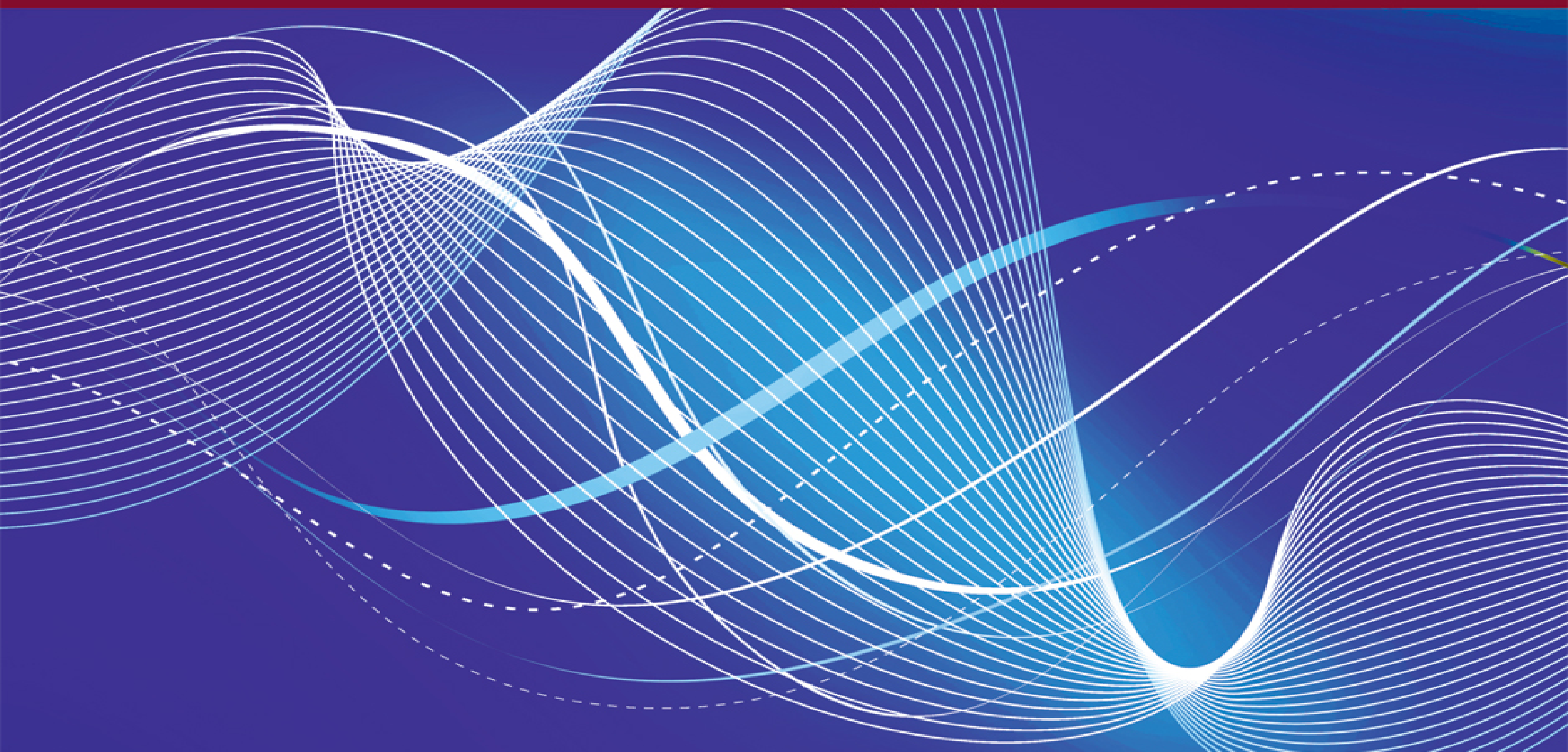
MATHEMATICS AND STATISTICS SERIES

Basic Stochastic Processes

Pierre Devolder
Jacques Janssen
Raimondo Manca



ISTE

WILEY

Introduction

This book will present basic stochastic processes for building models in insurance, especially in life and non-life insurance as well as credit risk for insurance companies. Of course, stochastic methods are quite numerous; so we have deliberately chosen to consider to use those induced by two big families of stochastic processes: stochastic calculus including Lévy processes and Markov and semi-Markov models. From the financial point of view, essential concepts such as the Black and Scholes model, VaR indicators, actuarial evaluation, market values and fair pricing play a key role, and they will be presented in this volume.

This book is organized into seven chapters. Chapter 1 presents the essential probability tools for the understanding of stochastic models in insurance. The next three chapters are, respectively, devoted to renewal processes (Chapter 2), Markov chains (Chapter 3) and semi-Markov processes both homogeneous and non-time homogeneous (Chapter 4) in time. This fact is important as new non-homogeneous time models are now becoming more and more used to build realistic models for insurance problems.

Chapter 5 gives the bases of stochastic calculus including stochastic differential equations, diffusion processes and changes of probability measures, therefore giving results that will be used in Chapter 6 devoted to Lévy processes. Chapter 6 is devoted to Lévy processes. This chapter also presents an alternative to basic stochastic models using Brownian motion as Lévy processes keep the properties of independent and stationary increments but without the normality assumption.

Finally, Chapter 7 presents a summary of Solvency II rules, actuarial evaluation, using stochastic instantaneous interest rate models, and VaR methodology in risk management.

Our main audience is formed by actuaries and particularly those specialized in entreprise risk management, insurance risk managers, Master's degree students in mathematics or economics, and people involved in Solvency II for insurance companies and in Basel II and III for banks. Let us finally add that this book can also be used as a standard reference for the basic information in stochastic processes for students in actuarial science.

1
Basic Probabilistic Tools for Stochastic Modeling

In this chapter, the readers will find a brief summary of the basic probability tools intensively used in this book. A more detailed version including proofs can be found in [JAN 06].

