Lyon PhD Course Actuarial Science

Chapter 1 Context, motivations, and basics

Professor : Etienne Marceau

May 15, 2017

Contents

1 Introduction

In this chapiter, we provide a brief description of a portfolio of risks and its characteristics (expectation and variance).

Then, we introduce the risk measures VaR and TVaR (and their properties) and we explain their pertinence in actuarial science.

We also look at the risk pooling.

Finally, we provide a brief introduction to simple approximations to the distribution of sum of independent rvs.

One is the normal approximation, based on the Central Limit Theorem.

Then, we give a short introduction to Monte Carlo simulation methods which prove to be useful in non life actuarial models.

2 Random variables

Non-life actuarial models are defined in terms of rvs.

2.1 Definitions

In actuarial science, a risk is represented by a rv X :

- the number of claim for an insurance contract
- the amount of a claim
- the aggregate amount for the contract
- the aggregate amount for an insurance port

- the losses for an investment
- etc.

The cdf of a rv X is defined by

$$F_X(x) = \Pr(X \leq x).$$

The properties of F_X are the following :

- F_X is non-decreasing ;
- F_X is right semi-continuous ;
- $F_X(-\infty) = 0$ and $F_X(\infty) = 1$;
- $0 \leq F_X(x) \leq 1$, $x \in \mathbb{R}$.

The survival function of the rv X, denoted by $\overline{F}_X(x)$, is defined by

$$\overline{F}_X(x) = 1 - F_X(x).$$

There are two important classed of rvs:

- discrete rvs ;
- continuous rvs.

2.2 Discrete random variables

A rv is said to discrete if it takes a finite number of values or a counting number of values in the set $A_X = \{x_1, x_2, x_3, ...\}$.

Its pmf is defined by f_X where

$$f_X(x) = \Pr(X = x), \ x \in A_X \subset \mathbb{R},$$

with $0 \leq f_X(x) \leq 1, \ x \in A_X$ and $\sum_{x \in A_X} f_X(x) = 1$.

The support A_X and the pmf provide a complete description of the random behavior of X.

An important class of discrete rvs are defined over an arithmetical support $A_X = \{0, 1h, 2h, ...\}$ where h is strictly positive scalar ($h > 0$).

Main discrete distributions:

- Poisson
- Binomial
- Negative Binomial
- Geometric

2.3 Continuous random variables

A rv is said to be (absolutely) continuous if it exists a function $f_X(x) \geq 0$, defined over $\mathbb{R}$, such that

$$\Pr(X \in \mathbb{R}) = \int_{\mathbb{R}} f_X(x)\,\mathrm{d}x = 1$$

and

$$\Pr(X \in C) = \int_{C} f_X(x)\,\mathrm{d}x,$$

where C is subset of $\mathbb{R}$.

If $C =]-\infty, a]$, then

$$\Pr(X \in C) = \Pr(X \in]-\infty, a]) = \Pr(X \leq a) = F_X(a).$$

If $C =]a,b]$, then

$$\begin{aligned}\Pr(X \in C) &= \Pr(X \in]a,b]) = \Pr(a < X \leq b) \\ &= \Pr(X \leq b) - \Pr(X \leq a) = F_X(b) - F_X(a).\end{aligned}$$

The function f_X is called the pdf of X.

We have the following relation between F_X amd f_X :

$$F_X(x) = \Pr(X \leq x) = \Pr(X \in]-\infty, x]) = \int_{-\infty}^{x} f_X(y)\, \mathrm{d}y.$$

Also, we have

$$\frac{\mathrm{d}}{\mathrm{d}x} F_X(x) = \frac{\mathrm{d}}{\mathrm{d}x} \int_{-\infty}^{x} f_X(y)\, \mathrm{d}y = f_X(x).$$

Main continuous distributions:

- Exponential
- Gamma
- Erlang
- Normal (gaussian)
- Lognormal
- Etc.

3 Portfolio of risks – Individual components

3.1 Context

Non-life actuarial models are defined in terms of rvs.

We consider a portfolio of n risks $X_1, ..., X_n$ for an insurance company or a financial firm over a fixed period of time.

The cdf of the rv X_i is denoted by F_{X_i}, with $i = 1, 2, ..., n$.

We define the aggregate amount for the whole portfolio by the rv S_n where

$$S_n = X_1 + ... + X_n.$$

The cdf of the rv S is denoted by F_S.

We use risk measures in order to quantify the risk associated to each rvs X_1, ..., X_n and to the rv S.

For the moment, we don't go into details of the specific models for the rvs X_1, ..., X_n in the context of non-life actuarial science.

3.2 Expectation

Usually in actuarial science, we assume that the expectation of the rv X_i exists for $i = 1, ..., n$.

Let us first consider one risk, represented by the rv X.

Then, we consider the expectation of the rv S_n (sometime denoted without "n", simply S).

3.2.1 Discrete random variables

For a finite support, we have

$$E[X] = \sum_{i=1}^{m} x_i \Pr(X = x_i) = \sum_{i=1}^{m} x_i \; f_X(x_i).$$

For counting support $A_X = \{x_1, x_2, x_3, ...\} \in \mathbb{R}$,it becomes

$$E[X] = \sum_{x \in A_X} x \Pr(X = x) = \sum_{i=1}^{\infty} x_i \; f_X(x_i),$$

assuming convergence.

Examples :

- Poisson distribution: $E[X] = \lambda$
- Binomial distribution: $E[X] = nq$
- Negative binomial distribution: $E[X] = r\frac{1-q}{q}$.

3.2.2 Continuous random variables

We have

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) \mathrm{d}x.$$

Examples :

- Exponential distribution: $E[X] = \frac{1}{\beta}$
- Gamma distribution: $E[X] = \frac{\alpha}{\beta}$
- Lognormal distribution: $E[X] = \exp\left(\mu + \frac{1}{2}\sigma^2\right)$

3.2.3 Linear transform

Let $a, b \in \mathbb{R}$.

Let X be a discrete or a continuous rv.

Let the rv Y be defined as a linear transform of X, i.e. $Y = aX + b$.

Then, we have

$$E[Y] = E[aX + b] = E[aX] + b = aE[X] + b.$$

3.2.4 Pure premium

Let the rv X corresponds to the eventual total claim amount for an insurance contract.

Then $E[X]$ is often called the pure premium of the insurance contract.

3.2.5 Positive continuous random variable

Proposition 1 *Let X be a positive continuous rv for which the expectation exists. Then, we have*

$$E[X] = \int_0^\infty \overline{F}_X(x)\mathrm{d}x.$$

Proof. We have

$$E[X] = \int_0^\infty y f_X(y)\mathrm{d}y = \int_0^\infty \int_0^y f_X(y)\mathrm{d}x\mathrm{d}y$$

whihc becomes

$$E[X] = \int_0^\infty \int_x^\infty f_X(y)\mathrm{d}y\mathrm{d}x = \int_0^\infty \overline{F}_X(x)\mathrm{d}x.$$

$\blacksquare$

3.2.6 Positive discrete random variable

Proposition 2 *Let X be a discrete rv defined over $\{0, 1h, 2h, ...\}$ for which the expectation exists. Then, we have*

$$E[X] = h\sum_{k=0}^{\infty} \overline{F}_X(kh).$$

Proof. We have

$$E[X] = \sum_{j=0}^{\infty} jhf_X(jh) = h\sum_{j=0}^{\infty}\sum_{k=0}^{j} f_X(jh)$$

and it follows that

$$E[X] = h\sum_{k=0}^{\infty}\sum_{j=k+1}^{\infty} f_X(jh) = h\sum_{k=0}^{\infty} \overline{F}_X(kh).$$

■

3.3 Other expectations for a single rv

3.3.1 Expectation of a function

We examine the expectation of a function g of X, denoted by $E\left[g\left(X\right)\right]$, if it exists.

If $X \in A_X = \{x_1, x_2, x_3, ..., x_m\}$, then

$$E\left[g\left(X\right)\right] = \sum_{i=1}^{m} g\left(x_i\right) f_X(x_i).$$

More generally, if $X \in A_X = \{x_1, x_2, x_3, ...\}$, then

$$E[g(X)] = \sum_{i=1}^{\infty} g(x_i) f_X(x_i).$$

If X is a continuous rv with pdf f_X, then

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)\ f_X(x)\,\mathrm{d}x.$$

3.3.2 Moment of order n

Let $g(y) = y^n$, pour $n \in \mathbb{N}^+$.

We define $E[X^n]$ as the moment of order n of the rv X, if it exists.

If $X \in A_X = \{x_1, x_2, x_3, ..., x_m\}$, then

$$E[X^n] = \sum_{i=1}^{m} x_i^n f_X(x_i).$$

More generally, If $X \in A_X = \{x_1, x_2, x_3, ...\}$, then

$$E[X^n] = \sum_{i=1}^{\infty} x_i^n f_X(x_i).$$

If X is a continuous rv with pdf f_X, then

$$E[g(X)] = E[X^n] = \int_{-\infty}^{\infty} x^n \; f_X(x) \, \mathrm{d}x.$$

3.3.3 Moment generating function

The moment generating function (mgf) of a rv X is defined by

$$M_X(t) = E\left[\mathrm{e}^{tX}\right],$$

assuming that the expectation exixts for a set of values $t \neq 0$.

If X is a discrete rv with a finite or a countable support, we have

$$M_X(t) = E\left[\mathrm{e}^{tX}\right] = \sum_{i=1}^{m} \mathrm{e}^{tx_i} f_X(x_i)$$

or

$$M_X(t) = E\left[\mathrm{e}^{tX}\right] = \sum_{i=1}^{\infty} \mathrm{e}^{tx_i} f_X(x_i),$$

assuming that the sum converges of a set of values of $t \neq 0$.

If X is a continuous rv, then we have

$$M_X(t) = E\left[\mathrm{e}^{tX}\right] = \int_{-\infty}^{\infty} \mathrm{e}^{tx} f_X(x) \, \mathrm{d}x,$$

assuming that the integral exists for some values $t \neq 0$.

The mgf does not exist for all distributions.

Examples of continuous distributions for which the mgf exists:

- Exponential : $M_X(t) = \left(\frac{\beta}{\beta - t}\right)$
- Gamma : $M_X(t) = \left(\frac{\beta}{\beta - t}\right)^{\alpha}$
- Normal (Gaussian) : $M_X(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$

3.3.4 Probability generating function

Later, we will consider distribution for which the mfg does not exists.

As we shall see later, probability generating functions are also useful.

They are defined for discrete rvs.

Let X be a discrete rv defined on $\mathbb{N}$.

The probability generating function (pgf) of the rv X is defined by

$$P_X(t) = E\left[t^X\right] = \sum_{k=0}^{\infty} t^k f_X(k) \tag{1}$$

if the sum converges for a set of values of $t \neq 1$.

They are closely related to the mgfs.

Indeed, we have

$$P_X(t) = E\left[t^X\right] = E\left[\exp\left(X \ln(t)\right)\right] = M_X\left(\ln(t)\right).$$

Examples of pgfs:

- Poisson distribution : $P_X(t) = \exp\{\lambda(t-1)\}$
- Binomial distribution : $P_X(t) = (qt + 1 - q)^n$
- Negative binomial distribution : $P_X(t) = \left(\frac{q}{1-(1-q)t}\right)^r$

3.4 Variance of a single rv

We assume that the variance of the rv X_i exists for $i = 1, 2, ..., n$.

Let us first consider one insurance contract (or one risk), represented by the rv X.

Then, we will consider the variance of the rv S_n (or simply S).

3.4.1 Definition

Let X be a rv such that $E\left[X^2\right]$ exists.

The variance of X is defined by

$$\mathrm{Var}(X) = E\left[(X - E[X])^2\right]. \tag{2}$$

From (2), we find

$$\mathrm{Var}(X) = E\left[X^2\right] - E[X]^2. \tag{3}$$

3.4.2 Discrete rvs

If X is a discrete rv with a finite or a countable support, (2) becomes

$$\mathrm{Var}(X) = \sum_{i=1}^{m} (x_i - E[X])^2 f_X(x_i)$$

or

$$\mathsf{Var}\left(X\right)=\sum_{i=1}^{\infty}\left(x_i-E\left[X\right]\right)^2 f_X\left(x_i\right).$$

Examples :

- Poisson distribution: $Var\left(X\right)=\lambda$
- Binomial distribution: $Var\left(X\right)=nq\left(1-q\right)$
- Negative Binomial distribution: $Var\left(X\right)=r\frac{1-q}{q^2}$.

3.4.3 Continous rvs

If X is a continuous rv, then we have

$$\mathsf{Var}\left(X\right)=\int_{-\infty}^{\infty}\left(x-E\left[X\right]\right)^2 f_X\left(x\right)\mathrm{d}x.$$

Examples :

- Exponential distribution: $\mathrm{Var}(X) = \frac{1}{\beta^2}$
- Gamma distribution: $\mathrm{Var}(X) = \frac{\alpha}{\beta^2}$
- Lognormal distribution: $\mathrm{Var}(X) = \mathrm{e}^{2\mu+\sigma^2}\left(\mathrm{e}^{\sigma^2} - 1\right)$

3.4.4 Interpretation

The variance of X can be interpreted as a measure of variability of the rv X.

