CHAPTER ONE

Motivation for Heavy-Tailed Models

1.1 Structure of the Book

This book is split into a few core components covering fundamental concepts:

- Chapter 1 motivates the need to consider heavy-tailed loss process models in operational risk (OpRisk) and insurance modeling frameworks. It provides a basic introduction to the concept of separating the modeling of the central loss process and the tails of the loss process through splice models. It also sets out the key statistical questions one must consider studying when performing analysis and modeling of high consequence rare-event loss processes.
- Chapter 2 covers all the fundamental properties one may require in univariate loss process modeling under an extreme value theory (EVT) approach. Of importance is the detailed discussion on the associated statistical assumptions that must be made regarding the properties of any data utilized in model estimation when working with EVT models. This chapter provides a relatively advanced coverage of generalized extreme value (GEV) family of models, block maximum and peaks over threshold frameworks. It provides detailed discussion on statistical estimation that should be utilized in practice for such models and how one may adapt such methods to small sample settings that may arise in OpRisk settings. In the process, the chapter details clearly how to construct several loss distributional approach models based on EVT analysis. It then concludes with results of EVT in the context of compound processes.
- Chapter 3 provides a set of formal mathematical definitions for different notions regarding a heavy-tailed or fat-tailed loss distribution and its properties. It is important that when modeling such loss processes, especially the asymptotic properties of compound process

Advances in Heavy Tailed Risk Modeling: A Handbook of Operational Risk,
First Edition. Gareth W. Peters and Pavel V. Shevchenko.

models built with heavy-tailed loss models, a clear understanding of the tail properties of such loss models is understood. In this regard, we discuss the family of sub-exponential loss models, the family of regularly varying and slowly varying models. There are within these large classes of models sub-categorizations that are often of use to understand when thinking about risk measures resulting from such loss models, these are also detailed, for example, long-tailed models, subversively varying models and extended regular variation. In addition, the chapter opens with a basic introduction to key notations and properties of asymptotic notations that are utilized throughout the book.

- Chapter 4 begins with a basic introduction to properties of mathematical representations and characterizations of heavy-tailed loss models through the characteristic function and its representation. It then details the notions of divisibility, self-decomposability and the resulting consequences such distributional properties have on loss distributional approach compound process models. The remainder of the chapter provides a detailed coverage of the family of univariate α-stable models, detailing their characterization, the parameterizations, density and distribution representations and parameter estimation. Such a family of models is becoming increasingly interesting for OpRisk modeling and insurance. It is recognized that such a family of models possesses many relevant and useful features that will capture aspects of OpRisk and insurance loss processes accurately and with advantageous features when used in a compound process model under a loss distributional approach structure.
- Chapter 5 provides the representations of flexible severity models based on tempering or exponential tilting of the α-stable family of loss models. Under this concept, there are many families of tempered stable models available; this chapter characterizes each and discusses the mathematical properties of each sub-class of models and how they may be used in compound process models for heavy-tailed loss models in OpRisk and insurance. In addition, it discusses the aspects of model estimation and simulation for such models. The chapter then finishes with a detailed discussion on quantile-transformed-based heavy-tailed loss models for OpRisk and insurance, such as the Tukey transforms and the sub-family of the g-and-h distributions that have been popular in OpRisk.
- Chapter 6 discusses compound processes and convolutional semi-group structures. This then leads to developing representations of closed-form compound process loss distributions and densities that admit heavy-tailed loss processes. The chapter characterizes several classes of such models that can be used in practice, which avoid the need for computationally costly Monte Carlo simulation when working with such models.
- Chapter 7 discusses many properties of different classes of heavy-tailed loss processes with regard to asymptotic representations and properties of the tail of both partial sums and compound random sums. It does so under first-, second- and third-order asymptotic expansions for the tail process of such heavy-tailed loss processes. This is achieved under many different assumptions relating to the frequency and severity distribution and the possible dependence structures in such loss processes.
- Chapter 8 extends the results of the asymptotics for the tail of heavy-tailed loss processes partial sums and compound random sums to the asymptotics of risk measures developed from such loss processes. In particular, it discusses closed-form single-loss approximations and first-order, second-order and higher order expansion representations. It covers value-at-risk, expected shortfall and spectral risk measure asymptotics. This chapter also covers some alternative risk measure asymptotic results based on EVT known as *penultimate approximations*.

- Chapter 9 rounds off the book with a coverage of numerical simulation and estimation procedures for rare-event simulations in heavy-tailed loss processes, primarily for the estimation of properties of risk measures that provides an efficient numerical alternative procedure to utilization of such asymptotic closed-form representations.

1.2 Dominance of the Heaviest Tail Risks

In this book, we develop and discuss models for OpRisk to better understand statistical properties and capital frameworks which incorporate risk classes in which infrequent, though catastrophic or high consequence loss events may occur. This is particularly relevant in OpRisk as can be illustrated by the historical events which demonstrate just how significant the appropriate modeling of OpRisk can be to a financial institution.

Examples of large recent OpRisk losses are

- J.P. Morgan, GBP 3760 million in 2013—US authorities demand money because of mis-sold securities to Fannie Mae and Freddie Mac;
- Madoff and investors, GBP 40,819 million in 2008—B. Madoff's Ponzi scheme;
- Société Générale, GBP 4548 million in 2008—a trader entered futures positions circumventing internal regulations.

