well and it is here that the Lévy-Grigelionis-Jacod characteristics of IIP's may prove very useful.

In the previous section we considered stochastically continuous SIIP's as weak approximations. Let L be a stochastically continuous SIIP; its characteristic function can be written in the form

$$\begin{aligned}\mathbf{E}[e^{i\theta L_t}] &= \exp\left[i\theta\mu t - \frac{1}{2}\theta^2\sigma^2 t + t\int_{\mathbb{R}\setminus\{0\}}(e^{i\theta x} - 1 - i\theta\psi(x))dM(x)\right] \\ &= \exp\left[i\theta B_t - \frac{1}{2}\theta^2 C_t + \int_0^t\int_{\mathbb{R}\setminus\{0\}}(e^{i\theta x} - 1 - i\theta\psi(x))d\nu(s,x)\right]\end{aligned}$$

where

$$\begin{aligned}B_t &= \mu t \\ C_t &= \sigma^2 t \\ d\nu(s,x) &= ds \otimes dM(x)\end{aligned}$$

for some $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}_+$, $M \in \mathcal{M}^{1\wedge x^2}_{\mathbb{R}\setminus\{0\}}$, where ds is Lebesgue measure, and $\psi(x)$ is as in the Lévy-Khintchine Theorem. The triplet (B, C, ν) uniquely determines the process. The stationarity of L is the reason the time derivatives $\dot{B}_t = \mu$ and $\dot{C}_t = \sigma^2$ are constant and the jump measure $d\nu(s,x)$ is a product measure of ds and $dM(x)$; the drift perturbation arrival rate, the variance perturbation arrival rate, the jump arrival rate and jump size distribution are all constant in time.

If one drops the stationarity assumption and replaces it with a "local stationarity" then B_t and C_t may have non-constant time derivatives and $d\nu$ might not be a product measure, reflecting a dependence of jump arrival rate and size on time. The interesting feature of the characteristics is that they appear to be amenable to a statistical fitting procedure. Furthermore, it turns out that B_t, C_t, and $d\nu$ are deterministic if and only if the process is an IIP.

Let X be a stochastically continuous IIP with time index set $[0, T]$ for $T < \infty$. This is true for càdlàg processes since the set of discontinuities is at most countable. For $n \in \mathbb{N}$ let

$$\Pi^{(n)} = \left\{ \{t_k^{(n)}\} : 0 = t_0^{(n)} < \ldots < t_{k^{(n)}}^{(n)} = T, k^{(n)} \in \mathbb{N} \right\}$$

be a sequence of partitions of $[0, T]$ such that $\lim_{n\to\infty} \max_{1 \leqslant k \leqslant k^{(n)}} |t_k^{(n)} - t_{k-1}^{(n)}| = 0$ and $\Pi^{(n)} \subset \Pi^{(n+1)}$, ie, a normal, refining sequence of partitions of $[0, T]$. For the process X, define corresponding sequences of increments by

$$\Delta_k^{(n)} X = X_{t_k^{(n)}} - X_{t_{k-1}^{(n)}} \quad n \in \mathbb{N}, 1 \leqslant k \leqslant k^{(n)}$$

Now, define

$$B_t^{(n)} = \sum_{k:t_k^{(n)} \leqslant t} \mathbf{E}[\psi(\Delta_k^{(n)} X)] \tag{4.16}$$

$$V_t^{(n)} = \sum_{k:t_k^{(n)} \leqslant t} \{\mathbf{E}[\psi(\Delta_k^{(n)} X)]^2 - (\mathbf{E}[\psi(\Delta_k^{(n)} X)])^2\} \tag{4.17}$$

$$Ef_t^{(n)} = \sum_{k:t_k^{(n)} \leqslant t} \mathbf{E}[f(\Delta_k^{(n)} X)] \tag{4.18}$$

where $f : \mathbb{R} \to \mathbb{R}$ is a bounded continuous function that is zero in some neighborhood of zero and $\psi(x)$ is as in the Lévy-Khintchine Theorem.