Also, it can be seen as measure of dispersion of the rv X about its expectation.

3.4.5 Linear transform

Let the rv Y be defined as a linear transform of X, i.e. $Y = aX + b$.

Then, we have

$$\mathsf{Var}(Y) = \mathsf{Var}(aX + b) = \mathsf{Var}(aX) = a^2 \mathsf{Var}(X).$$

4 Portfolio of risks – interaction between the components

4.1 Context

Recall that we consider a portfolio of n risks $X_1, ..., X_n$ for an insurance company or a financial firm over a fixed period of time.

The aggregate amount for the whole portfolio is defined by the rv S_n where

$$S_n = X_1 + ... + X_n.$$

The cdf of the rv S is denoted by F_S.

Now, we have a first look at the behavior of the rv S.

How does the interaction between the components influence the behavior of the rv S ?

4.2 Expectation of the aggregate claim amount

The expectation of the rv S, the aggregate claim amount rv, is given by

$$E[S] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]. \tag{4}$$

The relation in (4) is always true, whether the rvs (risks such as insurance contracts, etc.) $X_1, ..., X_n$ are independent or not.

The interaction between the rvs X_1, ..., X_n does not have an impact on the expectation of the rv S.

Interpretation : $E[S] =$ expected aggregate claim amount for the whole portfolio.

For example, the expected aggregate claim amount appears on the balance sheet of an insurance company.

4.3 Variance of S, a sum of rvs (first part)

The rv S is a sum of the rvs X_1, ..., X_n.

The rvs X_1, ..., X_n may be independent or dependent.

In the computation of the variance of S (sum of rvs), we need to take into account the linear correlation between all the pairs of rvs $\left(X_i X_j\right)$ with the set of rvs $\{X_1, ..., X_n\}$.

The linear correlation is measured with the covariance between the rvs X_i and X_j, $i \neq j \in \{1, 2, ..., n\}$.

First, we need to recall some notions about the multivariate distributions for vector of rvs and the correlation.

4.4 Joint distributions of n random variables

4.4.1 Definitions and properties

Let $\underline{X} = (X_1, ..., X_n)$ be vector of rvs.

The joint cdf of $\underline{X}$ is defined by

$$F_{\underline{X}}(\underline{x}) = \Pr(X_1 \leq x_1, ..., X_n \leq x_n).$$

First, we need to define the operator $\triangle_{a_i,b_i}$ where

$$\begin{aligned}\triangle_{a_i,b_i} F_{\underline{X}}(\underline{x}) &= F_{\underline{X}}(x_1, ..., b_i, ..., x_n) - F_{\underline{X}}(x_1, ..., a_i, ..., x_n) \\ &= \Pr(X_1 \leq x_1, ..., a_i < X_i \leq b_i, ..., X_n \leq x_n).\end{aligned}$$

It follows that

$$\begin{aligned}
\triangle_{a_j,b_j} \triangle_{a_i,b_i} F_{\underline{X}}(\underline{x}) &= \triangle_{a_j,b_j} F_{\underline{X}}(x_1,...,b_i,...,x_n) - \triangle_{a_j,b_j} F_{\underline{X}}(x_1,...,a_i,...,x_n) \\
&= F_{\underline{X}}(x_1,...,b_i,...,b_j...,x_n) - F_{\underline{X}}(x_1,...,b_i,...,a_j...,x_n) \\
&\quad - F_{\underline{X}}(x_1,...,a_i,...,b_j...,x_n) + F_{\underline{X}}(x_1,...,a_i,...,a_j...,x_n).
\end{aligned}$$

For $n = 2$, we write

$$\begin{aligned}
\Pr(\underline{a} < \underline{X} \leq \underline{b}) &= \Pr(\underline{X} \in (a_1, b_1] \times (a_2, b_2]) \\
&= \triangle_{a_1,b_1} \triangle_{a_2,b_2} F_{\underline{X}}(\underline{x}) \\
&= F_{\underline{X}}(b_1, b_2) - F_{\underline{X}}(a_1, b_2) \\
&\quad - F_{\underline{X}}(b_1, a_2) + F_{\underline{X}}(a_1, a_2).
\end{aligned}$$

Definition 3 *A multivariate cdf $F_{\underline{X}}$ is an application from $\mathbb{R}^n$ to $[0,1]$ such that :*

1. $F_{\underline{X}}$ *is non decreasing on* R^n *;*
2. $F_{\underline{X}}$ *is right-continuous on* R^n *;*
3. $\lim_{x_i \to -\infty} F_{\underline{X}}(x_1, ..., x_n) = 0$, *for* $i = 1, ..., n$ *;*
4. $\lim_{x_1 \to \infty, ..., x_n \to \infty} F_{\underline{X}}(x_1, ..., x_n) = 1$ *;*
5. *for all* x, *we have*

$$\triangle_{a_1,b_1} \cdots \triangle_{a_n,b_n} F_{\underline{X}}(\underline{x}) \geq 0 \tag{5}$$

such that

$$\Pr(\underline{a} < \underline{X} \leq \underline{b}) = \Pr(\underline{X} \in (a_1, b_1] \times ... \times (a_n, b_n]) \geq 0.$$

If $F_{\underline{X}}$ is differentiable, (5) is equivalent to

$$\frac{\partial^n}{\partial x_1 ... \partial x_n} F_{\underline{X}}(x_1, ..., x_n) \geq 0$$

on $\mathbb{R}^n$.

We define the marginal cdf of X_i by

$$F_{X_i}(x_i) = \Pr(X_i \leq x_i) = F_{X_1,...,X_n}(\infty, ..., \infty, x_i, \infty, ..., \infty),$$

for all $i = 1, ..., n$.

Frequently, we use the term "marginal" for "marginal cdf".

The joint survival function of $\underline{X}$ is defined by

$$\overline{F}_{X_1,...,X_n}(x_1, ..., x_n) = \Pr(X_1 > x_1, ..., X_n > x_n).$$

For $n = 2$ and $n = 3$, we have

$$\overline{F}_{X_1,X_2}(x_1, x_2) = 1 - F_{X_1}(x_1) - F_{X_2}(x_2) + F_{X_1,X_2}(x_1, x_2)$$

and

$$\begin{aligned}\overline{F}_{X_1,X_2,X_3}(x_1,x_2,x_3) &= 1-F_{X_1}(x_1)-F_{X_2}(x_2)-F_{X_3}(x_3)\\ &\quad +F_{X_1,X_2}(x_1,x_2)+F_{X_1,X_3}(x_1,x_3)\\ &\quad +F_{X_2,X_3}(x_2,x_3)-F_{X_1,X_2,X_3}(x_1,x_2,x_3).\end{aligned}$$

The marginal survival function of the rv X_i is defined by

$$\overline{F}_{X_i}(x_i)=\Pr(X_i>x_i)=\overline{F}_{X_1,\ldots,X_n}(0,\ldots,0,x_i,0,\ldots,0),$$

for $i=1,\ldots,n$.

The multivariate mgf is defined by

$$M_{X_1,\ldots,X_n}(t_1,\ldots,t_n)=E\left[\mathrm{e}^{t_1X_1}\ldots\mathrm{e}^{t_nX_n}\right],$$

if the expectation exists.

The rvs X_1, ..., X_n are independent if and only if we have

$$\begin{aligned} F_{X_1,...,X_n}(x_1,...,x_n) &= F_{X_1}(x_1)...F_{X_n}(x_n), \\ \overline{F}_{X_1,...,X_n}(x_1,...,x_n) &= \overline{F}_{X_n}(x_n)...\overline{F}_{X_1}(x_n). \end{aligned}$$

Also, if the rvs X_1, ..., X_n are independent, we have

$$E[g_1(X_1)...g_n(X_n)] = E[g_1(X_1)]...E[g_n(X_n)],$$

for all integrable functions $g_1, ..., g_n$.

For example, the expectation of the product of the rvs X_i and X_j becomes

$$E\left[X_i X_j\right] = E[X_i]\, E\left[X_j\right],$$

for $i \neq j \in \{1, ..., n\}$.

Dependence modeling is treated in details later.

4.4.2 Vector of discrete rvs

Let $\underline{X} = (X_1, ..., X_n)$ be vector of discrete rvs.

The multivariate pmf $(X_1, ..., X_n)$, denoted by $f_{X_1,...,X_n}$, is defined by

$$f_{X_1,...,X_n}(x_1, ..., x_n) = \Pr(X_1 = x_1, ..., X_n = x_n).$$

The function $f_{X_1,...,X_n}$ takes value in $[0, 1]$.

To simplify the presentation, we consider the special case where $X_i \in \{0, 1h, 2h, ...\}$, $i = 1, 2, ...$, and h is a strictly positive scalar.

We have the following expressions:

$$\begin{aligned} & E\left[g_1\left(X_1\right)...g_n\left(X_n\right)\right] \\ = & \sum_{m_1=0}^{\infty}...\sum_{m_n=0}^{\infty} g_1\left(m_1 h\right)...g_n\left(m_n h\right) f_{X_1,...,X_n}\left(m_1 h,...,m_n h\right) \end{aligned}$$

and

$$\begin{aligned} & E\left[g\left(X_1,...,X_n\right)\right] \\ = & \sum_{m_1=0}^{\infty}...\sum_{m_n=0}^{\infty} g_1\left(m_1 h,...,m_n h\right) f_{X_1,...,X_n}\left(m_1 h,...,m_n h\right). \end{aligned}$$

The multivariate mgf is given by

$$M_{X_1,...,X_n}\left(t_1,...,t_n\right) = \sum_{m_1=0}^{\infty}...\sum_{m_n=0}^{\infty} \mathrm{e}^{m_1 h t_1}...\mathrm{e}^{m_n h t_n} f_{X_1,...,X_n}\left(m_1 h,...,m_n h\right).$$

Also, we define the multivariate pgf by

$$P_{X_1,...,X_n}(t_1,...,t_n) = \sum_{m_1=0}^{\infty} ... \sum_{m_1=0}^{\infty} t_1^{m_1 h} ... t_n^{m_n h} f_{X_1,...,X_n}(m_1 h,...,m_n h).$$

For $n = 2$, we have the following relation for the bivariate mpf of (X_1, X_2) :

$$\begin{aligned} f_{X_1,X_2}(m_1 h, m_2 h) &= F_{X_1,X_2}(m_1 h, m_2 h) - F_{X_1,X_2}(m_1 h - h, m_2 h) \\ &\quad - F_{X_1,X_2}(m_1 h, m_2 h - h) \\ &\quad + F_{M_1,M_2 X_1,X_2}(m_1 h - h, m_2 h - h), \end{aligned}$$

for $(m_1, m_2) \in \mathbb{N} \times \mathbb{N}$ and with $F_{M_1,M_2}(m_1 h, m_2 h) = 0$ if $m_1 < 0$ or $m_2 < 0$.

4.4.3 Vector of continuous rvs

Let $(X_1, ..., X_n)$ be a vector of n continuous rvs.

The multivariate pdf is defined by

$$f_{X_1,...,X_n}(x_1, ..., x_n) = \frac{\partial n}{\partial x_1 ... \partial x_n} F_{X_1,...,X_n}(x_1, ..., x_n).$$

The pdf $f_{X_1,...,X_n}$ takes positive values which may be greater than 1 and it satisfies the property

$$\int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f_{X_1,...,X_n}(x_1, ..., x_n) \, \mathrm{d}x_1 ... \mathrm{d}x_n = 1.$$

An heuristic interpretation of the pdf $f_{X_1,\ldots,X_n}$ is

$$f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)\,\mathrm{d}x_1\ldots\mathrm{d}x_n \simeq \Pr\left(\bigcap_{i=1}^{n} X_i \in (x_i, x_i+\mathrm{d}x_i]\right).$$

Notably, we have

$$\Pr\left(\bigcap_{i=1}^{n} X_i \in (a_i, b_i]\right) = \int_{a_1}^{b_1}\ldots\int_{a_n}^{b_n} f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)\,\mathrm{d}x_1\ldots\mathrm{d}x_n,$$

$$E\left[\prod_{i=1}^{n} g_i(X_i)\right] = \int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}\prod_{i=1}^{n} g_i(x_i)\, f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)\,\mathrm{d}x_1\ldots\mathrm{d}x_n,$$

$$E\left[g(X_1,\ldots,X_n)\right] = \int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} g_1(x_1,\ldots,x_n)\, f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)\,\mathrm{d}x_1\ldots\mathrm{d}x_n$$

and

$$M_{X_1,\ldots,X_n}(t_1,\ldots,t_n) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \mathrm{e}^{t_1 x_1} \ldots \mathrm{e}^{t_n x_n} f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)\, \mathrm{d}x_1 \ldots \mathrm{d}x_n.$$

4.5 Covariance and linear correlation

4.5.1 Definition

Let (X, Y) be a pair of rvs.