Other well-known examples of OpRisk-related events include the 1995 Barings Bank loss of around GBP 1.3 billion; the 2001 Enron loss of around USD 2.2 billion and the 2004 National Australia Bank loss of AUD 360 m.

The impact that such significant losses have had on the financial industry and its perceived stability combined with the Basel II regulatory requirements BCBS (2006) have significantly changed the view that financial institutions have regarding OpRisk. Under the three pillars of the Basel II framework, internationally active banks are required to set aside capital reserves against risk, to implement risk management frameworks and processes for their continual review and to adhere to certain disclosure requirements. There are three broad approaches that a bank may use to calculate its minimal capital reserve, as specified in the first Basel II pillar. They are known as *basic indicator approach*, *standardized approach* and *advanced measurement approach* (AMA) discussed in detail in Cruz *et al.* (2015). AMA is of interest here because it is the most advanced framework with regards to statistical modeling. A bank adopting the AMA must develop a comprehensive internal risk quantification system. This approach is the most flexible from a quantitative perspective, as banks may use a variety of methods and models, which they believe are most suitable for their operating environment and culture, provided they can convince the local regulator (BCBS 2006, pp. 150–152). The key quantitative criterion is that a bank's models must sufficiently account for potentially high impact rare events. The most widely used approach for AMA is loss distribution approach (LDA) that involves modeling the severity and frequency distributions over a predetermined time horizon so that the overall loss Z of a risk over this time period (e.g. year) is

$$Z = X_1 + \cdots + X_N, \tag{1.1}$$

where N is the frequency modeled by random variable from discrete distribution and $X_1, X_2, \ldots$ the independent severities from continuous distribution $F_X(x)$. There are many important aspects of LDA such as estimation of frequency and severity distributions using data and expert judgements or modeling dependence between risks considered in detail in Cruz *et al.* (2015). In this book, we focus on modeling heavy-tailed severities.

Whilst many OpRisk events occur frequently and with low impact (indeed, are 'expected losses'), others are rare and their impact may be as extreme as the total collapse of the bank. The modeling and development of methodology to capture, classify and understand properties of operational losses is a new research area in the banking and finance sector. These rare losses are often referred to as *low frequency/high severity risks*. It is recognized that these risks have heavy-tailed (sub-exponential) severity distributions, that is, the distribution with the tail decaying to zero slower than any exponential.

In practice, heavy-tailed loss distribution typically means that the observed losses are ranging over several orders of magnitude, even for relatively small datasets. One of the main properties of heavy-tailed distributions is that if $X_1, \dots, X_n$ are independent random variables from common heavy-tailed distribution $F(x)$, then

$$\lim_{x\to\infty} \frac{\mathbb{P}\mathrm{r}[X_1 + \cdots + X_n > x]}{\mathbb{P}\mathrm{r}[\max(X_1, \dots, X_n) > x]} = 1. \tag{1.2}$$

This means that the tail of the sum of the random variables has the same order of magnitude as the tail of the maximum of these random variables, with interpretation that severe overall loss is due to a single large loss rather than due to accumulated small losses.

In OpRisk and insurance, we are often interested in the tail of distribution for the overall loss over a predetermined time horizon $Z = X_1 + \cdots + X_N$. In this case, if $X_1, X_2, \dots$ are independent severities from heavy-tailed distribution $F_X(x)$ and

$$\sum_{n=0}^{\infty} (1+\varepsilon)^n \mathbb{P}\mathrm{r}[N = n] < \infty$$

for some $\varepsilon > 0$ (which is satisfied, e.g. for Poisson and negative binomial distributions), then

$$1 - F_Z(x) = \mathbb{P}\mathrm{r}[X_1 + \cdots + X_N > x] \sim \mathbb{E}[N](1 - F_X(x)), \quad x \to \infty. \tag{1.3}$$

This can be used to approximate high quantiles of the distribution of Z as

$$F_Z^{-1}(q) \sim F_X^{-1}\left(1 - \frac{1-q}{\mathbb{E}[N]}\right), \quad q \to 1, \tag{1.4}$$

where q is the quantile level. This approximation is often referred to as the *single-loss approximation* because the compound distribution is expressed in terms of the single-loss distribution.

Heavy-tailed distributions include many well-known distributions. For example, the LogNormal distribution is heavy tailed. An important class of heavy-tailed distributions is the so-called regular varying tail distributions (often referred to as *power laws* or *Pareto distributions*)

$$1 - F(x) = x^{-\alpha} C(x), \quad x \to \infty,\ \alpha \geq 0, \tag{1.5}$$

where α is the so-called power tail index and $C(x)$ the slowly varying function that satisfies

$$\lim_{x\to\infty} C(tx)/C(x) = 1, \text{ for } t > 0. \tag{1.6}$$

Often, sub-exponential distributions provide a good fit to the real datasets of OpRisk and insurance losses. However, corresponding datasets are typically small and the estimation of these distributions is a difficult task with a large uncertainly in the estimates.

Remark 1.1 *From the perspective of capital calculation, the most important processes to model accurately are those which have relatively infrequent losses. However, when these losses do occur, they are distributed as a very heavy-tailed severity distribution such as members of the sub-exponential family. Therefore, the intention of this book is to present families of models suitable for such severity distribution modeling as well as their properties and estimators for the parameters that specify these models.*

The precise definition and properties of the heavy-tailed distributions is a subject of Chapter 3, and single-loss approximation is discussed in detail in Chapters 7 and 8. For a methodological insight, consider J independent risks, where each risk is modeled by a compound Poisson. Then, the sum of risks is a compound Poisson with the intensity and severity distribution given by the following proposition.