1.1. Probability space and random variables

Given a sample space Ω, the set of all possible events will be denoted by $\Im$, which is assumed to have the structure of a σ *-field* or a σ *-algebra. P* will represent a probability measure.

DEFINITION 1.1.– *A random variable (r.v.) with values in a topological space* (E,ψ) *is an application X from* Ω *to E such that:*

$$\forall B \in \psi : X^{-1}(B) \in \Im, \tag{1.1}$$

where $X^{-1}(B)$ is called the inverse image of the set B defined by:

$$X^{-1}(B) = \{\omega : X(\omega) \in B\}, X^{-1}(B) \in \Im. \tag{1.2}$$

Particular cases:

a) If $(E,\psi) = (\mathbb{R}, \beta)$, X is called a *real random variable*.

b) If $(E, \psi) = (\overline{\mathbb{R}}, \overline{\beta})$, where $\overline{\mathbb{R}}$ is the *extended real line* defined by $\mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$ and $\overline{\beta}$ is the *extended Borel σ -*

field of $\overline{\mathbb{R}}$, that is the minimal σ-field containing all the elements of β and the extended intervals:

$$\begin{aligned}&[-\infty,a),(-\infty,a],[-\infty,a],(-\infty,a),\\&[a,+\infty),(a,+\infty],[a,+\infty],(a,+\infty),\ a\in\mathbb{R},\end{aligned} \quad [1.3]$$

X is called a *real extended value* random variable.

c) If $E = \mathbb{R}^n$ (n>1) with the product σ-field $\beta^{(n)}$ of β, X is called an *n-dimensional real random variable*.

d) If $E = \overline{\mathbb{R}}^{(n)}$ (n>1) with the product σ-field $\beta^{(n)}$ of β, X is called a real *extended n-dimensional real random variable.*

A random variable X is called *discrete* or *continuous* accordingly as X takes at most a denumerable or a non-denumerable infinite set of values.

DEFINITION 1.2.– *The distribution function of the r.v. X, represented by F_X, is the function from* $\mathbb{R}\to[0,1]$ *defined by:*

$$F_X(x) = P\left(\{\omega : X(\omega) \le x\}\right). \quad [1.4]$$

Briefly, we write:

$$F_X(x) = P\left(X \le x\right). \quad [1.5]$$

This last definition can be extended to the multi-dimensional case with a r.v. X being an n-dimensional real vector: $X = (X_1,\ldots, X_n)$, a measurable application from $(\Omega, \Im, P)$ to $(\mathbb{R}^n, \beta^n)$.

DEFINITION 1.3.– *The distribution function of the r.v.* $X = (X_1,\ldots, X_n)$ *, represented by F_X, is the function from* $\mathbb{R}^n$ *to* [0,1] *defined by:*

$$F_X(x_1,\ldots,x_n) = P\left(\{\omega : X_1(\omega) \le x_1,\ldots,X_n(\omega) \le x_n\}\right). \quad [1.6]$$

Briefly, we write:

$$F_X(x_1,\ldots,x_n) = P(X_1 \le x_1,\ldots,X_n \le x_n). \quad [1.7]$$

Each component X_i ($i = 1,\ldots,n$) is itself a one-dimensional real r.v. whose d.f., called the *marginal d.f.*, is given by:

$$F_{X_i}(x_i) = F_X(+\infty,\ldots,+\infty,x_i,+\infty,\ldots,+\infty). \quad [1.8]$$

The concept of random variable is *stable* under a lot of mathematical operations; so any Borel function of a r.v. X is also a r.v.

Moreover, if X and Y are two r.v., so are:

$$\inf\{X,Y\}, \sup\{X,Y\}, X+Y, X-Y, X\cdot Y, \frac{X}{Y}, \quad [1.9]$$

provided, in the last case, that Y does not vanish.

Concerning the convergence properties, we must mention the property that, if $(X_n, n \ge 1)$ is a *convergent* sequence of r.v. – that is, for all $\omega\in\Omega$, the sequence $(X_n(\omega))$ converges to $X(\omega)$ – then the limit X is also a r.v. on Ω. This convergence, which may be called the *sure convergence*, can be weakened to give the concept of *almost sure* (a.s.) *convergence* of the given sequence.

DEFINITION 1.4.– *The sequence* $(X_n(\omega))$ *converges a.s. to* $X(\omega)$ *if:*

$$P\left(\{\omega : \lim X_n(\omega) = X(\omega)\}\right) = 1 \quad [1.10]$$

This last notion means that the possible set where the given sequence does not converge is a *null* set, that is, a set N belonging to $\Im$ such that:

$$P(N) = 0. \quad [1.11]$$

In general, let us remark that, given a null set, it is not true that every subset of it belongs to $\Im$ but of course if it belongs to $\Im$, it is clearly a null set. To avoid unnecessary complications, we will assume from here onward that any considered probability space is *complete*, i.e. all the subsets of a null set also belong to $\Im$ and thus their probability is zero.

1.2. Expectation and independence

Using the concept of integral, it is possible to define the *expectation* of a random variable X represented by:

$$E(X) = \int_{\Omega} X dP \left(= \int X dP \right), \quad [1.12]$$

provided that this integral exists. The computation of the integral:

$$\int_{\Omega} X dP \left(= \int X dP \right) \quad [1.13]$$

can be done using the induced measure μ on ($\mathbb{R}$, β), defined by [1.4] and then using the distribution function F of X.

Indeed, we can write:

$$E(X) \left(= \int_{\Omega} X dP \right), \quad [1.14]$$

and if F_X is the d.f. of X, it can be shown that:

$$E(X) = \int_R x dF_X(x). \quad [1.15]$$

The last integral is a Lebesgue–Stieltjes integral.