We define the covariance between X and Y by

$$\text{Cov}(X, Y) = E\left[(X - E[X])(Y - E[Y])\right] = E[XY] - E[X]E[Y].$$

The covariance for the pair (X, Y) measures the linear dependence relation between these two rvs.

4.5.2 Discrete rvs

Let the rvs X and Y be discrete and countable.

Then, the expression for $E[XY]$ is given by

$$E[XY] = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} x_i \, y_j \, f_{X,Y}(x_i, y_j).$$

We also have

$$\mathsf{Cov}(X,Y) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} (x_i - E[X]) \left(y_j - E[Y]\right) f_{X,Y}(x_i, y_j).$$

4.5.3 Continuous rvs

Let the rvs X and Y be continuous.

Then, we obtain

$$E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy \ f_{X,Y}(x,y)\,\mathsf{d}x\mathsf{d}y.$$

Moreover, we have

$$\mathsf{Cov}(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - E[X])(y - E[Y]) \ f_{X,Y}(x,y)\,\mathsf{d}x\mathsf{d}y.$$

4.5.4 Interpretation

The covariance may be positive, zero, or negative.

A positive (negative) value means a positive (negative) dependence relation between X and Y.

4.5.5 Independence

When the rvs X and Y are independent, we have $E[XY] = E[X]E[Y]$ and it implies that the covariance is zero.

4.5.6 Uncorrelated rvs

However, as it is illustrated in the next example, a covariance of zero between the rvs X and Y does not implies that the rvs are independent.

Example 4 *Let* (X, Y) *be a pair of discrete rvs* (X, Y) *where* $X \in \{-1, 0, 1\}$ *and* $Y \in \{-2, 0, 2\}$.

In the following table, we provide the values of its pmf $f_{X,Y}$ *:*

$x \backslash y$	-2	0	2
-1	0	$\frac{1}{3}$	0
0	$\frac{1}{6}$	0	$\frac{1}{6}$
1	0	$\frac{1}{3}$	0

Clearly, $f_X(-1) = f_X(0) = f_X(1) = \frac{1}{3}$, $f_Y(-2) = f_Y(2) = \frac{1}{6}$ *and* $f_Y(0) = \frac{2}{3}$.

We find that $\mathrm{Cov}(X, Y) = 0$.

However, we have e.g. $f_{X,Y}(0, 0) \neq f_X(0) f_Y(0) = \frac{2}{9}$.

Then, the rvs X *and* Y *are not independent although the covariance between them is zero.* $\square$

If the covariance is 0, it means that the rvs X and Y are not linearly correlated.

4.5.7 Properties of the covariance

We provide below the properties of the covariance.

Let the rvs X, Y, Z, W and the scalars a, b, c, $d \in \mathbb{R}$.

Then, we have

$$\begin{aligned}
\mathrm{Cov}(X, X) &= \mathrm{Var}(X), \\
\mathrm{Cov}(X, Y) &= \mathrm{Cov}(Y, X), \\
\mathrm{Cov}(aX, bY) &= ab\mathrm{Cov}(X, Y), \\
\mathrm{Cov}(X + c, Y + d) &= \mathrm{Cov}(X, Y), \\
\mathrm{Cov}(aX + c, bY + d) &= ab\mathrm{Cov}(X, Y), \\
\mathrm{Cov}(X, Y + Z) &= \mathrm{Cov}(X, Y) + \mathrm{Cov}(X, Z), \\
\mathrm{Cov}(X + W, Y + Z) &= \mathrm{Cov}(X, Y) + \mathrm{Cov}(W, Y) \\
&\quad + \mathrm{Cov}(X, Z) + \mathrm{Cov}(W, Z).
\end{aligned}$$

4.5.8 Pearson coefficient of linear correlation

We define the Pearson coefficient of linear correlation ρ_P by

$$\rho_P(X,Y) = \frac{\mathsf{Cov}(X,Y)}{\sqrt{\mathsf{Var}(X)\,\mathsf{Var}(Y)}},$$

where

$$-1 \leq \rho_P(X,Y) \leq 1.$$

If the rvs X and Y are independent, then $\rho_P(X,Y) = 0$.

However, since ρ_P only provides a measure of the linear dependence between the rvs X and Y.

If $\rho_P(X,Y) = 0$, the rvs X and Y are said to be non linearly correlated.

4.5.9 Last comment

Dependence modeling and measures of dependence are very important in actuarial science.

They are treated in details in later.

4.6 Variance of S, a sum of rvs (second part)

4.6.1 Portfolio of rvs

Recall that the aggregate claim amount rv S is a sum of rvs X_1, ..., X_n.

As mentioned earlier, the rvs X_1, ..., X_n may be independent or dependent.

We want to measure the variability of the aggregate claim amount (rv S) of the entire portfolio.

We use the following property.

4.6.2 Notations

Let $\underline{X} = (X_1, ..., X_n)$ be vector of n rvs.

The variance-covariance matrix of the vector of random variables $\underline{X}$ is defined by

$$\begin{pmatrix} \mathsf{Var}(X_1) & \mathsf{Cov}(X_1, X_2) & \cdots & \mathsf{Cov}(X_1, X_n) \\ \mathsf{Cov}(X_2, X_1) & \mathsf{Var}(X_2) & \cdots & \mathsf{Cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathsf{Cov}(X_n, X_1) & \mathsf{Cov}(X_n, X_2) & \cdots & \mathsf{Var}(X_n) \end{pmatrix}.$$

The matrix of linear correlation contains the Pearson coefficients of linear correlation for all pairs within the vector of rvs $\underline{X}$:

$$\begin{pmatrix} 1 & \rho_P\left(X_1, X_2\right) & \cdots & \rho_P\left(X_1, X_n\right) \\ \rho_P\left(X_2, X_1\right) & 1 & \cdots & \rho_P\left(X_2, X_n\right) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_P\left(X_n, X_1\right) & \rho_P\left(X_n, X_1\right) & \cdots & 1 \end{pmatrix}.$$

4.6.3 Property

We define the rvs $S = \sum_{i=1}^{n} a_i X_i$ and $T = \sum_{i=1}^{n} b_i X_i$, where the scalars $a_1, ..., a_n$, $b_1, ..., b_n \in \mathbb{R}$.

Then, we have the following relations :

$$E[S] = \sum_{i=1}^{n} a_i E[X_i], \; E[T] = \sum_{i=1}^{n} b_i E[X_i],$$

$$\mathsf{Var}(S) = \sum_{i=1}^{n} a_i^2 \mathsf{Var}(X_i) + \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} a_i a_j \mathsf{Cov}\left(X_i, X_j\right),$$

$$\mathsf{Var}(T) = \sum_{i=1}^{n} b_i^2 \mathsf{Var}(X_i) + \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} b_i b_j \mathsf{Cov}\left(X_i, X_j\right)$$

and

$$\mathsf{Cov}\,(S,T)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_i b_j \mathsf{Cov}\left(X_i,X_j\right).$$

4.6.4 Variance of the aggregate claim amount

Clearly, the variance of the rv $S = X_1 + ... + X_n$ is

$$
\begin{aligned}
\mathsf{Var}(S) &= \mathsf{Var}\left(\sum_{i=1}^{n} X_i\right) \\
&= \sum_{i=1}^{n} \mathsf{Var}(X_i) + \sum_{\substack{i=1 \\ i \neq j}}^{n} \sum_{j=1}^{n} \mathsf{Cov}\left(X_i, X_j\right) \\
&= \sum_{i=1}^{n} \mathsf{Var}(X_i) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathsf{Cov}\left(X_i, X_j\right).
\end{aligned}
$$

5 Risk measures – motivations

In actuarial science and quantitative risk management, the two main applications for risk measures are the computation of the premiums for insurance contracts and for the economic capital for portfolios of an insurance company or a financial institution.

5.1 Premiums

Let the positive rv X corresponds to the aggregate claim amount covered by an insurance contract.

The risk measure is denoted $\varsigma_{\kappa}(X)$, where κ is a scalar to be defined later (often, the level of confidence).

The amount $\varsigma_{\kappa}(X)$ corresponds to the premium for the insurance contract

5.2 Economic Capital

We consider a portfolio of n policies or risks, represented by the (independent or dependent) rvs X_1

The aggregate claim amount for the policy i is defined by the rv X_i, for $i = 1, 2, ..., n$.

The aggregate claim amount for the entire portfolio is defined by the rv $S = \sum_{i=1}^{n} X_i$.

The total capital allocated to the entire portfolio is determined with $\varsigma_\kappa(S)$, where ς_κ designate a risk measure.

The index κ corresponds to the level of confidence (e.g. 99 % ou 99.5 %).

The economic capital can be seen as a "cushion" i.e.

$$CE_\kappa(S) = \varsigma_\kappa(S) - E[S],$$

which corresponds to a provision used to protect against an adverse experience of the portfolio.

These requirements are notably strongly recommended in Solvency 2 and Bale 2.

5.3 Two main risk measures: VaR and TVaR

The two main risk measures are the Value-at-Risk (VaR) and the Tail-Value-at-Risk (TVaR).

The additional ingredients which are required to define the VaR and the TVaR are the quantile function and the truncated expectation.

6 Ingredient #1 for risk measures – Quantile function

6.1 Basic Definition

We start with the basic definition of the quantile function.

Definition 5 *Let X be a rv with cdf F_X. The quantile function corresponds to the inverse function F_X^{-1} associated to F_X which is defined by*

$$F_X^{-1}(u) = \inf \{x \in \mathbb{R} : F_X(x) \geq u\} = \sup \{x \in \mathbb{R} : F_X(x) < u\},$$

for $u \in (0,1)$. By convention, $\inf \emptyset = +\infty$ and $\sup \emptyset = -\infty$.

Remark 6 *From Definition 5, we derive the following properties:*

Properties

1. F_X^{-1} *is non decreasing ;*
2. F_X^{-1} *is left semi-continuous ;*
3. $F_X^{-1}(F_X(x)) \leq x$ *;*
4. $F_X\left(F_X^{-1}(x)\right) \geq u$.

6.2 Continuous random variables

If X is a continuous rv, then F_X^{-1} corresponds to the unique value x_u such that $F_X(x_u) = u$.

For specific continuous probability distributions, an explicti expression for F_X^{-1} can be derived.

Example 8 *Let* $X \sim Exp(\beta)$ *with cdf* $F_X(x) = 1 - \mathrm{e}^{-\beta x}$*. The closed expression for* F_X^{-1} *is the solution* x *such that* $F_X(x) = u$*. We find* $F_X^{-1}(u) = -\frac{1}{\beta}\ln(1-u)$*.* $\square$

See Appendix for other distributions for which we can derive a closed-form expression for F_X^{-1}.

When no closed expression exists, one resorts to numerical optimization to compute values of F_X^{-1}.

For instance, for the mixture of exponentials, the cdf is

$$F_X(x) = 1 - \mathrm{e}^{-\beta x} \sum_{j=0}^{n-1} \frac{(\beta x)^j}{j!}.$$

Then, we need to use numerical optimization to find values of F_X^{-1}.

6.3 Discrete random variables

For discrete rvs, one only needs to apply the definition.

Example 9 *Let X be a discrete rv with*

$$\Pr(X=0)=0.2, \Pr(X=100)=0.3 \ \Pr(X=500)=0.35$$

et $\Pr(X=1000)=0.15$. *We find*

$$F_X(u)=\begin{cases} 0, & x<0 \\ 0.2, & 0\leq x<100 \\ 0.5, & 100\leq x<500 \\ 0.85, & 500\leq x<1000 \\ 1, & x\geq 1000 \end{cases} \quad \text{and } F_X^{-1}(u)=\begin{cases} 0, & 0\leq u\leq 0.2 \\ 100, & 0.2<u\leq 0.5 \\ 500, & 0.5<u\leq 0.85 \\ 1000, & 0.85<u\leq 1 \end{cases}.$$

6.4 Non standard cases

in the follwing example, we illustrate the derivation of F_X^{-1} when F_X is non standard.

Example 10 *Let X be rv with*

$$F_X(x) = \begin{cases} \left(\frac{x}{200}\right)^2, & 0 \leq x < 100 \\ 0.25, & 100 \leq x < 200 \\ 1 - \left(\frac{300}{300+x}\right)^2, & x \geq 200 \end{cases}.$$

Then, the expression for F_X^{-1} *is given by*

$$F_X^{-1}(u) = \begin{cases} 200\left(u^{\frac{1}{2}}\right), & 0 \leq u \leq 0.25 \\ 200, & 0.25 < u \leq 0.64 \\ 300\left((1-u)^{-\frac{1}{2}} - 1\right), & 0.64 \leq u < 1 \end{cases}.$$

□

6.5 Properties for the quantile function

6.5.1 Quantile Function Theorem

The following theorem is fundamental.