Proposition 1.1 *Consider J independent compound Poisson random variables*

$$Z^{(j)} = \sum_{s=1}^{N^{(j)}} X_s^{(j)}, \quad j = 1, \ldots, J, \tag{1.7}$$

where the frequencies $N^{(j)} \sim Poisson(\lambda_j)$ and the severities $X_s^{(j)} \sim F_j(x)$, $j = 1, \ldots, J$ and $s = 1, 2, \ldots$ are all independent. Then, the sum $Z = \sum_{j=1}^{J} Z^{(j)}$ is a compound Poisson random variable with the frequency distribution $Poisson(\lambda)$ and the severity distribution

$$F(x) = \sum_{j=1}^{J} \frac{\lambda_j}{\lambda} F_j(x),$$

where $\lambda = \lambda_1 + \cdots + \lambda_J$.

The proof is simple and can be found, for example, in Shevchenko (2011, proposition 7.1). Suppose that all severity distributions $F_j(x)$ are heavy tailed, that is,

$$\overline{F}_j(x) = x^{-\alpha_j} C_j(x),$$

where $\alpha_1 < \cdots < \alpha_J$ and $C_j(x)$ are the slowly varying functions as defined in Equation 1.6. Then, $F(x) = \sum_{j=1}^{J} (\lambda_j/\lambda) F_j(x)$ is a heavy-tailed distribution too, with the tail index α_1 for $x \to \infty$. Thus, using the result (Equation 1.3) for heavy-tailed distributions, we obtain that

$$\lim_{x\to\infty} \frac{\mathbb{P}\mathrm{r}[Z > x]}{1 - F_1(x)} = \lambda_1. \tag{1.8}$$

This means that high quantiles of the total loss are due to the high losses of the risk with the heaviest tail. For illustration of this phenomenon with the real data from ORX database, see Cope *et al.* (2009). In their example, $LogNormal$ $(\mu = 8, \sigma = 2.24)$ gave a good fit for 10 business lines with average 100 losses per year in each line using 10,000 observations. The estimated capital across these 10 business lines was Euro 634 million with 95% confidence interval (uncertainty in the capital estimate due to finite data size) of width Euro 98 million. Then, extra risk cell (corresponding to the "clients, products & business practices" event type in the 'corporate finance' business line) was added with one loss per year on an average and the $LogNormal(\mu = 9.67, \sigma = 3.83)$ severity estimated using 300 data points. The obtained estimate for the capital over the 10 business units plus the additional one was Euro 5260 million with 95% confidence interval of the width Euro 19 billion. This shows that one high severity risk cell contributes 88% to the capital estimate and 99.5% to the uncertainty range. In this example, the high severity unit accounts for 0.1% of the bank's losses.

Another important topic in modeling large losses is EVT that allows to extrapolate to losses beyond those historically observed and estimate their probability. There are two types of EVT: *block maxima* and *threshold exceedances*; both are considered in Chapter 2. EVT block maxima are focused on modeling the largest loss per time period of interest. Modeling of all large losses exceeding a large threshold is dealt by EVT threshold exceedances. The key result of EVT is that the largest losses or losses exceeding a large threshold can be approximated by some limiting distributions which are the same regardless of the underlying process. This allows to extrapolate to losses beyond those historically observed. However, EVT is an asymptotic theory. Whether the conditions validating the use of the asymptotic theory are satisfied is often a difficult question to answer. The convergence of some parametric models to EVT regime is very slow. In general, it should not preclude the use of other parametric distributions. In Chapter 4, we consider many useful flexible parametric heavy-tailed distributions.

It is important to mention that empirical data analysis for OpRisk often indicates stability of an infinite mean model for some risk cells (e.g. see Moscadelli (2004)), that is, the severity distribution is a Pareto-type distribution (Equation 1.5) with $0 < \alpha \leq 1$ that has infinite mean. For a discussion about infinite mean models in OpRisk, see discussions in Nešlehová *et al.* (2006). Often, practitioners question this type of model and apply different techniques such as truncation from the above but then the high quantiles become highly dependent on the cut-off level. Typically, the estimates of high quantiles for fat-tailed risks have a very large uncertainty and the overall analysis is less conclusive than in the case of thin-tailed risks; however, it is not the reason to avoid these models if the data analysis points to heavy-tailed behaviour. Recent experience of large losses in OpRisk, when one large loss may lead to the bankruptcy, certainly highlights the importance of the fat-tailed models.

1.3 Empirical Analysis Justifying Heavy-Tailed Loss Models in OpRisk

There are several well-known published empirical studies of OpRisk data such as Moscadelli (2004) analysing 2002 Loss Data Collection Exercise (LDCE) survey data across 89 banks from 19 countries; Dutta & Perry (2006) analysing 2004 LDCE for US banks and Lu & Guo (2013) analysing data in Chinese banks.

- Moscadelli (2004) analysed 2002 Loss Data Collection Exercise (LDCE) survey data with more than 47,000 observations across 89 banks from 19 countries in Europe, North and South Americas, Asia and Australasia. The data were mapped to the Basel II standard eight business lines and seven event types. To model severity distribution, this study considered generalized Pareto distribution (EVT distribution for threshold exceedances in the limit of large threshold) and many standard two-parameter distributions such as gamma, exponential, Gumbel and LogNormal. The analysis showed that EVT explains the tail behaviour of OpRisk data well.