Proposition 9 *Suppose that* $\{\Pi^{(n)}\}_{n\in\mathbb{N}}$ *is a normal, refining sequence of partitions of* $[0, T]$ *such that* $\cup_{n\in\mathbb{N}}\Pi^{(n)}$ *contains all points of stochastic discontinuity. Then, for* $t \in [0, T]$,

$$B_t = \lim_{n\to\infty} B_t^{(n)} \tag{4.19}$$

$$V_t = \lim_{n\to\infty} V_t^{(n)} \tag{4.20}$$

$$Ef_t = \lim_{n\to\infty} Ef_t^{(n)} \tag{4.21}$$

where the convergence is uniform on $[0, T]$.

Proof: See Kwapień & Woyczyński (1992) or Jacod & Shiryaev (1987).

B_t is the *first characteristic* of X. The *second characteristic* of X is the unique measure ν on $\mathbb{R} \setminus \{0\} \times [0,T]$ such that for any $t \in [0,T]$ and $f : \mathbb{R} \to \mathbb{R}$ continuous, bounded, and zero in a neighborhood of zero,

$$\int_{\mathbb{R}\setminus\{0\}} \int_0^t f(x) d\nu(s,x) = Ef_t = \lim_{n\to\infty} Ef_t^{(n)}$$

And, to define the *third characteristic* of X, we restrict ourselves to IIP's X such that B_t is of bounded variation (see Jacod & Shiryaev (1987) for the general case). In this case we redefine $V_t^{(n)} = \sum_{k:t_k^{(n)} \leqslant t} \mathbf{E}[\psi(\Delta_k^{(n)} X)]^2$. It can be shown that $V_t = \lim_{n\to\infty} V_t^{(n)}$ exists and the convergence is uniform on $[0,T]$. Now, one can define the *third characteristic* of X by

$$C_t = V_t - \int_{\mathbb{R}\setminus\{0\}} \int_0^t \psi(x)^2 d\nu(s,x)$$

These quantities are highly suggestive. One could estimate these quantities from a point process under consideration and perhaps by performing some kind of smoothing one would end up with smooth characteristics yielding a process that fits the point process well and is mathematically tractable. This is a topic for future study.

Characteristics have been significantly generalized to the case of semi-martingales where B, C, and ν are the unique (up to modification) predictable stochastic processes completely characterizing the semi-martingale. These are referred to as the Lévy-Jacod-Grigelionis semi-martingale characteristics. This result as well as many weak convergence results for IIP's and processes with conditionally independent increments are treated extensively in Jacod & Shiryaev (1987) and warrant further study.

The case of semi-martingale characteristics is similar but involves conditioning. Let X be an $F = \{\mathcal{F}_t\}_{t\in[0,T]}$-adapted càdlàg process. Suppose also that X satisfies the following property: $\forall \epsilon > 0\ \exists \delta > 0$ such that for any partition $0 = t_0 < t_1 < \ldots < t_n = T$,

$$\mathbf{P}\left\{ \sum_{k=1}^{(n)} \left| \mathbf{E}[\psi(X_{t_k} - X_{t_{k-1}}) | \mathcal{F}_{t_{k-1}}] \right| > \delta \right\} < \epsilon$$

Define

$$B_t^{(n)} = \sum_{k:t_k^{(n)} \leqslant t} \mathbf{E}[\psi(\Delta_k^{(n)} X) | \mathcal{F}_{t_k^{(n)}}] \tag{4.22}$$

$$V_t^{(n)} = \sum_{k:t_k^{(n)} \leqslant t} \mathbf{E}[\psi(\Delta_k^{(n)} X) | \mathcal{F}_{t_k^{(n)}}]^2 \tag{4.23}$$

$$Ef_t^{(n)} = \sum_{k:t_k^{(n)} \leqslant t} \mathbf{E}[f(\Delta_k^{(n)} X) | \mathcal{F}_{t_k^{(n)}}] \tag{4.24}$$