One of its applications is Monte Carlo simulation (see chapter **??**).

Theorem 11 Quantile Function Theorem. *Let* X *be a rv with cdf* F_X *and quantile function* F_X^{-1}*. Let* U *ba a rv such that* $U \sim U(0,1)$*. Then, the cdf of* $F_X^{-1}(U)$ *is* F_X*.*

Proof. We consider three cases.

First, we assume that the rv X is continuous, which implies

$$\Pr\left(F_X^{-1}(U) \leq x\right) = \Pr\left(U \leq F_X(x)\right),$$

since the events $\left\{F_X^{-1}(U) \leq x\right\}$ and $\{U \leq F_X(x)\}$ are equivalent. For the cdf of $U \sim U(0,1)$, we find

$$\Pr\left(F_X^{-1}(U) \leq x\right) = \Pr\left(U \leq F_X(x)\right) = F_X(x).$$

Secondly, we assume that the rv X is defined on $\mathbb{N}$. Then, for $k \in \mathbb{N}$, we have

$$\Pr\left(F_X^{-1}(U) = k\right) = \Pr\left(F_X(k-1) < U \leq F_X(k)\right),$$

since the event $\left\{F_X^{-1}(U)=k\right\}$ corresponds to the event $\{F_X(k-1)<U\leq F_X(k)\}$
Then, we find

$$\begin{aligned}\Pr\left(F_X^{-1}(U)=k\right) &= \Pr\left(F_X(k-1)<U\leq F_X(k)\right)\\ &= F_X(k)-F_X(k-1).\end{aligned}$$

The general case will be considered later. ■

6.5.2 Probability Integral Transform Theorem

Theorem 12 . Probability Integral Transform Theorem. *Let X be a continuous rv with cdf F_X, quantile function F_X^{-1}. Let $U \sim U(0,1)$. Then, $F_X(X) \sim U(0,1)$ (i.e. the rv $F_X(X)$ follows a standard uniform distribution).*

Proof. Important: $F_X(X)$ is a rv which takes value within $[0,1]$. For all $u \in (0,1)$, we develop the expression of the cdf of $F_X(X)$ which is given by

$$\Pr(F_X(X) \leq u) = \Pr\left(X \leq F_X^{-1}(u)\right) = F_X\left(F_X^{-1}(u)\right) = u,$$

which leads to the desired result. ■

6.5.3 Quantile function and increasing function

Proposition 13 *Let X be a rv. If φ is a strictly increasing continuous function, then we have*

$$F_{\varphi(X)}^{-1}(u) = \varphi\left(F_X^{-1}(u)\right),$$

for $u \in (0, 1)$.

Proof. See e.g. [**?**]. ■

Example 14 *Let X be a rv. We define the rv Y by*

$$Y = \mathrm{e}^X.$$

Then, we have

$$F_Y^{-1}(u) = F_{\mathrm{e}^X}^{-1}(u) = \mathrm{e}^{F_X^{-1}(u)}.$$

Example 15 *Let X be a rv. Let $\varphi(x) = \ln(x)$ and $Y = \ln(X)$. Then, we obtain*

$$F_Y^{-1}(u) = \ln\left(F_X^{-1}(u)\right).$$

6.5.4 Quantile function and decreasing function

Proposition 16 *Let X be a* **continuous** *rv. If φ is a strictly decreasing continuous function, then we have*

$$F_{\varphi(X)}^{-1}(u) = \varphi\left(F_X^{-1}(1-u)\right),$$

for $u \in (0, 1)$.

Proof. See e.g. [**?**]. ■

7 Ingredient #2 for risk measures – Truncated Expectation

7.1 Definition

Let g be a function such that $g(y) = y \times 1_{\{y \leq d\}}$.

We define the truncated expectation of X by $E\left[X \times 1_{\{X \leq d\}}\right]$.

We also have

$$E\left[X \times 1_{\{X > d\}}\right] = E[X] - E\left[X \times 1_{\{X \leq d\}}\right].$$

If X is a discrete rv with a finite support, we have

$$E\left[X \times 1_{\{X \leq d\}}\right] = \sum_{i=1}^{m} x_i 1_{\{x_i \leq d\}} f_X(x_i).$$

If X is a discrete rv with a countable support, we have

$$E\left[X \times 1_{\{X \leq d\}}\right] = \sum_{i=1}^{\infty} x_i 1_{\{x_i \leq d\}} f_X(x_i),$$

assuming convergence.

If X is a continuous rv, $E\left[X \times 1_{\{X \leq d\}}\right]$ is

$$E\left[X \times 1_{\{X \leq d\}}\right] = \int_{-\infty}^{d} x f_X(x)\,\mathrm{d}x.$$

We develop the expressions for

$$E\left[X \times 1_{\{X \leq d\}}\right]$$

and

$$E\left[X \times 1_{\{X > d\}}\right]$$

for specific continuous probability distributions.

7.2 Examples

7.2.1 Example – Exponential distribution

Let $X \sim Exp(\beta)$ with pdf $f_X(x) = \beta \mathrm{e}^{-\beta x}$, $x \in \mathbb{R}^+$.

Then, we have

$$E\left[X \times 1_{\{X \leq d\}}\right] = \int_0^d x f_X(x)\,\mathrm{d}x = \int_0^d x\beta \mathrm{e}^{-\beta x}\mathrm{d}x = -d\mathrm{e}^{-\beta d} + \frac{1}{\beta}\left(1 - \mathrm{e}^{-\beta d}\right).$$

From (**??**), we also find

$$E\left[X \times 1_{\{X > d\}}\right] = \int_d^\infty x f_X(x)\,\mathrm{d}x = \frac{1}{\beta} + d\mathrm{e}^{-\beta d} - \frac{1}{\beta}\left(1 - \mathrm{e}^{-\beta d}\right) = d\mathrm{e}^{-\beta d} + \frac{1}{\beta}\mathrm{e}^{-\beta d}. \qquad (6)$$

7.2.2 Example – Gamma distribution

Let $X \sim Ga(\alpha,\beta)$ with pdf $\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-\beta x}$, $x \in \mathbb{R}^{+}$, cdf denoted by $H(x;\alpha,\beta)$ and sf, $\overline{H}(x;\alpha,\beta)$. Then, we find

$$\begin{aligned} E\left[X \times 1_{\{X\leq d\}}\right] &= \int_0^d x\frac{\beta^{a}}{\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-\beta x}\mathrm{d}x \\ &= \frac{\alpha}{\beta}\int_0^d \frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)}x^{\alpha+1-1}\mathrm{e}^{-\beta x}\mathrm{d}x \\ &= \frac{\alpha}{\beta}H(d;\alpha+1,\beta) = E[X]\,H(d;\alpha+1,\beta). \end{aligned}$$

It follows that

$$\begin{aligned} E\left[X \times 1_{\{X>d\}}\right] &= E[X] - E\left[X \times 1_{\{X \le d\}}\right] \\ &= E[X]\left(1 - H(d; \alpha+1, \beta)\right) \\ &= E[X]\,\overline{H}(d; \alpha+1, \beta). \end{aligned} \tag{7}$$

7.2.3 Example – Normal distribution

Let $X \sim N\left(\mu, \sigma^2\right)$ with pdf $\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $x \in \mathbb{R}$.

We use the relation $X = \mu + \sigma Z$, where the rv Z follows a standard normal probability distribution i.e. $Z \sim N(0,1)$, where F_Z is denoted by Φ and the quantile function is denoted by Φ^{-1}.

It implies that

$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

The expression for $E\left[X \times 1_{\{X \leq d\}}\right]$ is given by

$$
\begin{aligned}
E\left[X \times 1_{\{X \leq d\}}\right] &= \int_{-\infty}^{d} x \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x \\
&= \int_{-\infty}^{d} (\mu + x - \mu) \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x \\
&= \int_{-\infty}^{d} \mu \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x \\
&\quad + \int_{-\infty}^{d} (x - \mu) \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x \\
&= \mu \Phi\left(\frac{d-\mu}{\sigma}\right) + \int_{-\infty}^{d} \left(-2\frac{(x-\mu)}{2\sigma^2}\right) (-1) \frac{1}{2} 2\sigma^2 \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} \\
&= \mu \Phi\left(\frac{d-\mu}{\sigma}\right) - \sigma \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{(d-\mu)^2}{2\sigma^2}}.
\end{aligned}
$$

It follows that

$$\begin{aligned} E\left[X \times 1_{\{X>d\}}\right] &= E[X] - E\left[X \times 1_{\{X\leq d\}}\right] \\ &= \mu - \left(\mu\Phi\left(\frac{d-\mu}{\sigma}\right) - \sigma\frac{1}{\sqrt{2\pi}}e^{-\frac{(d-\mu)^2}{2\sigma^2}}\right) \\ &= \mu\overline{\Phi}\left(\frac{d-\mu}{\sigma}\right) + \sigma\frac{1}{\sqrt{2\pi}}e^{-\frac{(d-\mu)^2}{2\sigma^2}}. \end{aligned} \tag{8}$$

7.2.4 Example – Lognormal distribution

Let $X \sim LN\left(\mu, \sigma^2\right)$. Since $X = \mathrm{e}^Y$ where $Y \sim N\left(\mu, \sigma^2\right)$, then

$$F_X(x) = F_Y(\ln(x)) = \Phi\left(\frac{\ln(x) - \mu}{\sigma}\right).$$

We find

$$
\begin{aligned}
E\left[X \times 1_{\{X \leq d\}}\right] &= E\left[\mathrm{e}^{Y} \times 1_{\left\{\mathrm{e}^{Y} \leq d\right\}}\right] \\
&= E\left[\mathrm{e}^{Y} \times 1_{\{Y \leq \ln(d)\}}\right] \\
&= \int_{-\infty}^{\ln(d)} \mathrm{e}^{x} \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x \\
&= \int_{-\infty}^{\ln(d)} \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{\left(x^2-2\mu x+\mu^2\right)+x2\sigma^2}{2\sigma^2}} \mathrm{d}x \\
&= \int_{-\infty}^{\ln(d)} \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2-2x\left(\mu+\sigma^2\right)+\mu^2+2\sigma^2\mu+\left(\sigma^2\right)^2-\left(2\sigma^2\mu+\left(\sigma^2\right)^2\right)}{2\sigma^2}} \mathrm{d}x \\
&= \\
&= \exp(\mu+\sigma^2/2) \int_{-\infty}^{\ln(d)} \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{\left(x-\left(\mu+\sigma^2\right)\right)^2}{2\sigma^2}} \mathrm{d}x
\end{aligned}
$$

which becomes

$$E\left[X \times 1_{\{X \leq d\}}\right] = \exp(\mu + \sigma^2/2)\Phi\left(\frac{\ln(d) - \mu - \sigma^2}{\sigma}\right).$$

Also, we have

$$E\left[X \times 1_{\{X > d\}}\right] = E[X]\,\overline{\Phi}\left(\frac{\ln(d) - \mu - \sigma^2}{\sigma}\right). \tag{9}$$

8 Link between quantiles and expectations

8.1 Quantile function and expectation

Proposition 17 *Let X be a rv with cdf F_X, quantile function F_X^{-1} and for which the expectation exists. Then, we have*

$$\begin{aligned}
\int_{\kappa}^{1} F_X^{-1}(u)\,\mathrm{d}u &= E\left[X \times 1_{\left\{X > F_X^{-1}(\kappa)\right\}}\right] \\
&\quad + F_X^{-1}(\kappa)\left(F_X\left(F_X^{-1}(\kappa)\right) - \kappa\right) \qquad (10)\\
\int_{0}^{\kappa} F_X^{-1}(u)\,\mathrm{d}u &= E\left[X \times 1_{\left\{X \leq F_X^{-1}(\kappa)\right\}}\right] \\
&\quad + F_X^{-1}(\kappa)\left(\kappa - F_X\left(F_X^{-1}(\kappa)\right)\right), \qquad (11)
\end{aligned}$$

which implies that

$$\int_0^1 F_X^{-1}(u)\,\mathrm{d}u = E[X].$$

Proof. The result is obvious for a discrete or a continuous rv. For the general case, see e.g. [**?**]. ■

8.2 Quantile function and expectation – Example

We illustrate the previous result in the following example.