- Dutta & Perry's (2006) study of US banking institutions considered the 2004 LDCE survey data and narrowed down the number of suitable candidate datasets from all institutions surveyed to just seven institutions for which it was deemed sufficient numbers of reported losses were acquired. The somewhat heuristic selection criterion that the authors utilized was that a total of at least 1,000 reported total losses were required and, in addition, each institution was required to have consistent and coherent risk profiles relative to each other, which would cover a range of business types and risk types as well as asset sizes for the institutions.
- Feng *et al.*'s (2012) study on the Chinese banking sector utilized less reliable data sources for loss data of Chinese commercial banks collected through the national media covering 1990–2010. In the process collecting data for banks which include the 4 major state-owned commercial banks (SOCBs), 9 joint-stock commercial banks (JSCBs), 35 city commercial banks (CCBs), 74 urban and rural credit cooperatives (URCCs) and 13 China Postal Savings subsidiaries (CPS). The authors also note that the highest single OpRisk loss amount is up to 7.4 billion yuan, whereas the lowest amount is 50,000 yuan. In addition, losses measured in foreign currency were converted back via the real exchange rate when the loss occurred to convert it to the equivalent amount in yuan. Details of the incidence bank, incidence bank location, type of OpRisk loss, amount of loss, incident time and time span and the sources of OpRisk events were noted.

In the following, we focus on the study of Dutta & Perry (2006), where the authors explored a number of key statistical questions relating to the modeling of OpRisk data in practical banking settings. As noted, a key concern for banks and financial institutions, when designing an LDA model, is the choice of model to use for modeling the severity (dollar value) of operational losses. In addition, a key concern for regulatory authorities is the question of whether institutions using different severity modeling techniques can arrive at very different (and inconsistent) estimates of their exposure. They found, not surprisingly, that using different models for the same institution can result in materially different capital estimates. However, on the more promising side for LDA modeling in OpRisk, they found that there are some models that yield consistent and plausible results for different institutions even when their data differs in some core characteristics related to the collection processes. This suggests that OpRisk data displays some regularity across institutions which can be modeled. In this analysis, the authors noted that they were careful to consider both the modeling of aggregate data at the enterprise level, which would group losses from different business lines and risk types and modeling the attributes of the individual business line and risk types under the recommended business lines of Basel II/Basel III.

On the basis of data collected from seven institutions, with each institution selected as it had at least 1,000 loss events in total, and the data was part of the 2004 LDCE, they performed a detailed statistical study of attributes of the data and flexible distributional models that could be considered for OpRisk models. On the basis of these seven data sources, over a range of different business units and risk types, they found that fitting all of the various datasets one would need to use a model that is flexible enough in its structure. Dutta & Perry (2006) considered modeling via several different means: parametric distributions, EVT models and non-parametric empirical models.

The study focused on models considered by financial institutions in Quantitative Impact Study 4 (QIS-4) submissions, which included one-, two- and four-parameter models. The one- and two-parameter distributions for the severity models included exponential, gamma, generalized Pareto, LogLogistic, truncated LogNormal and Weibull. The four-parameter distributions include the generalized Beta distribution of second kind (GB2) and the g-and-h distribution. These models were also considered in Peters & Sisson (2006a) for modeling severity models in OpRisk under a Bayesian framework.

Dutta & Perry (2006) discussed the importance of fitting distributions that are flexible but appropriate for the accurate modeling of OpRisk data; they focussed on the following five simple attributes in deciding on a suitable statistical model for the severity distribution.

1. *Good Fit*. Statistically, how well does the model fit the data?
2. *Realistic*. If a model fits well in a statistical sense, does it generate a loss distribution with a realistic capital estimate?
3. *Well Specified*. Are the characteristics of the fitted data similar to the loss data and logically consistent?
4. *Flexible*. How well is the model able to reasonably accommodate a wide variety of empirical loss data?
5. *Simple*. Is the model easy to apply in practice, and is it easy to generate random numbers for the purposes of loss simulation?

Their criterion was to regard any technique that is rejected as a poor statistical fit for the majority of institutions to be inferior for modeling OpRisk. The reason for this consideration was related to their desire to investigate the ability to find aspects of uniformity or universality in the OpRisk loss process that they studied. From the analysis undertaken, they concluded that such an approach would suggest OpRisk can be modeled, and there is regularity in the loss data across institutions. Whilst this approach combined elements of expert judgement and statistical hypothesis testing, it was partially heuristic and not the most formal statistical approach to address such problems. However, it does represent a plausible attempt given the limited data sources and resources as well as competing constraints mentioned in the measurement criterion they considered.

We note that an alternative purely statistical approach to such model selection processes was proposed in OpRisk modeling in the work of Peters & Sisson (2006a), whose approach to model selection was to consider a Bayesian model selection based on Bayesian methodology of the Bayes factor and information criterion for penalized model selection such as the Bayesian information criterion.

In either approach, it is generally acknowledged that accurate model selection of an appropriate severity model is paramount to appropriate modeling of the loss processes and, therefore, to the accurate estimation of capital.

Returning to the findings from the seven sources of OpRisk data studied in Dutta & Perry (2006), they found that the exponential, gamma and Weibull distributions are rejected as good fits to the loss data for virtually all institutions at the enterprise, business line and event type levels. This was decided based on formal one sample statistical goodness of fit tests for these models.