Proposition 10 *Suppose that $\{\Pi^{(n)}\}_{n \in \mathbf{N}}$ is a normal, refining sequence of partitions of $[0, T]$ such that $\cup_{n \in \mathbf{N}} \Pi^{(n)}$ contains all points of stochastic discontinuity of X and X satisfies condition B. Then, for $t \in [0, T]$,*

$$B_t = \lim_{n \to \infty} B_t^{(n)} \tag{4.25}$$

$$V_t = \lim_{n \to \infty} V_t^{(n)} \tag{4.26}$$

$$Ef_t = \lim_{n \to \infty} Ef_t^{(n)} \tag{4.27}$$

where the convergence is uniform in probability on $[0, T]$.

B_t is the *first characteristic.* The *second characteristic* of is the unique measure ν on $\mathbb{R} \setminus \{0\} \times [0, T]$ such that for any $\omega \in \Omega$, $t \in [0, T]$ and $f : \mathbb{R} \to \mathbb{R}$ continuous, bounded, and zero in a neighborhood of zero,

$$\int_{\mathbb{R}\setminus\{0\}} \int_0^t f(x) d\nu(s, x, \omega) = Ef_t(\omega) = \lim_{n \to \infty} E_f^{(n)}(t)(\omega)$$

And, the *third characteristic* of X is

$$C_t(\omega) = V_t(\omega) - \int_{\mathbb{R}\setminus\{0\}} \int_0^t \psi(x)^2 d\nu(s, x, \omega)$$

A recent result of Sørensen (1996) extends a classical martingale method for computing ruin probabilities to a large class of semi-martingale reserve models. The key property used is the following:

Proposition 11 *For a quasi-left-continuous semimartingale X with characteristics (B, C, ν), the process Y defined by*

$$Y_t(\omega) = \exp\left[i\theta X_t(\omega) - i\theta B_t(\omega) - \frac{1}{2}\theta^2 C_t(\omega) - \int_0^t \int_{\mathbb{R}\setminus\{0\}} (e^{i\theta x} - 1 - i\theta\psi(x))d\nu(s, x, \omega)\right]$$

is a local square-integrable martingale for each θ.

Proof: See Jacod & Shiryaev (1987).

Although there is no discussion of fit in Sørensen (1996), the methods he develops offer the possibility of using more interesting models for reserves. We believe that Lévy-Grigelionis-Jacod characteristics may be very helpful even in the case of random characteristics.

A related topic that warrants further study is the special form that characteristics of stochastic integrals take. The idea here is that a point process reserve model R could be expressed as a stochastic integral of some process H with respect to a compensated point process:

$$R_t = \int_0^t H_s d\tilde{N}_s$$

where $\tilde{N}_t = N_t - \lambda_t$ for some point process N having predictable compensator λ. In this framework, one could consider sequences of processes $\{H^{(n)}\}_{n\in\mathbb{N}}$ and $\{N^{(n)}\}_{n\in\mathbb{N}}$ and examine the convergence

$$\int_0^t H_s^{(n)} d\tilde{N}_s^{(n)} \Rightarrow \int_0^t H_s^{(\infty)} d\tilde{N}_s^{(\infty)}$$

in terms of the nice convergence results for characteristics of stochastic integrals (see Jacod & Shiryaev (1987)).

A particularly interesting possibility is if the limiting stochastic integral is with respect to a Wiener process. In this case, the process $\int_0^t H_s^{(\infty)} d\tilde{N}_s^{(\infty)}$ is a continuous martingale, a class of processes that exhibits a high degree of tractability and for which many results are known (Revuz & Yor 1992). Also, interesting work has appeared recently on the construction of strong solutions to stochastic differential equations driven by a compensated Poisson process which may enable the explicit calculation or efficient simulations of ruin probabilities (Ruiz de Chávez, Léon & Tudor 1996). Finally, the recent work on multilevel, bilinear stochastic differential equations, whose weak limits under various space/time rescalings are measure-valued processes (Dawson 1993), offer the possibility of modelling coupled claims and income processes, as required in the general model introduced in Section (2.1). Particularly of interest in this case is where the multilevel, bilinear system of SDE's is driven by point processes.