Let X be a discrete rv with

x	$\Pr(X=x)$	$F_X(x)$
0	0.3	0.3
100	0.2	0.5
200	0.25	0.75
500	0.15	0.90
1000	0.1	1

We observe that

$$F_X^{-1}(0.8) = 500.$$

Since F_X^{-1} is a scale function, we obtain

$$\begin{aligned}
\int_{0.8}^{1} F_X^{-1}(u)\,\mathrm{d}u &= 500\,(0.9-0.8) + 1000 \times (1-0.9) \\
&= F_X^{-1}(0.8) \times (F_X(800) - 0.8) \\
&\quad +1000 \times \Pr(X = 1000) \\
&= F_X^{-1}(0.8) \times \left(F_X\left(F_X^{-1}(0.8)\right) - 0.8\right) \\
&\quad +E\left[X \times 1_{\left\{X > F_X^{-1}(0.8)\right\}}\right].
\end{aligned}$$

Also, we observe

$$\begin{aligned}\int_0^1 F_X^{-1}(u)\,\mathrm{d}u &= 0\times(0.3-0)\\ &\quad +100\times(0.5-0.3)\\ &\quad +200\times(0.75-0.5)\\ &\quad +500\times(0.9-0.75)\\ &\quad +1000\times(1-0.9)\\ &= \sum_{i=1}^{5} x_i \Pr(X=x_i)\\ &= E[X].\end{aligned}$$

8.3 Quantile function and expectation – additional results

Corollary 18 *Let X be a discrete rv defined on the finite support $A = \{x_1, ..., x_m\}$. Then, we have*

$$\int_0^1 F_X^{-1}(u)\,\mathsf{d}u = \sum_{i=1}^{m} x_i \Pr(X = x_i) = E[X].$$

Corollary 19 *When the rv X is continuous, $F_X\left(F_X^{-1}(u)\right) = u$. Then, (10) and (11) become*

$$\int_\kappa^1 F_X^{-1}(u)\,\mathsf{d}u = E\left[X \times 1_{\left\{X > F_X^{-1}(\kappa)\right\}}\right] \quad (12)$$

$$\int_0^\kappa F_X^{-1}(u)\,\mathsf{d}u = E\left[X \times 1_{\left\{X \leq F_X^{-1}(\kappa)\right\}}\right]. \quad (13)$$

Also, we have

$$\int_0^1 F_X^{-1}(u)\,\mathrm{d}u = E[X].$$

Example 20 *Let $X \sim Exp(\beta)$. We obtain*

$$\begin{aligned}
\int_0^1 F_X^{-1}(u)\,\mathrm{d}u &= -\frac{1}{\beta}\int_0^1 \ln(1-u)\,\mathrm{d}u \\
&= -\frac{1}{\beta}\left(-(1-u)\ln(1-u)\big|_0^1 - \int_0^1 \frac{(1-u)}{1-u}\right)\mathrm{d}u \\
&= -\frac{1}{\beta}(-0+0-1) \\
&= \frac{1}{\beta}.
\end{aligned}$$

9 Risk measures – First steps

Now, we are ready to present the two main risk measures used in actuarial science and quantitative risk management :

- Value at Risk (VaR)
- Tail Value at Risk (TVaR)

We first consider an individual risk X.

9.1 Value at Risk

The risk measure called *Value at Risk* (VaR) is very popular.

Definition 21 *Let κ be the level of confidence such that* $0 < \kappa < 1$. *The VaR is defined by*

$$VaR_{\kappa}(X) = F_X^{-1}(\kappa).$$

The VaR does not give any information of the behavior of the rv X beyond $VaR_{\kappa}(X)$.

For example, if the rv X is continuous and for $\kappa = 99.99\%$, it means that the probability that the rv X exceeds $VaR_{99.99\%}(X)$ is equal to 0.01 %.

However, it does not specfiy to which extent X may exceed the VaR.

9.2 Tail Value at Risk

The *Tail Value at Risk* (TVaR) is proposed as an alternative to the VaR.

Definition 22 *Let κ be the level of confidence such that $0 < \kappa < 1$. The TVaR is defined by*

$$TVaR_{\kappa}(X) = \frac{1}{1-\kappa}\int_{\kappa}^{1} VaR_{u}(X)\,\mathrm{d}u, \tag{14}$$

with $TVaR_{0}(X) = \int_{0}^{1} VaR_{u}(X)\,\mathrm{d}u = E[X]$.

Using (10) with (14), we find

$$TVaR_{\kappa}(X) = \frac{E\left[X \times 1_{\{X>VaR_{\kappa}(X)\}}\right] + VaR_{\kappa}(X)\left(F_{X}\left(VaR_{\kappa}(X)\right) - \kappa\right)}{1-\kappa}, \tag{15}$$

where $E\left[X\times 1_{\{X>VaR_\kappa(X)\}}\right]$ is directly derived or obtained with

$$E\left[X\times 1_{\{X>VaR_\kappa(X)\}}\right]=E\left[X\right]-E\left[X\times 1_{\{X\leq VaR_\kappa(X)\}}\right].$$

According to (14), the risk measure $TVaR_\kappa\left(X\right)$ can be considered as the arithmetical mean of the risk measures $VaR_u\left(X\right)$ for values of u greater than κ.

While the VaR indicates a single point in the tail of a distribution, the TVaR provides a better knowledge of the behavior of the tail.

If the rv X is continuous, then $F_X\left(VaR_\kappa\left(X\right)\right)-\kappa=0$ and it follows that

$$TVaR_\kappa\left(X\right)=\frac{E\left[X\times 1_{\{X>VaR_\kappa(X)\}}\right]}{1-\kappa}. \tag{16}$$

We also provide the definition of the *Conditional Tail Expectation* (CTE).

9.3 Conditional Tail Expectation

Definition 23 *Let κ be the level of confidence such that $0 < \kappa < 1$. The CTE is defined by*

$$CTE_{\kappa}(X) = E\left[X | \, X > VaR_{\kappa}(X)\right],$$

which can be written as follows:

$$CTE_{\kappa}\left(X\right) = \frac{E\left[X \times 1_{\{X > VaR_{\kappa}(X)\}}\right]}{\Pr\left(X > VaR_{\kappa}\left(X\right)\right)}.$$

Interpretation: The CTE corresponds to the expectation of the rv X when the rv X takes value greater than the VaR.

We have $CTE_0(X) = E\left[X\right]$.

9.4 TVaR vs CTE : TVaR is better

We analyze the relation between the TVaR and the CTE.

If the rv X is continuous, then

$$TVaR_{\kappa}(X) = \frac{E\left[X \times 1_{\{X > VaR_{\kappa}(X)\}}\right]}{1-\kappa}.$$

Also, since

$$\Pr(X > VaR_{\kappa}(X)) = \overline{F}_X(VaR_{\kappa}(X)) = 1-\kappa,$$

we obtain

$$TVaR_{\kappa}(X) = \frac{E\left[X \times 1_{\{X > VaR_{\kappa}(X)\}}\right]}{1-\kappa} = CTE_{\kappa}(X). \tag{17}$$

When the rv X is continuous, the CTE is equal to the TVaR.

However, this is generally not the case.

In fact, the TVaR was initially proposed according to the Definition 21 and (17) has lead to the definition of the CTE.

To avoid any confusion, we only use the TVaR and we will not use the CTE.

In quantitative risk management, the TVaR is often called *expected-shortfall*.

9.5 Examples

9.5.1 Example – Discrete distribution

Let X be a discrete rv with support

$$\{0, 100, 200, 500, 1000\}$$

and the following values of pmf :

x	0	100	200	500	1000
$f_X(x)$	0.3	0.4	0.15	0.10	0.05
$F_X(x)$	0.3	0.7	0.85	0.95	1.00

The value of $E[X]$ is 170.

In the following table, we provide the values of $VaR_\kappa(X)$, $TVaR_\kappa(X)$ and $CTE_\kappa(X)$ for $\kappa = 0.22$, 0.3, 0.39, 0.8501, 0.95 and 0.9999 :

κ	0.22	0.3	0.39	0.8501	0.95	0.9999
$VaR_\kappa(X)$	0	0	100	500	500	1000
$TVaR_\kappa(X)$	217.9487	242.8571	263.9344	666.7779	1000	1000
$CTE_\kappa(X)$	242.8571	242.8571	433.3333	1000	1000	not defined

In the case of $VaR_{0.39}(X) = 100$, we obtain

$$\begin{aligned} TVaR_{0.39}(X) &= \frac{E[X \times 1_{\{X > VaR_{0.39}(X)\}}]}{1-0.39} \\ &\quad + \frac{VaR_{0.39}(X)(F_X(VaR_{0.39}(X)) - 0.39)}{1-0.39} \\ &= \frac{200 \times 0.15 + 500 \times 0.1 + 1000 \times 0.05 + 100(0.7-0.39)}{1-0.39} \\ &= 263.934426 \end{aligned}$$

and

$$
\begin{aligned}
CTE_{0.39}(X) &= E[X|X > VaR_{0.39}(X)] \\
&= \sum_{x\in\{200,500,1000\}} \frac{x \Pr(X = x)}{\Pr(X > 100)} \\
&= \frac{200 \times 0.15 + 500 \times 0.1 + 1000 \times 0.05}{1 - 0.7} \\
&= 433.333333.
\end{aligned}
$$

We observe that the value of $CTE_{\kappa}(X)$ for $\kappa > 0.95$ is not defined.

It provides an illustration that it is better to use the TVaR rather than the CTE. $\square$

9.5.2 Example – Exponential distribution

Let $X \sim Exp(\beta)$.

Then,

$$VaR_\kappa(X) = -\frac{1}{\beta}\ln(1-\kappa).$$

If $d = VaR_\kappa(X) = -\frac{1}{\beta}\ln(1-\kappa)$ in (6), then $e^{-\beta VaR_\kappa(X)} = 1-\kappa$.

We replace the latter in (16) and we find

$$TVaR_\kappa(X) = \frac{E\left[X \times 1_{\{X>VaR_\kappa(X)\}}\right]}{1-\kappa} = VaR_\kappa(X) + \frac{1}{\beta}.$$

9.5.3 Example – Gamma distribution

Let $X \sim Ga(\alpha, \beta)$.

There is no closed-form for the VaR. Its value is computed by optimization.

In R, the quantile function is qgamma().

From (7) and (16), we obtain

$$TVaR_\kappa(X) = \frac{E\left[X \times 1_{\{X > VaR_\kappa(X)\}}\right]}{1-\kappa} = \frac{E[X]\,\overline{H}\left(VaR_\kappa(X);\alpha+1,\beta\right)}{1-\kappa}.$$

9.5.4 Example – Normal (gaussian) distribution

Let $X \sim N\left(\mu, \sigma^2\right)$.

We have $X = \mu + \sigma Z$, where $Z \sim N(0,1)$, with F_Z and F_Z^{-1} are respectively denoted by Φ and Φ^{-1}.

Then, the VaR de X is

$$VaR_\kappa(X) = \mu + \sigma VaR_\kappa(Z) = \mu + \sigma\Phi^{-1}(\kappa).$$

There is no closed form for $\Phi^{-1}(\kappa)$, but it is coded in e.g. `R` et `Excel`®.

With $d = VaR_\kappa(X) = \mu + \sigma\Phi^{-1}(\kappa)$ in (8) and (16), we find

$$
\begin{aligned}
TVaR_\kappa(X) &= \frac{E\left[X \times 1_{\{X > VaR_\kappa(X)\}}\right]}{1-\kappa} \\
&= \frac{\mu\overline{\Phi}\left(\frac{\mu+\sigma F_Z^{-1}(\kappa)-\mu}{\sigma}\right) + \sigma\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{\left(\mu+\sigma F_Z^{-1}(\kappa)-\mu\right)^2}{2\sigma^2}}}{1-\kappa}
\end{aligned}
$$

which becomes

$$
TVaR_\kappa(X) = \mu + \frac{1}{1-\kappa}\sigma\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{\left(F_Z^{-1}(\kappa)\right)^2}{2}}. \tag{18}
$$

From (18), we find

$$TVaR_{\kappa}(Z) = \frac{1}{1-\kappa}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{\left(F_Z^{-1}(\kappa)\right)^2}{2}},$$

which allows to obtain

$$TVaR_{\kappa}(X) = \mu + \sigma TVaR_{\kappa}(Z),$$

confirming Properties 27 and 28 for the TVaR.

9.5.5 Example – Lognormal distribution

Let $X \sim LN(\mu, \sigma)$. Since $X = e^Y$ where $Y \sim N\left(\mu, \sigma^2\right)$, we derive $VaR_\kappa(X) = \mathrm{e}^{\mu+\sigma\Phi^{-1}(\kappa)}$by Proposition 24.

Applying jointly (9) and (16) with $d = VaR_\kappa(X)$, we obtain

$$
\begin{aligned}
TVaR_\kappa(X) &= \frac{E[X]\,\overline{\Phi}\left(\frac{\ln(d)-\mu-\sigma^2}{\sigma}\right)}{1-\kappa} = \frac{E[X]\,\overline{\Phi}\left(\frac{\mu+\sigma\Phi^{-1}(\kappa)-\mu-\sigma^2}{\sigma}\right)}{1-\kappa} \\
&= \frac{E[X]\,\overline{\Phi}\left(\Phi^{-1}(\kappa)-\sigma\right)}{1-\kappa}.
\end{aligned}
$$

9.6 Properties

9.6.1 VaR and increasing function

The following proposition is important for the VaR.