When considering the g-and-h distribution, they did not perform the standard hypothesis test for goodness of fit instead opting for a comparison of quantile–quantile (Q–Q) plots and diagnostics based on the five criteria posed above. In all the situations, they found that the g-and-h distribution fits as well as other distributions on the Q–Q plot. The next most preferred distributions were the GB2, LogLogistic, truncated LogNormal and generalized Pareto models, indicating the importance of considering flexible severity loss models. However, only g-and-h distribution resulted in realistic and consistent capital estimates across all seven institutions. In addition, they noted that the EVT models fitted under an EVT threshold exceedances framework were also generally suitable fits for the tails, consistent with the discussions and findings in Lu & Guo (2013) for OpRisk data in the Chinese banking sector and with the results in Moscadelli (2004) analysing 2002 LDCE.

1.4 Motivating Parametric, Spliced and Non-Parametric Severity Models

In this section, we discuss the different approaches that have been adopted in the literature to model aspects of heavy-tailed loss processes. Primarily we focus on the modeling of the severity process in an OpRisk LDA framework; however, we note that many of these approaches can also be adopted for modeling of the annual loss process should sufficient data be available. Before discussing these approaches, it is important to understand some of the basic implications associated with subscribing to such modeling frameworks. We detail two of the most fundamental of these in the following.

Basic Statistical Assumptions to be Considered in Practice

1. It is typical from the statistical perspective to apply the models to be discussed later on the proviso that the underlying process under consideration is actually arising from a single physical process responsible for the losses to be observed. However, in practice, several authors have discussed the impracticality of such assumptions in real-world financial environments, which unlike their physical extreme process counterparts often studied in atmospheric science, hydrology and meteorology, such financial processes are difficult to attribute to a fundamental single 'physical' driving force. Discussion on such issues and their consequences to the suitability of such modeling approaches is provided in Cope *et al.* (2009) and Chavez-Demoulin *et al.* (2006).
2. The other typical statistical assumption that will have potential consequences to application of such modeling paradigms to be discussed later relates to the assumptions made on the temporal characteristics of the underlying loss process driving the heavy-tailed behaviour. In most modeling frameworks discussed later, the parameters causing the loss process will typically be considered unknown but static over time. However, it is likely that in dynamically evolving commercial environments in which financial institutions, disappear, appear and merge on a global scale, whilst regulation continually adapts to the corporate and political landscape, such loss processes driving the heavy-tailed behaviour may not have parameters which are static over time. For example, it is common that under severe losses from an event such as rogue trading, one would typically see the financial institution involved take significant measures to modify the process with the aim to prevent such losses in the same manner again in future, by changing the financial controls, policies and regulatory oversight. This has practical consequences for the ability to satisfy the typical statistical assumptions one would like to adopt with such heavy-tailed models.
3. Typically, the application and development of theoretical properties of the models to be developed, including the classical estimators developed for the parameters of such models under either a frequentist or a Bayesian modeling paradigm, revolve around the assumption that the losses observed are independent and identically distributed. Again,

several authors have developed frameworks motivating the necessity to capture dependence features adequately in OpRisk and insurance modeling of heavy-tailed data, see Böcker & Klüppelberg (2008), Chavez-Demoulin *et al.* (2006) and Peters *et al.* (2009a). In practice, the models presented later can be adapted to incorporate dependence, once a fundamental understanding of their properties and representations is understood for the independently and identically distributed (i.i.d.) cases and this is an active field of research in OpRisk at present.

4. Finally, there is also, typically for several high consequence loss processes, a potential upper limit of the total loss that may be experienced by such a loss process. Again, this is practically important to consider before developing such models to be presented.

The actuarial literature has undertaken several approaches to attempt to address aspects of modeling when such assumptions are violated. For example, should one believe that the underlying risk process is a consequence of multiple driving exposures and processes, then it is common to develop what are known as mixture loss processes. Where if one can identify key loss processes that are combining to create the observed loss process in the OpRisk framework under study, then fitting a mixture model in which there is one component per driving process (potentially with different heavy-tailed features) is a possibility. Another approach that can be adopted and we discuss in some detail throughout next section is the method known as *splicing*. In such a framework, a flexible severity distribution is created, which aims to account for two or more driving processes that give rise to the observed loss process. This is achieved under a splicing framework under the consideration that the loss processes combining to create the observed process actually may differ significantly in the amounts of losses they generate and also in OpRisk perhaps in the frequency at which these losses are observed. Therefore, a splicing approach adopts different models for particular intervals of the observed loss magnitudes. Therefore, small losses may be modeled by one parametric model over a particular interval of loss magnitudes and large severe losses captured by a second model fitted directly to the losses observed in the adjacent loss magnitude partition of the loss domain. These will be discussed in some detail in the following chapter.

In general, it is a serious challenge for the risk managers in practice to try to reconcile such assumptions into a consistent, robust and defensible modeling framework. Therefore, we proceed with an understanding that such assumptions may not all be satisfied jointly under any given model when developing the frameworks to be discussed later. However, in several cases, the models we will present will in many respects provide a conservative modeling framework for OpRisk regulatory reporting and capital estimation should these assumptions be violated as discussed earlier.