Moreover, if F_X is absolutely continuous with f_X as density, we obtain:

$$E(X) = \int_{-\infty}^{+\infty} x f_x(x) dx. \quad [1.16]$$

If g is a Borel function, then we also have (see, e.g. [CHU 00] and [LOÈ 63]):

$$E(g(X)) = \int_{-\infty}^{+\infty} g(x) dF_X \quad [1.17]$$

and with a density for *X:*

$$E(g(X)) = \int_{-\infty}^{+\infty} g(x) f_X(x) dx. \quad [1.18]$$

It is clear that the expectation is a linear operator on integrable functions.

DEFINITION 1.5.– *Let a be a real number and r be a positive real number, then the expectation:*

$$E\left(|X-a|^r\right) \quad [1.19]$$

is called the absolute moment of X, of order r, centered on a.

The moments are said to be centered moments of order r if $a=E(X)$. In particular, for $r = 2$, we get the *variance* of X represented by σ^2 (var(X)) :

$$\sigma^2 = E\left(|X-m|^2\right). \qquad [1.20]$$

REMARK 1.1.– From the linearity of the expectation, it is easy to prove that:

$$\sigma^2 = E(X^2) - (E(X))^2, \qquad [1.21]$$

and so:

$$\sigma^2 \le E(X^2), \qquad [1.22]$$

and, more generally, it can be proved that the variance is the smallest moment of order 2, whatever the number *a* is.

The set of all real r.v. such that the moment of order r exists is represented by L^r.

The last fundamental concept that we will now introduce in this section is *stochastic independence*, or more simply *independence*.

DEFINITION 1.6.– *The events* $A_1, \ldots, A_n$, $(n > 1)$ *are stochastically independent or independent iff:*

$$\forall m = 2,\ldots,n, \forall n_k = 1,\ldots,n : n_1 \neq n_2 \neq \cdots \neq n_k : P\left(\bigcap_{k=1}^{m} A_{n_k}\right) = \prod_{k=1}^{m} P(A_{n_k}). \qquad [1.23]$$

For $n = 2$, relation [1.23] reduces to:

$$P(A_1 \cap A_2) = P(A_1)P(A_2). \qquad [1.24]$$

Let us remark that piecewise independence of the events $A_1, \ldots, A_n$, $(n > 1)$ does not necessarily imply the independence of these sets and, thus, not the stochastic independence of these n events.

From relation [1.23], we find that:

$$P(X_1 \leq x_1, \ldots, X_n \leq x_n) = P(X_1 \leq x_1) \cdots P(X_n \leq x_n), \forall (x_1, \ldots, x_n) \in \mathbb{R}^n . \quad [1.25]$$

If the functions F_X, $F_{X_1}, \ldots, F_{X_n}$ are the distribution functions of the r.v. $X = (X_1, \ldots, X_n)$, $X_1, \ldots, X_n$, we can write the preceding relation as follows:

$$F_X(x_1, \ldots, x_n) = F_{X_1}(x_1) \cdot \cdots \cdot F_{X_n}(x_n), \forall (x_1, \ldots, x_n) \in \mathbb{R}^n . \quad [1.26]$$

It can be shown that this last condition is also sufficient for the independence of $X = (X_1, \ldots, X_n)$, $X_1, \ldots, X_n$. If these d.f. have densities f_X, $f_{X_1}, \ldots, f_{X_n}$, relation [1.24]is equivalent to:

$$f_X(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n), \forall (x_1, \ldots, x_n) \in \mathbb{R}^n . \quad [1.27]$$

In case of the integrability of the n real r.v $X_1, X_2, \ldots, X_n$, a direct consequence of relation [1.26] is that we have a very important property for the expectation of the product of n independent r.v.:

$$E\left(\prod_{k=1}^{n} X_k\right) = \prod_{k=1}^{n} E(X_k) . \quad [1.28]$$

The notion of independence gives the possibility of proving the result called the *strong law of* large *numbers*, which states that if $(X_n, n \geq 1)$ is a sequence of integrable independent and identically distributed r.v., then:

$$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{a.s.} E(X) . \quad [1.29]$$

The next section will present the most useful distribution functions for stochastic modeling.

DEFINITION 1.7 (SKEWNESS AND KURTOSIS COEFFICIENTS).–

a) The *skewness coefficient of Fisher* is defined as follows:

$$\gamma_1 = \frac{E\left[(X - E(X))^3\right]}{\sigma^3}$$

From the odd value of this exponent, it follows that:

$-\gamma_1>0$ gives a *left dissymmetry* giving a maximum of the density function situated to the left and a distribution with a right heavy queue, $\gamma_1 = 0$ gives *symmetric* distribution with respect to the mean;

$-\gamma_1<0$ gives a *right dissymmetry* giving a maximum of the density function situated to the right and a distribution with a left heavy queue.

b) The *kurtosis coefficient* also due to Fisher is defined as follows:

$$\gamma_2 = \frac{E\left[(X - E(X))^4\right]}{\sigma^4}$$

Its interpretation refers to the normal distribution for which its value is 3. Also some authors refer to the *excess of kurtosis* given by γ_1-3 of course null in the normal case.

For $\gamma_2<3$, distributions are called *leptokurtic*, being more plated around the mean than in the normal case and with heavy queues.

For $\gamma_2>3$, distributions are less plated around the mean than in the normal case and with heavy queues.

1.3. Main distribution probabilities

In this section, we will restrict ourselves to presenting the principal distribution probabilities related to real random variables.