Proposition 24 *Let X be a rv. For an increasing continuous function φ of X, we have*

$$VaR_{\kappa}\left(\varphi\left(X\right)\right)=\varphi\left(VaR_{\kappa}\left(X\right)\right).$$

Proof. The result follows from Proposition 16. ■

Here is an illustration of Proposition 24.

Example 25 *Let.* X *be a rv. We define* $Y = \mathrm{e}^{X}$. *Then,* $VaR_{\kappa}(Y) = \mathrm{e}^{VaR_{\kappa}(X)}$, *since* $\varphi(x) = \mathrm{e}^{x}$ *is an increasing continuous function.* $\square$

9.6.2 VaR and decreasing function

Proposition 26 *Let X be a* **continuous** *rv. If φ is a strictly decreasing continuous function, then we have*

$$VaR_\kappa(\varphi(X)) = \varphi(VaR_{1-\kappa}(X)),$$

for $u \in (0,1)$.

Proof. See e.g. [**?**]. ■

9.6.3 Positive Homogeneity and Translation invariant

From Proposition 24, the two following properties are deducted.

Property 27 Positive Homogeneity (Invariance to the multiplication by positive scalar). *For a scalar $a \in \mathbb{R}^+$, we have*

$$VaR_\kappa(aX) = aVaR_\kappa(X)$$

and

$$TVaR_\kappa(aX) = aTVaR_\kappa(X).$$

Proof. Since $\varphi(x) = ax$ is increasing and continuous, Proposition 24 leads to the first result. The second result follows from the definition of the TVaR and the first result. ■

Property 28 Translation invariant. *For a scalar $b \in \mathbb{R}$, we have*

$$VaR_{\kappa}(X+b) = VaR_{\kappa}(X) + b$$

and

$$TVaR_{\kappa}(X+b) = TVaR_{\kappa}(X) + b.$$

Proof. The function $\varphi(x) = x + b$ is increasing and continuous. Then, by Proposition 24, we obtain the first result. From the definition of the TVaR with the first result, we obtain the second result. ■

These two properties are treated in detailed in Chapter 3.

We derive the VaR and the TVaR for specifc continuous distribution.

10 Risk measures and sum of rvs

10.1 Motivation

Now, we need to compute the VaR and the TVaR associated to the aggregate claim amount for the whole portfolio, represented by the rv S.

Recall that the rv S is a sum of the (independent or dependent) rvs X_1, ..., X_n.

We need to know the distribution of the rv S i.e. we need to identify or approximate F_S.

Most of the times, it is not an easy task.

That's one of the strong abilities of the actuaries.

Recall that the main objective is to compute F_S.

An accurate knowledge of F_S is crucial to compute the popular risk measures VaR and TVaR, as well as risk quantities such as stop-loss premiums.

For specific contexts, we can find closed-form expressions for F_S.

In other contexts, we can identify the distribution of S.

Generally, we need to use approximation methods in order to compute F_S and the associated risk measures.

In actuarial science and quantitative risk management, a number of situations arise where the moments of a distribution can be rather easily obtained but the analytical form of this distribution can not be derived.

As an example, one may think of the risk quantification involving sums of dependent random variables for which only in rare cases can a closed-form expression be found for risk measures and risk quantities.

It implies that one has to use approximation methods which have been suggested over the years to provide some guidance as to the level of the risk assessment.

Among others, we may roughly classify those methods in the following three categories:

- moments-based approximations ;
- methods based on Monte-Carlo simulations ;
- recursive and numerical methods.

10.2 Sum of two independent rvs

10.2.1 Introduction

Let $S = \sum_{i=1}^{n} X_i$, where X_1, ..., X_n are independent rvs.

By the property of linearity of the expectation, we have

$$E[S] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i].$$

Since the rvs X_1, ..., X_n are independent, it implies that

$$\mathsf{Var}(S) = \mathsf{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathsf{Var}(X_i).$$

In actuarial science and in quantitative risk management, the distribution of S is of great importance.

In forecoming chapers, we treat this important subject.

10.2.2 Sum of 2 discrete independent rvs

Let X_1 and X_2 be two independent rvs defined on $\{0, 1h, 2h, 3h, ...\}$ wiht $f_{X_1}(kh) = \Pr(X_1 = kh)$ and $f_{X_2}(kh) = \Pr(X_2 = kh)$, for $k \in \mathbb{N}$.

We want to compute the pmf of $S = X_1 + X_2$, denoted by $f_S(kh) = \Pr(S = kh)$ for $k \in \mathbb{N}$.

Conditioning on the possible values of X_1 and due to the assumption of independence between X_1 and X_2, we find

$$f_S(kh) = \sum_{j=0}^{k} f_{X_1}(jh) f_{X_2}((k-j)h). \tag{19}$$

This operation is called the convolution product of f_{X_1} and f_{X_2}, which is denoted by $f_{X_1} * f_{X_2}$.

Then, we have $f_S(kh) = f_{X_1} * f_{X_2}(kh)$, $k \in \mathbb{N}$.

Example 29 *Let X_1 and X_2 be two independent rvs such that $X_1 \sim Bin(5, 0.3)$ and $X_2 \sim Bin(10, 0.2)$. We define $S = X_1 + X_2$. Using (19) with $h = 1$,* *the obtained values of* $f_S(k)$ *for* $k = 0, 1, ..., 13$ *are given in the following table :*

k	0	1	2	3	4	5	6
$f_S(k)$	0.01805	0.08379	0.18058	0.23967	0.21909	0.14614	0.07348
k	7	8	9	10	11	12	13
$f_S(k)$	0.02837	0.00848	0.00196	0.00035	0.00005	0.00001	0.00000

Also, $f_S(k) = 0$*, for* $k = 14, 15$*. The rv* S *does not follow a known discrete distribution. It is possible to verify that* $E[S] = \sum_{k=0}^{15} k f_S(k) = 3.5$ *which is also equal to* $E[X_1] + E[X_2] = 1.5 + 2 = 3.5$. $\square$

10.2.3 Sum of 2 independent continuous rvs

We consider two independent positive continuous rvs X_1 and X_2 with pdfs f_{X_1} and f_{X_2} respectively.

The expression of f_S where $S = X_1 + X_2$ is given by

$$f_S(x) = \int_0^x f_{X_1}(y) f_{X_2}(x-y)\,\mathrm{d}y,$$

which is obtained by conditioning on X_1 and due to the assumption of independence between X_1 and X_2.

Note that f_S corresponds to the convolution product of f_{X_1} and f_{X_2}, which is denoted by $f_S(x) = f_{X_1} * f_{X_2}(x)$, $x \in \mathbb{R}^+$.

Example 30 *Let* $S = X_1 + X_2$ *where* X_1 *and* X_2 *are independent rvs with* $X_i \sim U(0, b_i)$, $i = 1, 2$ *and* $0 < b_1 \leq b_2$. *Then, we obtain*

$$\begin{aligned} f_S(x) &= \int_0^x f_{X_1}(y) f_{X_2}(x-y)\,\mathrm{d}y \\ &= \int_0^x \frac{1}{b_1} I_{\{x\in[0,b_1]\}} \frac{1}{b_2} I_{\{(x-y)\in[0,b_2]\}} \mathrm{d}y \\ &= \begin{cases} \frac{x}{b_1 b_2}, & 0 \leq x \leq b_1 \\ \frac{b_1}{b_1 b_2}, & b_1 \leq x \leq b_2 \\ \frac{b_1+b_2-x}{b_1 b_2}, & b_2 \leq x \leq b_1 + b_2 \end{cases} . \end{aligned}$$

□

Example 31 *Let* $S = X_1 + X_2$ *where* X_1 *and* X_2 *are independent rvs with* $X_i \sim Exp(\beta_i)$, $i = 1, 2$ *and* $\beta_1 \neq \beta_2$. *Then we find*

$$f_S(x) = \int_0^x \beta_1 \mathrm{e}^{-\beta_1 y} \beta_2 \mathrm{e}^{-\beta_2(x-y)} \mathrm{d}y = \frac{\beta_2}{\beta_2 - \beta_1} \beta_1 \mathrm{e}^{-\beta_1 x} + \frac{\beta_1}{\beta_1 - \beta_2} \beta_2 \mathrm{e}^{-\beta_2 x},$$

which corresponds to the pdf of the Generalized Erlang distribution. Clearly, $E[S] = \int_0^\infty x f_S(x)\,\mathrm{d}x = \frac{1}{\beta_1} + \frac{1}{\beta_2}$ *which also corresponds to* $E[X_1] + E[X_2]$.
$\square$

10.2.4 Sum of independent rvs and mgf

Moment generating functions are useful to identify the distribution of a sum of rvs.

Let $S = \sum_{i=1}^{n} X_i$, where X_1, ..., X_n are independent rvs. In certain circumstances, it could be useful to use the mgf in order to find the distribution of S.

Indeed, we have

$$E\left[\mathrm{e}^{tS}\right] = E\left[\mathrm{e}^{t\sum_{i=1}^{n} X_i}\right] = \prod_{i=1}^{n} E\left[\mathrm{e}^{tX_i}\right]. \tag{20}$$

The idea is to identify the distribution of S the expression obtained in (20).

This approach is illustrated in the following examples.

The results obtained in those examples will be useful throughout the book.

Example 32 *Let* $X_i \sim Pois(\lambda_i)$ *for* $i = 1, ..., n$. *Then,* $S = \sum_{i=1}^{n} X_i \sim Pois(\lambda_1 + ... + \lambda_n)$. *Indeed, we apply (20)*

$$E\left[\mathrm{e}^{tS}\right] = \prod_{i=1}^{n} E\left[\mathrm{e}^{tX_i}\right] = \prod_{i=1}^{n} \mathrm{e}^{\lambda_i\left(\mathrm{e}^{t}-1\right)} = \mathrm{e}^{(\lambda_1+...+\lambda_n)\left(\mathrm{e}^{t}-1\right)}$$

which corresponds to the mgf of a Poisson distribution with parameter $\lambda_1 + ... + \lambda_n$. $\square$

We use the procedure provided in example 32 to find the desired results in the following examples.

Example 33 *Let* $X_i \sim Bin(r_i, q)$ *for* $i = 1, ..., n$. *Then,* $S = \sum_{i=1}^{n} X_i \sim Bin(r_1 + ... + r_n, q)$. $\square$

Example 34 *Let* $X_i \sim NB(r_i, q)$ *for* $i = 1, ..., n$. *Then,* $S = \sum_{i=1}^{n} X_i \sim NB(r_1 + ... + r_n, q)$. $\square$

Example 35 *Let* $X_i \sim Ga(\alpha_i, \beta)$ *for* $i = 1, ..., n$. *Then,* $S = \sum_{i=1}^{n} X_i \sim Ga(\alpha_1 + ... + \alpha_n, \beta)$. $\square$

Example 36 *Let* $X_i \sim Exp(\beta)$ *for* $i = 1, ..., n$. *Then,* $S = \sum_{i=1}^{n} X_i \sim Erl(n, \beta)$. $\square$

10.2.5 Sum of two dependent positive continous rvs

We consider a pair of positive continuous rv (X_1, X_2) for which the bivariate pdf is f_{X_1,X_2}.

Then, the pdf of $S = X_1 + X_2$ is defined in terms of f_{X_1,X_2} i.e.

$$f_S(s) = \int_0^s f_{X_1,X_2}(x_1, s - x_1)\, \mathrm{d}x_1. \tag{21}$$

10.2.6 Sum of two dependent positive discrete rvs

Let (X_1, X_2) be a pair of rvs with arithmetical support, i.e. where $X_i \in \{0, 1h, 2h, ...\}$ and h is strictly positive scalar. The bivariate pmf is denoted by $f_{X_1,X_2}(m_1 h, m_2 h)$.

Then, the pmf of $S = X_1 + X_2$ is given by

$$f_S(kh) = \sum_{m_1=0}^{k} f_{X_1,X_2}(m_1 h, kh - m_1 h). \tag{22}$$

11 Portfolio of risks and risk measures

11.1 Context

The basic risk measures are the *Value-at-Risk* (VaR) and the *Tail-Value-at-Risk* (TVaR).

We recall their definitions.

With a level of confidence, the VaR associated to the rv X is defined by

$$VaR_{\kappa}(X) = F_X^{-1}(\kappa) = \inf\{x \in \mathbb{R} : F_X(x) \geq u\}.$$

The definition of TVaR is given by

$$\begin{aligned} TVaR_\kappa(X) &= \frac{1}{1-\kappa}\int_\kappa^1 VaR_u(X)\,\mathrm{d}u \\ &= \frac{E\left[X \times 1_{\{X>VaR_\kappa(X)\}}\right]}{1-\kappa} \\ &\quad + \frac{VaR_\kappa(X)\left(F_X\left(VaR_\kappa(X)\right)-\kappa\right)}{1-\kappa}. \end{aligned}$$

11.2 Desirable properties and coherence

We present the desirable properties for a risk measure ς_κ.