Bibliography

Asmussen, S. (1984), 'Approximations for the probability of ruin within finite time', *Scandanavian Actuarial Journal* pp. 31–57.

Bernouilli, D. (1738), 'Specimen theoriae novae de mensura sortis', *Commentarii Academiae Scientarum Imperialis Petropolitanae* **5**, 175–192. [translation : Exposition of a New Theory on the Measurement Of Risk, 1954, *Econometrica*, **22**, pp. 22-36].

Billingsley, P. (1968), *Convergence of Probability Measures*, John Wiley and Sons, New York.

Billingsley, P. (1995), *Probability and Measure*, 3rd ed., John Wiley and Sons, New York.

Chambers, J., Mallows, C. & Stuck, B. (1976), 'A Method for Simulating Stable Random Variables', *Journal of the American Statistical Association* **71**, 340–344.

Chukova, S., Dimitrov, B. & Garrido, J. (1993), 'Renewal and nonhomogeneous Poisson processes generated by distributions with periodic failure rate', *Statistics and Probability Letters* **17**, 19–25.

Cramér, H. (1930), 'On the Mathematical Theory of Risk', *Jubilee Volume of Försäkringsakliebolaget Skandia, Nordiska Bokhandeln* .

Darling, D. & Siegert, A. (1953), 'The first passage problem for a continuous Markov process', *Annals of Mathematical Statistics* **24**, 624–639.

Davidson, J. (1994), *Stochastic Limit Theory*, Oxford University Press, New York.

Dawson, D. (1993), 'Measure Valued Markov Processes', *Lecture Notes in Mathematics* **1541**.

DeMoivre, A. (1738), *The Doctrine of Chances*, London.

Doob, J. L. (1953), *Stochastic Processes*, John Wiley and Sons, New York.

Elliot, R. (1982), *Stochastic Calculus and Applications*, Springer-Verlag, New York.

Embrechts, P. & Veraverbeeke, N. (1982), 'Estimates for the probability of ruin with emphasis on the possibility of large claims', *Insurance: Mathematics and Economics* pp. 55–72.

Ethier, S. & Kurtz, T. (1986), *Markov Processes - Characterization and convergence*, John Wiley and Sons, New York.

Feller, W. (1966), *An Introduction to Probability Theory and its Applications*, Vol. II, John Wiley and Sons, New York.

Folland, G. (1984), *Real Analysis-Modern Techniques and their Applications*, John Wiley and Sons, New York.

Furrer, H. (1996), *private communication* .

Furrer, H., Michna, Z. & Weron, A. (1996), 'Stable Lévy Motion Approximation in Collective Risk Theory', *to appear in Insurance : Mathematics & Economics* .

Garrido, J. (1987), Diffusion Models for Risk Processes with Interest and Inflation, PhD thesis, University of Waterloo, Waterloo, Ontario.

Gerber, H. (1979), *An Introduction to Mathematical Risk Theory*, S.S. Huebner Foundation, Univeristy of Pennsylvania.

Gluckman, P. (1979), Applications of diffusion approximations to the collective theory of risk, PhD thesis, Standford University.

Grandell, J. (1991), *Aspects of Risk Theory*, Springer-Verlag, New York.

Hida, T. (1980), *Brownian Motion*, Springer-Verlag, New York.

Hogg, R. & Klugman, S. (1984), *Loss Distributions*, Wiley, New York.

Iglehart, D. (1969), 'Diffusion Approximations in Collective Risk Theory', *Journal of Applied Probability* **6**, 285–292.