1.3.1. *Binomial distribution*

Let X be a discrete random variable, whose distribution (p_i, i = 0,..., n) with:

$$p_i = P(X = i),\ i = 1,\ldots,n \qquad [1.30]$$

is called a *binomial* distribution with parameters (*n,p*) if:

$$p_i = \binom{n}{i} p^i q^{n-i},\ i = 0,\ldots,n, \qquad [1.31]$$

a result from which we get:

$$E(X) = np, \operatorname{var}(X) = npq. \qquad [1.32]$$

The characteristic function and the generating function, when the latter exists, of X, respectively, defined by:

$$\varphi_X(t) = E(e^{itX}), \qquad [1.33]$$
$$g_X(t) = E(e^{tX})$$

are given by:

$$\varphi_X(t) = (pe^{it} + q)^n, \qquad [1.34]$$
$$g_X(t) = (pe^t + q)^n.$$

1.3.2. *Negative exponential distribution*

This is defined by:

$$\Pr[X \le x] = \begin{cases} 0 & x \le 0 \\ 1 - e^{-\frac{1}{\beta}x} & x \ge 0 \end{cases} \qquad [1.35]$$

$$E[X] = \beta,$$

$$\text{variance}[X] = \beta^2,$$

$$\gamma_1 = 2, \gamma_2 = 9.$$

1.3.3. *Normal (or Laplace–Gauss) distribution*

The real r.v. X has a normal (or Laplace–Gauss) distribution of parameters$(\mu,\sigma^2),\mu\in\mathbb{R},\sigma^2>0$, if its density function is given by:

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in \mathbb{R}. \qquad [1.36]$$

From now on, we will use the notation $X \prec N(\mu,\sigma^2)$.

The main parameters of this distribution are:

$$E(X) = \mu, \ \ \text{var}(X) = \sigma^2, \qquad [1.37]$$

$$\varphi_X(t) = \exp\left(i\mu t - \frac{\sigma^2 t^2}{2}\right), \ g_X(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right).$$

If $\mu = 0$, $\sigma^2 = 1$, the distribution of X is called a *reduced* or *standard normal distribution*. In fact, if X has a normal distribution (μ,σ^2), $\mu \in R, \sigma^2 > 0$, then the so-called reduced r.v. Y defined by:

$$Y = \frac{X - \mu}{\sigma} \qquad [1.38]$$

has a standard normal distribution, thus from [1.36] with mean 0 and variance 1.

Let $\exists k > 0 : \lim_{x\to\infty} \frac{1-F(x)}{1-F(ux)} = u^k$, for all $u > 0$ be the distribution function of the standard normal distribution; it is possible to express the distribution function of anynormal r.v. X of parameters (μ, σ^2), $\mu \in \mathbb{R}$, $\sigma^2 > 0$, as follows:

$$F_X(x) = P(X \le x) = P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right). \quad [1.39]$$

In addition, from the numerical point of view, it is sufficient to know the numerical values for the standard distribution.

From relation [1.38], we also deduce that:

$$f_X(x) = \frac{1}{\sigma}\Phi'\left(\frac{x-\mu}{\sigma}\right), \quad [1.40]$$

where, of course, from [1.35]:

$$\Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \quad [1.41]$$

From the definition of Φ, we have:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}} dy, \; x \in \mathbb{R} \quad [1.42]$$

and so:

$$\Phi(-x) = 1 - \Phi(x), x > 0, \quad [1.43]$$

and consequently, for X normally distributed with parameters (0, 1), we obtain:

$$P(|X| \le x) = \Phi(x) - \Phi(-x) = 2\Phi(x) - 1, \; x > 0. \quad [1.44]$$

In particular, let us mention the following numerical results:

$$P\left(|X-m| \le \frac{2}{3}\sigma\right) = 0.4972(\approx 50\%), \quad [1.45]$$
$$P\left(|X-m| \le \sigma\right) = 0.6826(\approx 68\%),$$
$$P\left(|X-m| \le 2\sigma\right) = 0.9544(\approx 95\%),$$
$$P\left(|X-m| \le 3\sigma\right) = 0.9974(\approx 99\%).$$

The normal distribution is one of the most often used distributions by virtue of the *central limit theorem*, which states that if (X_n, $n \ge 1$) is a sequence of independent and identically distributed (i.i.d.) r.v. with mean m and variance σ^2, then the sequence of r.v. is defined by:

$$\frac{S_n - nm}{\sigma\sqrt{n}} \quad [1.46]$$

with:

$$S_n = X_1 + \cdots + X_n, \quad n > 0 \quad [1.47]$$

converging in law to a standard normal distribution.

This means that the sequence of the distribution functions of the variables defined by [3.20] converges to Φ.

This theorem was used by the Nobel Prize winner H. Markowitz [MAR 159] to justify that the return of a diversified portfolio of assets has a normal distribution. As a particular case of the central limit theorem, let us mention the *de Moivre's theorem* starting with:

$$X_n = \begin{cases} 1, \text{ with prob. } p, \\ 0, \text{ with prob. } 1-p, \end{cases} \quad [1.48]$$

so that, for each *n,* the r.v. defined by relation [1.47] has a binomial distribution with parameters (*n,p*).

Now by applying the central limit theorem, we get the following result:

$$\frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow[n \to +\infty]{law} N(0,1), \quad [1.49]$$

which is called de Moivre's result.

1.3.4. *Poisson distribution*

This is a discrete distribution with the following characteristics:

$$\begin{aligned} &P(\xi = n) = e^{-\lambda} \frac{\lambda^n}{n!}, n = 0,1,\ldots \\ &m = \sigma^2 = \lambda, \\ &\gamma_1 = \frac{1}{\sqrt{\lambda}}, \gamma_2 = \frac{1}{\lambda} + 3. \end{aligned} \quad [1.50]$$

This is one of the most important distributions for all applications. For example, if we consider an insurance company looking at the total number of claims in one year, this variable may often be considered to be a Poisson variable.