Property 37 Homogeneity. *Let the rv X be a risk and and $a \in \mathbb{R}^{+}$ be a strictly positive scalar. A risk measure ς_{κ} is homogeneous if*

$$\varsigma_{\kappa}(aX) = a\varsigma_{\kappa}(X),$$

for $0 < \kappa < 1$.

A modification of the currency of the risk X should lead to a similar change to the value obtained with the risk measure.

Property 38 Invariance to translation. *Let the risk X and a scalar $a \in \mathbb{R}$. A risk measure ς_{κ} is invariant to translation if*

$$\varsigma_{\kappa}(X + a) = \varsigma_{\kappa}(X) + a,$$

for $0 < \kappa < 1$.

For example, for a loss rv $L = X - \pi$ where π corresponds to the premium associée à X, it is justified that $\varsigma_\kappa(L) = \varsigma_\kappa(X) - \pi$.

Property 39 Monotonocity. *Let the rv X_1 and X_2 be two risks such that*

$$\Pr(X_1 \leq X_2) = 1.$$

A risk measure ς_κ is monotone if

$$\varsigma_\kappa(X_1) \leq \varsigma_\kappa(X_2),$$

for $0 < \kappa < 1$.

If the risk X_2 is more dangerous than the risk X_1, it is reasonable that the risk measure produces a greater capital for X_2 than the capital for X_1.

Property 40 Subadditivity. *Let X_1 and X_2 be two risks. The risk measure ς_κ is subadditive if*

$$\varsigma_\kappa(X_1 + X_2) \leq \varsigma_\kappa(X_1) + \varsigma_\kappa(X_2),$$

for $0 < \kappa < 1$.

The property of subadditivity is very important regarding risk pooling.

Indeed, it is interesting to examine the benefit of risk pooling that corresponds to

$$B_\kappa^\varsigma(S) = \sum_{i=1}^{n} \varsigma_\kappa(X_i) - \varsigma_\kappa(S)$$

and that results by pooling risks X_1, ..., X_n.

As risk pooling is the foundation of the insurance, it would be desirable that $B_{\kappa}^{\varsigma}(S)$ is positive.

The notion of coherent risk measure is due to [[**?**]].

Definition 41 Coherent risk measure. *A risk measure ς_{κ} si said to be* coherent *if Properties 37, 38, 39 and 40 are satisfied.*

11.3 Other desirable properties

In actuarial science (see e.g. [**?**]), the three following properties are considered to be desirable.

Property 42 No excessive risk margin (no rip-off). *The risk measure* ς_κ *should not induce an excessive risk margin. If* $X \leq x_{\mathsf{max}}$*, then we have*

$$\varsigma_\kappa (X) \leq x_{\mathsf{max}},$$

for $0 < \kappa < 1$*.*

It is unjustified to hold a capital in excess of the maximum amount that the aggregate claim of a policy or the entire portfolio can take.

Property 43 Positive Risk Margin. *We should have* $\varsigma_\kappa(X) \geq E[X]$, *for* $0 < \kappa < 1$.

The minimal capital must exceed the expected aggregate claim amount (for a policy or a portfolio), otherwise the ruin will certain.

Property 44 Justified Risk Margin. *Let* $a \in \mathbb{R}$ *be a scalar. We should have* $\varsigma_\kappa(a) = a$, *for* $0 < \kappa < 1$.

It is not justified to allocate a capital different from a if the aggregate claim amount for a portfolio corresponds to a scalar a.

11.4 Value-at-Risk

11.4.1 Invariant, homogeneous and monotone

Since $\varphi(x) = ax$ and $\varphi(x) = x + a$ are increasing and continuous functions, we deduct with the help of Proposition 16 that the VaR is invariant to translation and positive homogeneous.

The risk measure VaR is also monotone. If $\Pr(X_1 \leq X_2) = 1$, them we observe $F_{X_1}(x) \geq F_{X_2}(x)$ for all x, which implies

$$VaR_\kappa(X_1) \leq VaR_\kappa(X_2)$$

for all κ.

11.4.2 Not subadditive

The risk measure VaR is not subadditive as it is illustrated with the following counterexample.

Example 45 *Let X_1 and X_2 be two independent rvs X_1 and X_2 where $X_i \sim Exp\left(0.1\right)$. We now that $X_1 + X_2 \sim Gamma\left(2, 0.1\right)$ (which is equivalent to Erlang (2,0.1)). With the values of $VaR_\kappa\left(X_1 + X_2\right)$ and $VaR_k\left(X_1\right) + VaR_k\left(X_2\right)$ provided in the table below, it is that the VaR is not subadditive :*

κ	0.1	0.2	0.5	0.8	0.9
$VaR_k\left(X_1\right) + VaR_k\left(X_2\right)$	0.2107	0.4463	1.3863	3.2189	4.6052
$VaR_k\left(X_1 + X_2\right)$	0.5318	0.8244	1.6783	2.9943	3.8897

In numerous examples, we obtain the following inequality (for values of κ):

$$VaR_\kappa\left(S\right) > VaR_\kappa\left(X_1\right) + ... + VaR_\kappa\left(X_n\right).$$

It means that the VaR is not subadditive and, that for given values of κ, it leads to negative benefits of risk pooling.

Since $VaR_{\kappa}(X) \leq x_{\mathsf{max}}$ for all κ and $VaR_{\kappa}(a) = a$, the VaR does not introduce excessive risk margin and the risk margin which is introduced with the VaR is justified.

However, since $VaR_{\kappa}(X) \leq E[X]$ for values of κ, then the VaR may introduce negative risk margin.

11.5 Tail-Value-at-Risk

11.5.1 Positive homogeneous

The TVaR measure is positive homogeneous.

Indeed, for $a \geq 0$, we have

$$\begin{aligned} TVaR_{\kappa}(aX) &= \frac{1}{1-\kappa}\int_{\kappa}^{1} VaR_{u}(aX)\,du \\ &= a\frac{1}{1-\kappa}\int_{\kappa}^{1} VaR_{u}(X)\,du = aTVaR_{\kappa}(X), \end{aligned}$$

which implies that the TVaR is homogeneous.

11.5.2 Invariant to translation

The TVaR is invariant to the translation.

Indeed, we obtain the desire relation by proceeding in the following way:

$$\begin{aligned} TVaR_{\kappa}(X+a) &= \frac{1}{1-\kappa}\int_{\kappa}^{1} VaR_{u}(X+a)\,du \\ &= \frac{1}{1-\kappa}\int_{\kappa}^{1} VaR_{u}(X)\,du + \frac{1}{1-\kappa}\int_{\kappa}^{1} a du \\ &= TVaR_{\kappa}(X) + a. \end{aligned}$$

11.5.3 Monotone

The TVaR is monotone.

Since the inequality $VaR_\kappa(X_1) \leq VaR_\kappa(X_2)$ is verified for all κ, then it is clear that $TVaR_\kappa(X_1) \leq TVaR_\kappa(X_2)$ for all κ.

11.5.4 Subbadditive

In the aim to verify that the TVaR is subadditive, we follow the approach explained in [**?**].

We need the following lemma.

Lemma 46 *Let X be a rv with cdf F_X.*

Let X_1, ..., X_n be a sequence of iid rvs where $X_i \sim X$, for $i = 1, 2, ..., n$.

Then, we have

$$TVaR_{\kappa}(X) = \lim_{n\rightarrow\infty} \frac{\sum_{j=[n\kappa]+1}^{n} X_{j:n}}{[n(1-\kappa)]} \text{ (a.s.)}, \tag{23}$$

where

$$X_{1:n} \leq X_{2:n} \leq ... \leq X_{n-1:n} \leq X_{n:n}$$

are the order statistics of

$$X_1, ..., X_n$$

and $[u]$ *corresponds at the integer part of* u*.*

Proof. See the proof of Proposition 4.1 in [**?**] (including pages 1494 and 1495 as well). ■

The result of Lemma 46 provide an interesting interpretation of the TVaR.

An adapted version of this lemma is used in Section 13.5 when Monte-Carlo simulation is applied to evaluate the TVaR.

Proposition 47 *The TVaR is subadditive.*

Proof. We follow the proof provided in [**?**].

It is also recommende to consult the proof presented in Appendix A of [**?**].

Another proof is provided in [**?**].

Let X_1, ..., X_n be a sequence of iid rvs where $X_i \sim X$, $i = 1, 2, ..., n$.

For an integer m such that $1 \leq m + 1 \leq n$, the following equality is satisfied:

$$\sum_{j=m+1}^{n} X_{j:n} = \sup \left\{ X_{i_{m+1}} + ... + X_{i_n}; 1 \leq i_{m+1} < ... < i_n \leq n \right\}.$$

Let (X, Y) be a pair of rvs with bivariate cdf denoted by $F_{X,Y}$.

Let (X_1, Y_1), (X_2, Y_2), ..., (X_n, Y_n) be sequence of iid pairs of rvs where $(X_i, Y_i) \sim (X, Y)$ for $i = 1, 2, ...n$.

We define $S = X + Y$ and $S_i = X_i + Y_i$, for $i = 1, 2, ..., n$.

By Lemma 46, we have

$$TVaR_\kappa (S) = \lim_{n \to \infty} \frac{\sum_{j=[n\kappa]+1}^{n} S_{j:n}}{[n(1-\kappa)]} \text{ (a.s.)},$$

where

$$S_{1:n} \leq S_{2:n} \leq ... \leq S_{n-1:n} \leq S_{n:n}$$

are order statistics of S_1, ..., S_n and $[u]$ corresponds to the integer part of u.

We have

$$\begin{aligned}
\sum_{j=[n\kappa]+1}^{n} S_{j:n} &= \sup\left\{S_{i_{[n\kappa]+1}} + ... + S_{i_n}; 1 \leq i_{[n\kappa]+1} \leq ... \leq i_n \leq n\right\} \\
&\leq \sup\left\{X_{i_{[n\kappa]+1}} + ... + X_{i_n}; 1 \leq i_{[n\kappa]+1} \leq ... \leq i_n \leq n\right\} \\
&\quad + \sup\left\{Y_{i_{[n\kappa]+1}} + ... + Y_{i_n}; 1 \leq i_{[n\kappa]+1} \leq ... \leq i_n \leq n\right\} \\
&= \sum_{j=[n\kappa]+1}^{n} X_{j:n} + \sum_{j=[n\kappa]+1}^{n} Y_{j:n}.
\end{aligned}$$

We divide by $[n(1-\kappa)]$, let $n \to \infty$ and we obtain the desired result, applying Lemma 46.

■

11.5.5 Positive risk margin

We know that $TVaR_0(X) = E[X]$.

To verify that $TVaR_\kappa(X) \geq E[X]$, we take the derivative of the expression for the TVaR is with respect to κ and we obtain

$$\begin{aligned}
\frac{dTVaR_\kappa(X)}{d\kappa} &= \frac{d\left(\frac{1}{1-\kappa}\int_\kappa^1 VaR_u(X)\,du\right)}{d\kappa} \\
&= \frac{-VaR_\kappa(X)}{1-\kappa} + \frac{TVaR_\kappa(X)}{1-\kappa}.
\end{aligned}$$

Since

$$\begin{aligned} TVaR_\kappa(X) &= \frac{1}{1-\kappa}\int_\kappa^1 VaR_u(X)\,du \\ &\geq \frac{1}{1-\kappa}\int_\kappa^1 VaR_\kappa(X)\,du = VaR_\kappa(X), \end{aligned}$$

then we have $\frac{dTVaR_\kappa(X)}{d\kappa} \geq 0$, which implies that the TVaR is increasing in κ.

We may conclude that the TVaR introduce a positive risk margin.

11.5.6 No excessive risk margin

The TVaR does not produce an excessive risk margin because we have

$$\begin{aligned} TVaR_{\kappa}(X) &= \frac{1}{1-\kappa}\int_{\kappa}^{1} VaR_u(X)\,du \\ &\leq \frac{1}{1-\kappa}\int_{\kappa}^{1} x_{\mathsf{max}}du = x_{\mathsf{max}}. \end{aligned}$$

11.5.7 Justified Risk Margin

Since we observe

$$TVaR_{\kappa}(a) = \frac{1}{1-\kappa}\int_{\kappa}^{1} a du = a,$$

we conclude that the TVaR introduce a risk margin that it is appropriate.

12 Risk Pooling – First steps

It is well known that insurance activities are based on risk pooling.

12.1 Benefit of risk pooling and risk measures

12.1.1 Definition

Let ς_{κ} be a risk measure.

The benefit of pooling n risks (rvs X_1, ..., X_n) is measured by

$$B_{\kappa}^{\varsigma}\left(X_1, ..., X_n\right) = \sum_{i=1}^{n} \varsigma_{\kappa}\left(X_i\right) - \varsigma_{\kappa}\left(S\right).$$

Since the TVaR is subadditive, the benefit of risk pooling is always positive i.e.

$$B_{\kappa}^{TVaR}\left(X_1, ..., X_n\right) \geq 0.$$

12.1.2 Numerical illustration

In the following numerical example, we illustrate risk pooling using the VaR and the TVaR.