Statistical Modeling Approaches to Heavy-Tailed Loss Processes:

The five basic statistical approaches to modeling the severity distribution for a single-loss process that will be considered throughout this book are:

1. EVT methods for modeling explicitly the tail behaviour of the severity distribution in the loss process: ‘block maxima’ and ‘points over threshold’ models.
2. Spliced parametric distributional models combining exponential family members with EVT model tail representations: mixtures and composite distributions.
3. Spliced non-parametric kernel density estimators with EVT tail representations.
4. Flexible parametric models for the entire severity distribution considered from sub-exponential family members: α-stable, tempered and generalized tempered α-stable,

generalized hypergeometric (normal inverse Gaussian), GB2, generalized Champernowne and quantile distributions (g-and-h).

5. Spliced parametric distributional models examples combining exponential family members with sub-exponential family parametric models.

As is evident from the survey of different approaches to modeling heavy-tailed loss processes, mentioned earlier, there is a large variety of models and techniques developed to study and understand such important phenomena as heavy-tailed processes. In the context of OpRisk, the consequences of failing to model adequately the possible heavy-tailed behaviour of certain OpRisk loss processes could result in significant under estimation of the required capital to guard against realizations of such losses in a commercial banking environment and the subsequent failure or insolvency of the institution.

1.5 Creating Flexible Heavy-Tailed Models via Splicing

In this section, we briefly detail the basic approaches to create a spliced distribution and the motivation for such models. These will then be significantly elaborated in the proceeding models when they are incorporated with various modeling approaches to capture heavy-tailed behaviour of a loss process.

It is common in practice for actuarial scientist and risk managers to consider the class of flexible distributional models known as *spliced distributions*. In fact, there are standard packages implemented in several widely utilized software platforms for statistical and risk modeling that incorporate at least basic features of spliced models. The basic k-component spliced distribution as presented in Klugman *et al.* (1998, section 5.2.6) is defined according to Definition 1.1.

Definition 1.1 (Spliced Distribution) *A random variable $X \in \mathbb{R}^+$ representing the loss of a particular risk process can be modeled by a k-component spliced distribution, defined according to the density function partitioned over the loss magnitudes according to the intervals $\cup_{i=1}^{k}[x_{i-1}, x_i) = \mathbb{R}^+$ and given by*

$$f_X(x) = \begin{cases} w_1\, f_1(x), & 0 \le x < x_1, \\ w_2\, f_2(x), & x_1 \le x < x_2, \\ \vdots & \\ w_{k-1}\, f_{k-1}(x), & x_{k-2} \le x < x_{k-1}, \\ w_k\, f_k(x), & x_{k-1} \le x < \infty, \end{cases} \tag{1.9}$$

where the weight parameters $w_i \ge 0$, $i = 1, \dots, k$ satisfy $w_1 + \cdots + w_k = 1$, and $f_1(x), \dots, f_k(x)$ are proper density functions, that is, $\int f_i(x)\,dx = 1$, $i = 1, \dots, k$.

To illustrate this, consider the typically applied model involving the choice of $k = 2$ in which the loss processes have loss magnitudes which are partitioned into two regions $[0, x_{\min}) \cup [x_{\min}, \infty)$. The interpretation being that two driving processes give rise to the risk processes under study. Less frequent but more severe loss processes would typically experience losses exceeding $x_{\min}$. Therefore, we may utilise a lighter tailed parametric model $f_1(x)$ in the region $[0, x_{\min})$ and an associated normalization for the truncation of the distribution over this region. This would be followed by a heavier tailed perhaps parametric model $f_2(x)$ in the region $[x_{\min}, \infty)$, which would also be standardized by w_2 to ensure that the total resulting density on $\mathbb{R}^+$ was appropriately normalized. Clearly, there are several approaches that can be adopted to achieve this, for example, one may wish to ensure continuity or smoothness of the joint distribution such as at the boundary points between adjacent partitions. This will impose restrictions on the parameters controlling the distributional models; in other settings, such concerns will not be of consequence. Example illustrations of such models are provided in Examples 1.1–1.4, which illustrate a discontinuous model and continuous models, respectively.

EXAMPLE 1.1 Parametric Body and Parametric Tail

Assume that losses $X_1, X_2, \dots, X_K$ are independent and identically distributed. If we want to model the losses above a selected threshold $x_{\min}$ using some parametric distribution $G_2(x)$ with density $g_2(x)$ defined on $x > 0$ (e.g. LogNormal distribution) and the losses below using another parametric distribution $G_1(x)$ with density $g_1(x)$ defined on $x > 0$ (e.g. Gamma distribution), then corresponding density $f(x)$ and distribution $F(x)$ for spliced model to fit are

$$f(x) = wf_1(x) + (1-w)f_2(x),$$

$$F(x) = \begin{cases} wF_1(x), & 0 < x < x_{\min}, \\ w + (1-w)F_2(x), & x \geq x_{\min}, \end{cases}$$

where $w \in [0, 1]$ is the weight parameter and the proper densities $f_1(x)$ and $f_2(x)$ (and their distribution functions $F_1(x)$ and $F_2(x)$) correspond to the densities $g_1(x)$ and $g_2(x)$ truncated above and below $x_{\min}$, respectively:

$$f_1(x) = \frac{g_1(x)}{G_1(x_{\min})}\mathbb{I}_{x<x_{\min}}, \quad F_1(x) = \frac{G_1(x)}{G_1(x_{\min})},\ x < x_{\min},$$

$$f_2(x) = \frac{g_2(x)}{1 - G_2(x_{\min})}\mathbb{I}_{x\geq x_{\min}}, \quad F_2(x) = \frac{G_2(x) - G_2(x_{\min})}{1 - G_2(x_{\min})},\ x \geq x_{\min}.$$