1.3.5. *Lognormal distribution*

This is a continuous distribution on the positive half line with the following characteristics:

$$\Pr[\ln X \le x] = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad [1.51]$$

$$E[X] = e^{\mu + \frac{\sigma^2}{2}},$$

$$\text{variance}[X] = e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right),$$

$$\gamma_1 = (e^{\sigma^2} + 2)\sqrt{(e^{\sigma^2} - 1)}$$

$$\gamma_2 = \omega^4 + 2\omega^3 + 3\omega^2 - 3, \omega = e^{\sigma^2}.$$

Let us say that the lognormal distribution has no generating function and that the characteristic function has no explicit form. When $\sigma < 0.3$, some authors recommend a normal approximation with parameters (μ, σ^2).

The normal distribution is *stable* under the addition of independent random variables; this property means that the sum of *n* independent normal r.v. is still normal. This is no longer the case with the lognormal distribution which is stable under *multiplication*, which means that for two independent lognormal r.v. X_1, X_2, we have

$$X_i \prec LN(\mu_i, \sigma_i), i = 1, 2 \Rightarrow X_1 \times X_2 \prec LN\left(\mu_1 + \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2}\right). \quad [1.52]$$

1.3.6. *Gamma distribution*

This is a continuous distribution on the positive half line having the following characteristics:

$$\Pr\left[x < X \leq x + \Delta x\right] = \frac{\theta^{\nu}}{(\nu - 1)!} x^{\nu - 1} e^{-\theta x} \Delta x \qquad [1.53]$$

$$E\left[X\right] = \frac{\nu}{\theta},$$

$$\text{variance}\left[X\right] = \frac{\nu}{\theta^2},$$

$$\gamma_1 = 2 / \sqrt{\nu},$$

$$\gamma_2 = \frac{6}{\nu} + 3.$$

For the gamma law with parameters (ν, θ) denoted $\gamma(\nu, \theta)$, an additivity property exists:

$$\gamma(\nu, \theta) + \gamma(\nu', \theta) \succ \gamma(\nu + \nu', \theta).$$

1.3.7. *Pareto distribution*

The non-negative r.v. X has a *Pareto distribution* if its distribution function is given by:

$$F_X(x) = 1 - \left(\frac{k}{x}\right)^{\alpha}, \quad x > k;\ k > 0,\ \alpha > 0. \qquad [1.54]$$

Its support is $(k, +\infty)$.

The corresponding density function is given by:

$$f_X(x) = \frac{\alpha k^{\alpha}}{x^{\alpha+1}}, x \geq k. \qquad [1.55]$$

The Pareto distribution has centered moments of order r provided that $r < \alpha$; and in this case:

$$E\left[X^r\right] = \frac{\alpha k^r}{\alpha - r}, \ r < \alpha$$

and so:

$$E[X] = \frac{\alpha k}{\alpha - 1}, \alpha > 1,$$

$$\operatorname{var} X = \frac{\alpha k^2}{(\alpha - 1)^2 (\alpha - 2)}, \alpha > 2.$$

$$\gamma_1 = \frac{\alpha + 1}{\alpha - 3}\sqrt{1 - \frac{2}{\alpha}}, \alpha > 3; \gamma_2 = \frac{3(\alpha - 2)(3\alpha^2 + \alpha + 2)}{\alpha(\alpha - 3)(\alpha - 4)}, \alpha > 4.$$

These values explain why this distribution is considered to be *dangerous* as for some values of the parameters it is not excluded to observe large values of X in a random experiment.

For $\alpha < 1$, the mean is infinite, and for $1 < \alpha < 2$, although the mean is finite, the variance is not.

The problem of this distribution also comes from the fact that the function $1 - F(x)$ decreases in a polynomial way for large x (distribution with heavy queue) and no longer exponentially like the other presented distributions, except, of course, for the Cauchy distribution.

In non-life insurance, it is used for modeling large claims and catastrophic events.

REMARK 1.2.– We also have:

$$\ln(1 - F_X(x)) = \ln\left(\frac{k}{x}\right)^{\alpha}, \ x > k;\ k > 0,\ \alpha > 0,$$

ou

$$\ln(1 - F_X(x)) = \alpha(\ln k - \ln x).$$

REMARK 1.3.– If we compare the form:

$$F_X(x) = 1 - \left(\frac{k}{x}\right)^{\alpha}, \quad x > k;\ k > 0,\ \alpha > 0. \qquad [1.56]$$

with:

$$F_{X'}(x) = \begin{cases} 1 - \left(\dfrac{\theta}{x+\theta}\right)^{\beta}, & x \geq 0, \\ 0, & x < 0, \end{cases} \qquad [1.57]$$

we have another form of the Pareto distribution with as support all the positive half line. This is possible with the change of variable:

$$Y = X + k$$

Of course, the variances remain the same but not the means.

$$m_X = \frac{\alpha k}{\alpha - 1}(\alpha > 1),\ m_{X'} = \frac{\theta}{\beta - 1}(\beta > 1) \qquad [1.58]$$

$$\sigma_X^2 = \frac{\alpha k^2}{(\alpha-1)^2(\alpha-2)}(\alpha > 2), \quad \sigma_{X'}^2 = \frac{\theta^2 \beta}{(\beta-1)^2(\beta-2)}(\beta > 2) = \frac{m_{X'}^2 \beta}{\beta - 2}$$

Here are two graphs of the distribution function showing the impact of the dangerous parameters.