Ikeda, N. & Watananbe, S. (1989), *Stochastic Differential Equations and Diffusion Processes*, 2nd ed., Elsevier Publishing Company, New York.

Jacod, J. & Shiryaev, A. (1987), *Limit Theorems for Stochastic Processes*, Springer-Verlag, Berlin.

Janicki, A. & Weron, A. (1994), *Simulation and Chaotic Behaviour of α-Stable Stochastic Processes*, Marcel Dekker, New York.

Janssen, J. (1981), 'Generalized Risk Models', *Cahier du C.E.R.O., Bruxelles,* **23**, 225–243.

Kallianpur, G. (1980), *Stochastic Filtering Theory*, Springer-Verlag, New York.

Kanter, M. (1975), 'Stable densities under change of scale and total variation inequalities', *Annals of Probability* **31**, 697–707.

Karlin, S. & Taylor, H. (1981), *A Second Course in Stochastic Processes*, Academic Press, New York.

Kloeden, P. & Platen, E. (1992), *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin.

Kolmogorov, A. & Gnedenko, B. (1968), *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley, New York.

Kwapień, S. & Woyczyński, W. (1992), *Random Series & Stochastic Integrals: Single & Multiple*, Birkhäuser, Boston.

Lindvall, T. (1973), 'Weak convergence of probability measures and random functions in the space $D[0,\infty)$', *Journal of Applied Probability* **10**, 109–121.

Linz, P. (1985), *Analytical & Numerical Method for Volterra Equations*, Studies in Applied Mathematics, SIAM.

Mijnheer, J. (1975), 'Sample path properties of stable processes', *Mathematical Centre Tracts* **59**.

Milstein, G. (1995), *Numerical Integration of Stochastic Differential Equations*, Kluwer Academic Publishers, London.

Pollard, D. (1984), *Convergence of Stochastic Processes*, Springer-Verlag, New York.

Prohorov, Y. (1956), 'Convergence of random processes and limit theorems in probability theory', *Theor. Probability Appl.* **1**, 157–214. English translation.

Protter, P. (1990), *Stochastic Integration and Differential Equations*, Springer-Verlag, New York.

Ramsay, C. & Usabel, M. (1997), 'Calculating ruin probabilities via product integration', *ASTIN Bulletin: to appear* .

Revuz, D. & Yor, M. (1992), *Continuous Martingales and Brownian Motion*, 2nd ed., Springer-Verlag, Berlin.

Ruiz de Chávez, J., Léon, J. & Tudor, C. (1996), 'Strong solution of anticipating stochastic differential equations on the poisson space', *Bol. Soc. Mat. Mexicana* **2**(3), 55–63.

Samorodnitsky, G. & Taqqu, M. (1994), *Stable Non-gaussian Random Processes*, Chapman and Hall, New York.

Seal, H. (1983), 'The poisson process: its failure in risk theory', *Insurance: Mathematics & Economics* **2**, 287–288.

Skorokhod, A. (1956), 'Limit theorems for stochastic processes with independent increments', *Theory of Probability and its Applications* **2**, 138–171.

Sørensen, M. (1996), 'A Semimartingale Approach To Some Problems In Risk Theory', *ASTIN Bulletin* **1**(26), 15–23.

Stone, C. (1963), 'Weak convergence of stochastic processes defined on semi-infinite time intervals', *Proc. Amer. Math. Soc.* **14**, 694–696.

Stroock, D. (1993), *Probability: An Analytic View*, Cambridge University Press, Cambridge.

Stroock, D. & Varadhan, S. (1979), *Multidimensional Diffusion Processes*, Springer-Verlag, New York.

Zolotarev, V. (1986), *One-dimension Stable Distributions*, Vol. 65 of *Translations of Mathematical Monographs*, American Mathematical Society, Providence, Rhode Island.

150mm

6"

APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989