■

EXAMPLE 1.2 Empirical Body and Parametric Tail

Assume that losses $X_1, X_2, \dots, X_K$ are independent and identically distributed. If we want to model the losses above a selected threshold $x_{\min}$ using some parametric distribution $G_2(x)$ with density $g_2(x)$ defined on $x > 0$ (e.g. LogNormal distribution) and the losses below using empirical distribution

$$G_1(x) = \frac{1}{K}\sum_{k=1}^{K}\mathbb{I}_{X_k \leq x},$$

then corresponding distribution $F(x)$ for the spliced model is

$$F(x) = \begin{cases} G_1(x), & 0 < x < x_{\min}, \\ G_1(x_{\min}) + (1 - G_1(x_{\min}))F_2(x), & x \geq x_{\min}, \end{cases}$$

where $F_2(x)$ is distribution $G_2(x)$ truncated below $x_{\min}$, that is,

$$F_2(x) = \frac{G_2(x) - G_2(x_{\min})}{1 - G_2(x_{\min})} \mathbb{I}_{x \geq x_{\min}}.$$

Comparing to Example 1.1, note that we selected weight parameter $w = G_1(x_{\min})$ to have a model consistent with the data below $x_{\min}$.

If the threshold $x_{\min}$ is large enough, then (under the regularity conditions of EVT threshold exceedances discussed in Chapter 2 and satisfied for most of the distributions used in practice), the truncated distribution $F_2(x)$ may be approximated by the generalized Pareto distribution

$$G_{\xi,\beta}(x - x_{\min}) = \begin{cases} 1 - (1 + \xi(x - x_{\min})/\beta)^{-1/\xi}, & \xi \neq 0, \\ 1 - \exp(-(x - x_{\min})/\beta), & \xi = 0. \end{cases}$$ ■

EXAMPLE 1.3 Gamma Body and Pareto Tail, Discontinuous Density

Consider a loss process with loss random variable $X \sim f_X(x)$ modeled according to a $k = 2$ component spliced distribution comprised of a gamma distribution over the intervals $[0, x_{\min})$ and a Pareto distribution over the interval $[x_{\min}, \infty)$. The resulting density, without any continuity restrictions at the partition boundary, is, therefore, given by

$$f_X(x) = \begin{cases} w_1 Z_1^{-1} \dfrac{\beta^{-\alpha}}{\Gamma(\alpha)} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right), & x \in [0, x_{\min}), \alpha, \beta > 0 \\ w_2 \dfrac{\gamma x_{\min}^{\gamma}}{x^{\gamma+1}}, & x \in [x_{\min}, \infty),\ \gamma > 0 \end{cases} \tag{1.10}$$

with

$$\begin{aligned} Z_1 = F_X(x; \alpha, \beta) &= \int_0^{x_{\min}} \frac{\beta^{-\alpha}}{\Gamma(\alpha)} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right) dx \\ &= \Gamma^{-1}(\alpha) \underbrace{\int_0^{x/\beta} t^{\alpha-1} \exp(-t)\, dt}_{\text{lower incomplete Gamma function}} \qquad (1.11) \\ &= \Gamma^{-1}(\alpha) \gamma\left(\alpha, \frac{x}{\beta}\right) \end{aligned}$$

and subject to the constraint that $\sum_{i=1}^{2} w_i = 1$. Furthermore, we may wish to consider cases typically in practice in which the mode of the first partitions distribution lies in the interval $(0, x_{\min})$ in which cases we further impose the restriction on the shape and scale parameters such that $(\alpha - 1)\beta \in (0, x_{\min})$. In the illustration, we consider the resulting density for the settings $x_{\min} = 100$, $w_1 = 0.5$, $\alpha = 2, \beta = 10$ and $\gamma = 0.4$ giving a density in Figure 1.1.

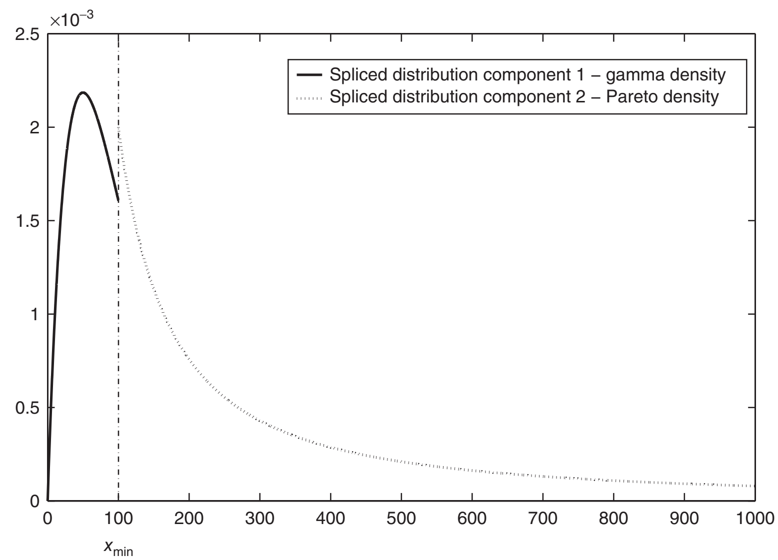

FIGURE 1.1 **Spliced density plot for gamma and Pareto distribution single risk severity model, no continuity constraints.**