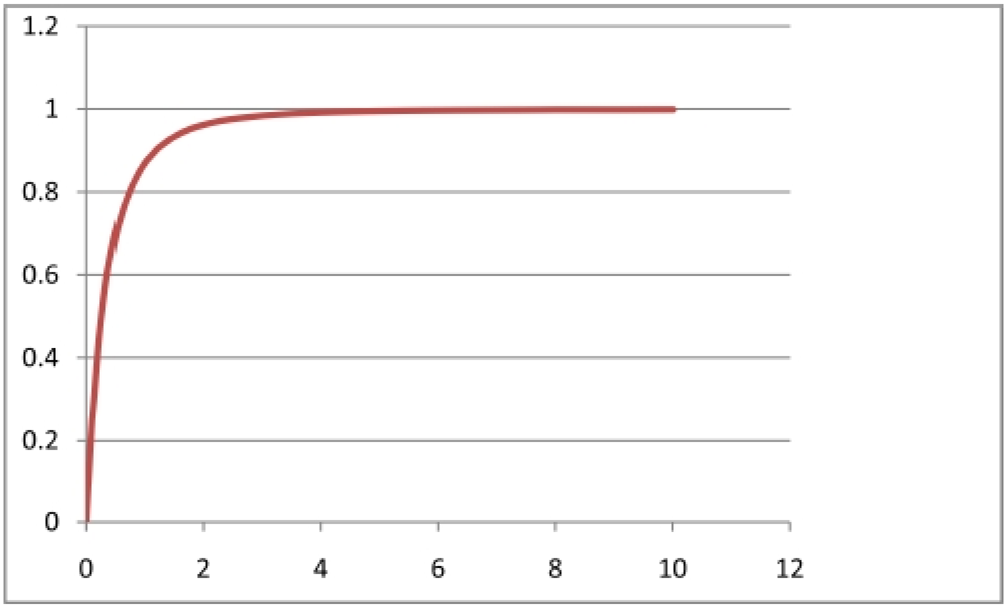


Figure 1.1. *Pareto distribution function with Θ=1,β=1*

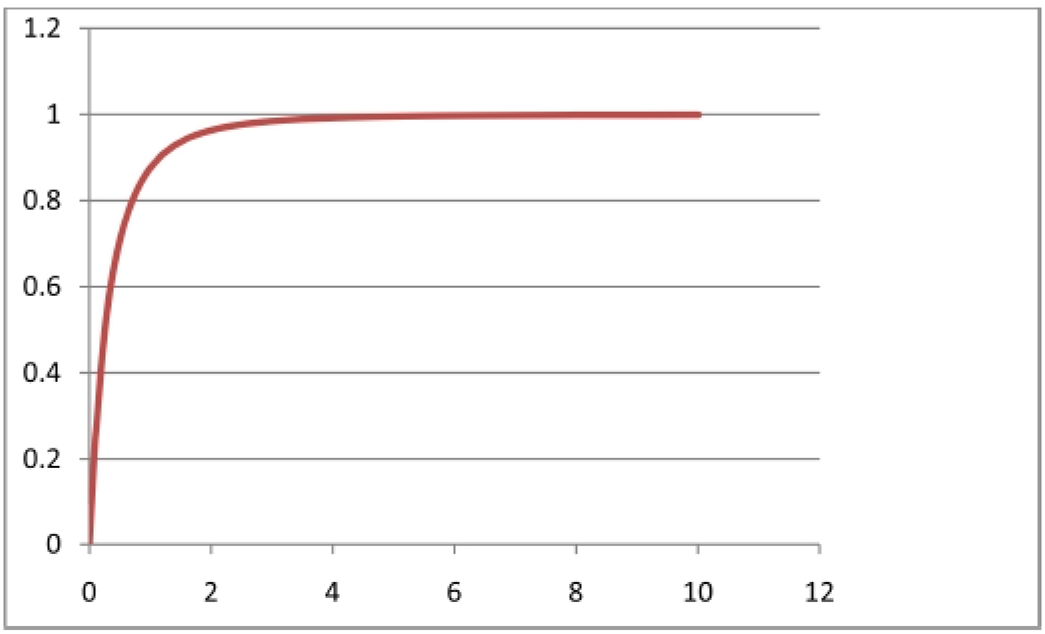


Figure 1.2. *Pareto distribution with β=3,θ=1*

REMARK 1.4.– As we have:

$$\ln(1-F_X(x)) = \ln\left(\frac{k}{x}\right)^{\alpha}, \; x > k;\; k > 0,\; \alpha > 0, \qquad [1.59]$$

or

$$\ln(1-F_X(x)) = \alpha(\ln k - \ln x).$$

The proportion of claims larger than x is a linear function of x in a double logarithmic scale with α as slope.

1.3.8. *Uniform distribution*

Its support is [a,b] on which the density is constant with the value 1/(b-a).

For basic parameters, we have:

$$m = \frac{b-a}{2}, \sigma^2 = \frac{(b-a)^2}{12}$$
$$\gamma_1 = 0, \gamma_2 = 1,8.$$

1.3.9. *Gumbel distribution*

This is related to a non-negative random variable with the following characteristics:

$$F(x) = e^{-e^{-x}}, \qquad [1.60]$$
$$f(x) = e^{(-x-e^{-x})},$$
$$E(X) = 0,57722...,$$
$$\mathrm{var}(X) = \frac{\pi^2}{6},$$
$$\gamma_1 = 1,29857, \gamma_2 = 5,4.$$

1.3.10. *Weibull distribution*

This is related to a non-negative random variable with the following characteristics:

$$F(x) = 1 - e^{-(\frac{x}{\alpha})^{\beta}}, \alpha, \beta > 0. \qquad [1.61]$$
$$f(x) = \alpha\beta^{-\alpha}x^{\alpha-1}e^{-(\frac{x}{\alpha})^{\beta}},$$
$$E(X) = \frac{\beta}{\alpha}\Gamma(\frac{1}{\alpha}),$$
$$\mathrm{var}(X) = \frac{\beta^2}{\alpha}\left[2\Gamma(\frac{2}{\alpha}) - \frac{1}{\alpha}(\Gamma(\frac{1}{\alpha})^2\right].$$

1.3.11. *Multi-dimensional normal distribution*

Let us consider an n-dimensional real r.v. X represented as a column vector of its n components $X = (X_1,\ldots, X_n)'$. Its d.f. is given by:

$$F_X(x_1,\ldots,x_n) = P(X_1 \le x_1,\ldots, X_n \le x_n). \qquad [1.62]$$

If the density function of X exists, the relations between the d.f. and the density function are:

$$f_X(x_1,\ldots,x_n) = \frac{\partial^n F_X}{\partial x_1 \ldots \partial x_n}(x_1,\ldots,x_n), \qquad [1.63]$$
$$F_X(x_1,\ldots,x_n) = \int_{-\infty}^{x_1}\ldots\int_{-\infty}^{x_n} f_X(\xi_1,\ldots,\xi_n)d\xi_1,\ldots,d\xi_n.$$

For the principal parameters, we will use the following notation:

$$E(X_k) = \mu_k, k = 1,\ldots,n, \qquad [1.64]$$
$$E((X_k - \mu_k)(X_l - \mu_l)) = \sigma_{kl}, k, l = 1,\ldots,n,$$
$$E((X_k - \mu_k))^2 = \sigma_k^2, k = 1,\ldots,n,$$
$$\rho_{kl} = \frac{E((X_k - \mu_k)(X_l - \mu_l))}{\sqrt{E((X_k - \mu_k)^2)E((X_k - \mu_k)^2)}}\left(= \frac{\sigma_{kl}}{\sigma_k\sigma_l}\right), k, l = 1,\ldots,n.$$

The parameters σ_{kl} are called the *covariances* between the r.v. X_k and X_l and the parameters ρ_{kl} are called the *correlation coefficients* between the r.v. X_k and X_l.

It is well known that the correlation coefficient ρ_{kl} measures a certain linear dependence between the two r.v. X_k and X_l. More precisely, if it is equal to 0, then there is no such dependence and the two variables are called *uncorrelated*; for the values +1 and –1, this dependence is certain.

With matrix notation, the following σ matrix:

$$\mathbf{\Sigma}_X = \left[\sigma_{ij} \right] \tag{1.65}$$

is called the *variance*–covariance matrix of X.

The characteristic function of X is defined as:

$$\varphi_X(t_1,\dots,t_n) = E\left(e^{i(t_1X_1+..+t_nX_n)}\right)\left(= E\left(e^{i\mathbf{t}'X}\right)\right). \tag{1.66}$$

Let **μ,Σ** be, respectively, an n-dimensional real vector and an $n \times n$ positive definite matrix. The n-dimensional real r.v. X has a *non-degenerated n-dimensional normal distribution* with parameters **μ,Σ** if its density function is given by:

$$f_X(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det \mathbf{\Sigma}}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}, \mathbf{x} \in \mathbb{R}^n. \tag{1.67}$$

Then, it can be shown by integration that parameters **μ,Σ** are indeed, respectively, the *mean vector* and the *variance–covariance matrix* of X.

As usual, we will use the notation $X \prec N_n(\boldsymbol{\mu},\mathbf{\Sigma})$.

The characteristic function of X is given by:

$$\varphi_X(\mathbf{t}) = e^{i\boldsymbol{\mu}'\mathbf{t} - \frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}} . \qquad [1.68]$$

The main fundamental properties of the *n*-dimensional normal distribution are:

– every subset of *k* r.v. of the set $\{X_1,\ldots,X_n\}$ also has a *k*-dimensional distribution which is also normal;

– the multi-dimensional normal distribution is *stable* under linear transformations of X and for the addition, we have that if $X_k \prec N_n(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$, $k = 1,\ldots, m$ and if these m random vectors are independent, then:

$$X_1 + \cdots + X_m \prec N_n(\boldsymbol{\mu}_1 + \cdots + \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_1 + \cdots + \boldsymbol{\Sigma}_m) . \qquad [1.69]$$

In the particular case of the *two-dimensional normal distribution*, we have:

$$\boldsymbol{\mu} = (\mu_1, \mu_2)', \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}, \rho = \frac{\sigma_{12}}{\sigma_1\sigma_2}, \qquad [1.70]$$

$$\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} \dfrac{1}{\sigma_1^2} & -\dfrac{\rho}{\varsigma_1\sigma_2} \\ -\dfrac{\rho}{\varsigma_1\sigma_2} & \dfrac{1}{\sigma_2^2} \end{bmatrix}, \det\boldsymbol{\Sigma} = \sigma_1\sigma_2\sqrt{1-\rho^2}.$$

From the first main fundamental properties of the *n*-dimensional normal distribution given above, we deduce that:

$$X_k \prec N_1(\mu_k, \sigma_k^2), k = 1, 2 . \qquad [1.71]$$

For the special degenerated case $|\rho| = 1$, it can be proved that:

$$\rho = 1: \frac{X_2 - \mu_2}{\sigma_2} = \frac{X_1 - \mu_1}{\sigma_1}, \qquad [1.72]$$

$$\rho = -1: \frac{X_2 - \mu_2}{\sigma_2} = -\frac{X_1 - \mu_1}{\sigma_1},$$

relations meaning that, in this case, the entire probability mass in the plan lies on a straight line so that the two random variables X_1 and X_2 are perfectly dependent, i.e. relations [1.72] are true with probability one.

To conclude this section, let us recall the well-known property saying that two independent r.v. are uncorrelated but the converse is not true except for the normal distribution.

1.3.12. *Extreme value distribution*

In this section, we present basic results on the theory of extreme values [EMB 08], well adapted for the large claims designing risks which can take large values even with small probabilities but far from the mean value.

1.3.12.1. *Definition*

Let $X_1, \dots, X_n$ be the independent realizations of the considered risk X and the risk of this sample can be measured by the largest claim value that is:

$$Z_n = \max\{X_1, \dots, X_n\}. \qquad [1.73]$$

If F is the distribution function of the r.v. X, we have from the independence assumption:

$$P(Z_n \le z) = F^n(x). \qquad [1.74]$$

As this probability tends to 0 with $n \to \infty$ and for all x, it is necessary to find asymptotic results giving a more precise view of what could happen for large n and that is, in fact, the aim of extreme theory.