EXAMPLE 1.4 Gamma Body and Pareto Tail, Continuous Density

Consider a loss process with loss random variable $X \sim f_X(x)$ modeled according to a $k = 2$ component spliced distribution comprised again of a gamma distribution over the intervals $[0, x_{\min})$ and a Pareto distribution over the interval $[x_{\min}, \infty)$. This time the resulting density is developed subject to the constraint that a certain degree of smoothness is present at the partition boundary connecting the two density functions as captured by equality of the first moments. The resulting density is then developed according to

$$f_X(x) = \begin{cases} w_1 Z_1^{-1} \dfrac{\beta^{-\alpha}}{\Gamma(\alpha)} x^{\alpha-1} \exp\left(-\dfrac{x}{\beta}\right), & x \in [0, x_{\min}), \alpha, \beta > 0, \\ w_2 \dfrac{\gamma x_{\min}^{\gamma}}{x^{\gamma+1}}, & x \in [x_{\min}, \infty), \gamma > 0, \end{cases} \tag{1.12}$$

with

$$\begin{aligned} Z_1 = F_X(x; \alpha, \beta) &= \int_0^{x_{\min}} \frac{\beta^{-\alpha}}{\Gamma(\alpha)} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right) dx \\ &= \Gamma^{-1}(\alpha) \underbrace{\int_0^{x/\beta} t^{\alpha-1} \exp(-t)\, dt}_{\text{lower incomplete gamma function}} \\ &= \Gamma^{-1}(\alpha) \gamma \left(\alpha, \frac{x}{\beta}\right) \end{aligned} \tag{1.13}$$

and subject to the constraint that $\sum_{i=1}^{2} w_i = 1$. Furthermore, we may wish to consider cases typically in practice in which the mode of the first partitions distribution lies in the interval $(0, x_{\min})$ in which cases we further impose the restriction on the shape and scale parameters such that $(\alpha - 1)\beta \in (0, x_{\min})$. In addition, the continuity constraint implies the following restriction on the two densities at $x_{\min}$,

$$\begin{aligned} f_1(x_{\min}) &= f_2(x_{\min}) \\ \left.\frac{df_1(x)}{dx}\right|_{x=x_{\min}} &= \left.\frac{df_2(x)}{dx}\right|_{x=x_{\min}}. \end{aligned} \tag{1.14}$$

Where these two restrictions create the following system of constraints that the model parameters must satisfy

$$\begin{aligned} f_1(x_{\min}) - f_2(x_{\min}) &= w_1 Z_1^{-1} \frac{\beta^{-\alpha}}{\Gamma(\alpha)} x_{\min}^{\alpha-1} \exp\left(-\frac{x_{\min}}{\beta}\right) - w_2 \frac{\gamma x_{\min}^{\gamma}}{x_{\min}^{\gamma+1}} = 0 \\ \left.\frac{df_1(x)}{dx}\right|_{x=x_{\min}} - \left.\frac{df_2(x)}{dx}\right|_{x=x_{\min}} &= \frac{w_1}{Z_1} \frac{\beta^{-\alpha}}{\Gamma(\alpha)} \left[(\alpha-1) x_{\min}^{\alpha-2} \exp\left(-\frac{x_{\min}}{\beta}\right) \right. \\ &\quad \left. - \frac{1}{\beta} x_{\min}^{\alpha-1} \exp\left(-\frac{x_{\min}}{\beta}\right)\right] + w_2 \gamma x_{\min}^{-2}(\gamma+1) = 0. \end{aligned} \tag{1.15}$$

In the following example illustration, we consider the resulting density for the settings $x_{\min} = 100$, $w_1 = 0.5$, $\alpha = 2$ and β, γ each set to satisfy these constraints giving a density given in Figure 1.2.

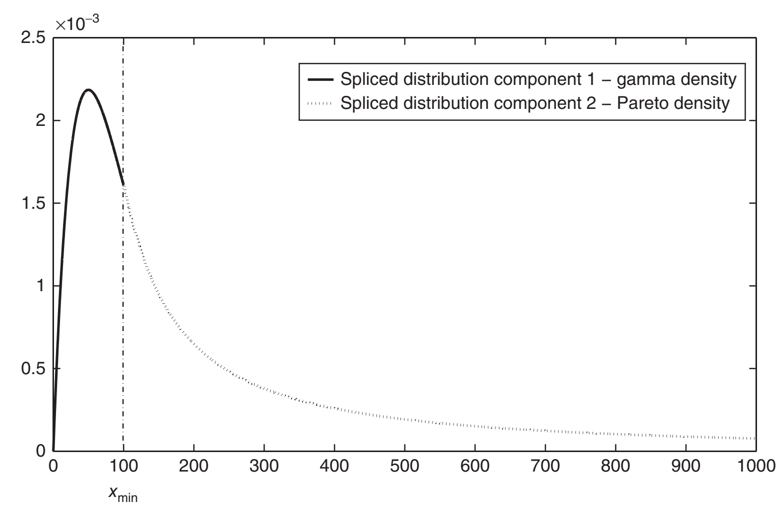

FIGURE 1.2 Spliced density plot for Gamma and Pareto distribution single risk severity model, with continuity constraints.