1.3.12.2. *Asymptotic results*

Fisher [FIS 28] and Gnedenko [GNE 43] proved that if there exist two sequences of real numbers $(c_n),(d_n)$ with $X_1,\ldots, X_n > 0$ for all n, such that the distribution of the following r.v.:

$$Y_n = \frac{\max\{X_1,\ldots X_n\} - d_n}{c_n} \qquad [1.75]$$

is not degenerated, then the limit distribution of Y_n must have one of the three following forms:

$$\text{i) Gumbel's law: } \Lambda(x) = e^{-e^{-x}}, x \in \mathbb{R}. \qquad [1.76]$$

$$\text{ii) Fréchet's law: } \Phi_\beta(x) = \begin{cases} 0, x \le 0, \\ e^{-x^{-\beta}}, x > 0, \end{cases}, \beta > 0. \qquad [1.77]$$

$$\text{iii) Weibull's law: } \psi_\beta(x) = \begin{cases} e^{-(-x)^\beta}, x < 0, \\ 1, \quad x \ge 0, \end{cases} \beta > 0. \qquad [1.78]$$

If we introduce the function

$$a(y) = \exp\left[-(1-\tau y)^{1/\tau}\right], \qquad [1.79]$$

($\tau = -\dfrac{1}{\beta}$ for Fréchet, $\tau = \dfrac{1}{\beta}$ for Weinbull and $\tau = 0$ for Gumbel),

the preceding characterizations of the three attractions domains can be given as follows:

$$\text{Fréchet}: F(y) = \begin{cases} 0, y \le 1/\tau \\ a(y), y > 1/\tau \end{cases} \tau < 0,$$
$$\text{Gumbel}: F(y) = a(y) = e^{-e^{-y}}, \tau = 0,$$
$$\text{Weibull}: F(y) = \begin{cases} a(y), y < 1/\tau \\ 1, y \ge 1/\tau \end{cases} \tau > 0. \quad [1.80]$$

To verify this result, for example, for the Fréchet law, we introduce z and k defined as:

$$x = 1 - \tau y, \beta = -\frac{1}{\tau}$$

with $\tau < 0$ as k is positive and $y > \frac{1}{\tau}$ as $x > 0$. Moreover, as from [1.77], we obtain well:

$$\Phi_\alpha(x) = \begin{cases} 0, y \le \frac{1}{\tau}, \\ e^{-(1-\tau y)^{1/\tau}}, y > \frac{1}{\tau}. \end{cases} \quad [1.81]$$

For the Weibull case, we have to define x and β as:

$$x = \tau y - 1, \beta = \frac{1}{\tau}$$

And for the Gumbel case, we have:

$$\lim_{\tau \searrow 0} \exp\left[-(1 - \tau y)^{1/\tau}\right] = \exp\left[-\lim_{n \to \infty}(1 - \frac{y}{n})^n\right]$$
$$= e^{-e^{-y}}.$$

REMARK 1.5.–

i) By Taylor expansion, we have for the Fréchet law:

$$\Phi_{\beta}(x) \approx x^{-\beta}, x \to \infty \qquad [1.82]$$

and so the tail of Φ_β decreases like a power law.

ii) We have the following equivalences:

$$H \prec \Phi_{\beta} \Leftrightarrow \ln X^{\beta} \prec \Delta \Leftrightarrow -X^{-1} \prec \Psi_{\beta} \qquad [1.83]$$

The parameter $E(Y|\Im_1) = E(Y|B)$ can be seen as a dispersion parameter and d_n as a localization parameter tending toward the mode, i.e. the maximum of the density function of the limit distribution. Gnedenko [GNE 43] characterized the three classes of the distribution function F of the considered risk called *attraction domains*; so if a risk has its distribution function in one of these three domains, we know what the limit distribution for this risk is.

To characterize these three attraction domains, let us introduce the concept of slowly varying function L.

Such a function with support $(0, \infty)$ is *slowly varying type* iff:

$$\lim_{x \to \infty} \frac{L(tx)}{L(x)} = 1, \forall t > 0 \qquad [1.84]$$

Moreover, if:

$$\lim_{x \to \infty} \frac{L(tx)}{L(x)} = t^{\alpha}, \forall t > 0 \qquad [1.85]$$

L is called *regularly varying at index α*.

We can now give the following characterization of d the three attraction domains (see [EMB 08]):

i) The *Gumbel attraction* domain contains the distribution functions F so that:

$$\exists x_0 :$$
$$F(x_0) = 1 \text{ et } F(x) < 1 \text{ for } x < x_0$$
$$\text{and a } z < x_0 \text{ such that :}$$

$$\text{ii) } \overline{F}(x) = c(x) \exp\left\{-\int_0^x \frac{g(t)}{a(t)} dt\right\}, z < x < x_0 \qquad [1.86]$$

where c and g are measurable functions such that

$\lim_{x \nearrow x_0} c(x) = c > 0$ and $\lim_{x \nearrow x_0} g(x) = 1 > 0$

and $a(x)$ a positive absolutely function with Lebesgue density a' such that $\lim_{x \nearrow x_0} a'(x) = 0$ as for example

$$a(x) = \int_x^{x_0} \frac{\overline{F}(t)}{\overline{F}(x)} dt, x < x_0.$$

Examples: normal law, exponential law, chi-square law, gamma law, lognormal law, Weibull distribution and laws with heavy queues decreasing to 0 faster than the exponential.

iii) The *Fréchet attraction domain* contains the distribution functions F so that $\overline{F}(x) = x^{-\alpha} L(x)$ with L slowly varying function:

$$\exists k > 0 : \lim_{x \to \infty} \frac{1 - F(x)}{1 - F(ux)} = u^k, \text{for all } u > 0. \qquad [1.87]$$

Examples: Student’s law, Cauchy’s law, Pareto’s law, laws with heavy queues decreasing to 0 slower than the exponential.