It also provide another example that the VaR is not subadditive.

We consider a homogeneous portfolio of m 1-period term life insurance contracts

The aggregate claim amount for the insurance contract i is defined by the rv

$$X_i = bI_i$$

where $b = 100000$ is the death benefit, for $i = 1, 2, ..., m$.

Let I_1, ..., I_m be iid rvs with

$$I_i \sim I \sim Bern(q)$$

for $i = 1, 2, ..., m$ and for $q = 0.0017$.

The rv I_i corresponds to the occurrence rv where

$$I_i = \begin{cases} 1 & \text{, if the insured dies over the next period} \\ 0 & \text{, otherwise} \end{cases} ,$$

for $i = 1, 2, ..., m$.

Then, the aggregate claim amount for the entire portfolio can be represented as follows

$$
\begin{aligned}
S &= \sum_{i=1}^{m} X_i \\
&= b\sum_{i=1}^{m} I_i \sim bN,
\end{aligned}
$$

where the rv N is binomially distributed i.e.

$$
N = \sum_{i=1}^{m} I_i \sim Binom\,(m; q)\,.
$$

For a single policy, we obtain the following results for $i = 1, 2, ..., n$:

- the pure premium for a policy is $E\,[X_i] = 100\ 000q = 170$;
- $VaR_{0.995}\,(X_i) = 0$;

- $TVaR_{0.995}(X_i) = 34\ 000$.

In the following table, we provide the values of $VaR_{0.995}\left(\sum_{i=1}^{n} X_i\right)$ and $TVaR_{0.995}\left(\sum_{i=1}^{n} X_i\right)$ for $n = 100$, $1000, 100\ 000$ and $1\ 000\ 000$:

n	100	1000	100 000	1 000 000
$VaR_{0.995}\left(\sum_{i=1}^{n} X_i\right)$	200 000	600 000	20 400 000	180 700 000
$TVaR_{0.995}\left(\sum_{i=1}^{n} X_i\right)$	214 640.5	646 349.1	20 889 484	182 036 101

The values in the table are exact.

For the VaR, we observe a negative benefit of risk pooling for portfolio size n:

$$B_{0.995}^{VaR}(S) = \sum_{i=1}^{n} VaR_{0.995}(X_i) - VaR_{0.995}\left(\sum_{i=1}^{n} X_i\right) \leq 0.$$

As expected, since the TVaR is subadditive, the benefit of risk pooling is positive for any portfolio size n:

$$B_{0.995}^{VaR}(S) = \sum_{i=1}^{n} TVaR_{0.995}(X_i) - TVaR_{0.995}\left(\sum_{i=1}^{n} X_i\right) \geq 0.$$

The use of individual VaRs to determine the total capital for the entire portfolio can lead to disastrous financial consequences.

Indeed, from the perspective of a risk manager for a financial institution, it is assumed that he asks each entity i ($i = 1,2, ..., n$) (whose risk is represented by the rv X_i) to establish the amount of capital to be set aside to deal with each risk and make the total.

If we use the VaR, it is not necessary to fund the risk of death because the probability of survival is greater than 99.5%.

As the amount of capital is 0 for each entity, it follows that their sum is 0.

Based on the TVaR, the amount to be set aside for each risk would be 34 000 compared to an expected aggregate claim amount per policy of 170.

The capital cost is exorbitant.

Taking advantage of the pooling of risks and n= 1000 (1 000 000), the aggregate capital based on TVaR is then 646 349.1 (182 036 101) or 646.3491 (182.0361) per policy.

Per policy the benefit of risk poling is 34 000 - 646.3491 = 33 353.6509 (34 000 - 182.0361 = 33 817. 9639).

From the viewpoint of a pricing actuary, the possible use of VaR to calculate the premium is also problematic.

According to the VaR, the premium for a single policy would be 0, which is lower than the pure premium.

Using the TVaR, a premium of 34 000 is disproportionate to the pure premium, which is 170.

However, assuming that n contracts are issued and that the rvs X_1, ..., X_n are iid, the benefit from risk pooling allows the insurance compagy to ask for a lower premium.

For e.g. $n = 100\ 000$ and using the $TVaR$, the premium for a single policy is 208.8948.

12.2 Risk pooling and average cost per contract

12.2.1 Definition

Through the mecanism of insurance, each insured transfers its actuarial risk to an insurance company.

The aggregation or pooling of all those actuarial risks corresponds to global risk for the insured.

To understand risk pooling, it is interesting to examine the stochastic behavior of the average cost per contract for an homogeneous portfolio of risks.

In homogeneous portfolios of m risks, the rvs X_1, ..., X_m share the same stochastic behavior. i.e. $X_i \sim X$.

We assume the expectation and the variance of the rv X are finite.

The average cost per contract is defined by the rv W_n, where

$$W_n = \frac{S_n}{n} = \frac{\text{Aggregate claim amount}}{\text{Number of contracts}}.$$

The average cost per contract corresponds to the share of the aggregated claim amount for the entire portfolio which is allocated to one policy (contract) of the portfolio.

We consider two cases:

- portfolio of independent risks
- portfolio of dependent risks

12.2.2 Portfolio of independent risks

The expectation of the rv W_n is

$$E[W_n] = E\left[\frac{S_n}{n}\right] = \frac{1}{n}E[S_n].$$

Since the rvs X_1, ..., X_n are identically distributed, $E[W_n]$ becomes

$$E[W_n] = \frac{nE[X]}{n} = E[X].$$

It implies that the expectation of the average cost per contract is equal to the expected cost for a single contract.

The variance of the rv W_n is given by

$$\mathsf{Var}\left(W_n\right) = \mathsf{Var}\left(\frac{S_n}{n}\right) = \frac{1}{n^2}\mathsf{Var}\left(S_n\right).$$

Since the rvs X_1, ..., X_n are iid, $Var\left(W_n\right)$ becomes

$$\mathsf{Var}\left(W_n\right) = \frac{1}{n^2} n \ \mathsf{Var}\left(X\right) = \frac{1}{n}\mathsf{Var}\left(X\right) < \mathsf{Var}\left(X\right),$$

We observe that the variance of the average cost per contract for a portfolio of $n > 1$ contract is lower to the variance of the cost for a single contract.

When n becomes very large, we observe that

$$\lim_{n\rightarrow\infty} \mathsf{Var}\left(W_n\right) = \lim_{n\rightarrow\infty} \frac{1}{n}\mathsf{Var}\left(X\right) = 0.$$

It implies that variance of the average cost per contract tends to 0 as the number of risks (contracts) considerably increases.

12.2.3 Portfolio of independent risks – Example

Let X_1 ,..., X_n be iid risks where X_i is defined by

$$X_i = 100000 I_i$$

where $I_i \sim Bern(0.001)$, for $i = 1, 2, ..., n$.

The expectation and the variance of the rv X_i are

$$E[X_i] = 100000q = 100$$

and

$$\begin{aligned} Var(X_i) &= 100000^2 q(1-q) \\ &= 9900000. \end{aligned}$$

Recall that

$$S_n = bN_n$$

where

$$N_n \sim Bin\,(n, q)\,.$$

We also know that

$$VaR_\kappa\,(W_n) = \frac{1}{n} VaR_\kappa\,(S_n) = \frac{b}{n} VaR_\kappa\,(N_n)$$

and

$$TVaR_\kappa\,(W_n) = \frac{1}{n} TVaR_\kappa\,(S_n) = \frac{b}{n} TVaR_\kappa\,(N_n)$$

The magnitude of the eventual losses (100000) are very large compared to its expected value (100).

An individual cannot deposit into a bank account an amount of 100000, in the case of the occurrence of the losses over the next period, which such a low probability (0.001).

Also, if the individual decides to deposit 100 (expected value) or 200 (its double), those amount are so negligible that they are useless.

Since the individual cannot face the eventual losses by himself/herself, the risk pooling is non-expensive and efficient way to be protected against the eventual losses.

For a portfolio of n independent insureds, the expectation of his/her share (average cost per contract) of eventual losses remains the same

$$E[W_n] = E[X] = 100.$$

However, the variance of his/her share (average cost per contract) of eventual losses decreases by a rate of $\frac{1}{n}$ i.e.

$$Var\left(W_n\right) = \frac{1}{n} Var\left(X\right).$$

Values of $Var\left(W_n\right)$, $VaR_{0.999}\left(W_n\right)$ and $TVaR_{0.999}\left(W_n\right)$ are provided in the following table :

n	$E\left[W_n\right]$	$Var\left(W_n\right)$	$VaR_{0.999}\left(W_n\right)$	$TVaR_{0.999}\left(W_n\right)$
1	100	9900000	0	100000
1000	100	9900	500	568.13
1000000	100	9.9	109.9	110.81

12.2.4 Portfolio of dependent risks

We consider an homogeneous portfolio of dependent risks.

The rvs X_1,..., X_n are assumed to be identically distributed.

The dependence relation between the rvs has no impact on the expectation of the rv W_n which is

$$E\left[W_n\right] = E\left[\frac{S_n}{n}\right] = \frac{1}{n}E\left[S_n\right].$$

Since the rvs X_1, ..., X_n are identically distributed, $E\left[W_n\right]$ becomes

$$E\left[W_n\right] = \frac{nE\left[X\right]}{n} = E\left[X\right].$$

However, the variance of W_n is given by

$$\begin{aligned}\mathrm{Var}\,(W_n) &= \frac{1}{n^2}\mathrm{Var}\,(S_n) = \frac{1}{n^2}\left[n\mathrm{Var}\,(X) + n(n-1)\mathrm{Cov}(X_1, X_2)\right] \\ &= \frac{\mathrm{Var}\,(X)}{n} + \frac{n-1}{n}\mathrm{Cov}(X_1, X_2).\end{aligned}$$

Also, when n becomes very large, we observe that

$$\lim_{n\to\infty} \mathrm{Var}\,(W_n) = \lim_{n\to\infty} \left\{\frac{\mathrm{Var}\,(X)}{n} + \frac{n-1}{n}\mathrm{Cov}(X_1, X_2)\right\} = \mathrm{Cov}(X_1, X_2).$$

It means that the variance of the average cost per contract is not eliminated when the number of insureds increases.

This stochastic behavior of the rv W_n via the VaR and the TVaR within the context of a portfolio of dependent risks will be illustrated later.

12.3 Risk pooling, central limit theorem and moment-based approximation

12.3.1 Context

We consider a portfolio of independent rvs X_1, ..., X_n, with finite expectations and variances.

The aggregate claim amount for the portfolio is defined by the rv

$$S = X_1 + ... + X_n.$$

It implies that we know the expectation and the variance of the S i.e.

$$E[S] = \sum_{i=1}^{n} E[X_i] \quad \text{and} \quad Var(S) = \sum_{i=1}^{n} Var(X_i).$$

However, assume that we cannot find a closed-form expression for F_S.

The idea behind moment-based approximation is to approximate F_S by F_T, where T is a rv for which we can easy compute the values of F_T and other related quantities of interest such as the risk measures VaR and the TVaR.

The distribution of the rv T must have two parameters.

These two parameters are fixed such that $E[T] = E[S]$ and $\mathrm{Var}(T) = \mathrm{Var}(S)$.

One of the main moment-based approximation is the approximation method based on the normal distribution, often called "normal approximation"

12.3.2 Approximation method based on the normal distribution

The approximation method based on the normal distribution is based on (justified by) the Central Limit Theorem.

The values of $E[S]$ and $VaR(S)$ are assumed to be known.

We approximate the rv S by the rv $T \sim Norm(E[T], \mathsf{Var}(T))$, where $E[T] = E[S]$ and $\mathsf{Var}(T) = \mathsf{Var}(S)$.

It follows that

$$
\begin{aligned}
F_S(x) &\simeq F_T(x) = \Phi\left(\frac{x - E[S]}{\sqrt{\mathsf{Var}(S)}}\right), \\
VaR_\kappa(S) &\simeq VaR_\kappa(T) = E[S] + \sqrt{\mathsf{Var}(S)}VaR_\kappa(Z), \\
TVaR_\kappa(S) &\simeq TVaR_\kappa(T) = E[S] + \sqrt{\mathsf{Var}(S)}TVaR_\kappa(Z), \\
&= E[S] + \sqrt{\mathsf{Var}(S)}\frac{e^{-\frac{(VaR_\kappa(Z))^2}{2}}}{(1-\kappa)\sqrt{2\pi}},
\end{aligned}
$$

where $Z \sim N(0,1)$.

The approximation is justified by the Central Limit Theorem.

We present two versions of the Central Limit Theorem :

- Sum of iid rvs with finite expectation and finite variance (well-known) ;
- Sum of independent rv with finite expectation and finite variance, and satisfying Lindeberg's condition.

Here is the classical and well-known version of the Central Limit Theorem.

Theorem 48 Central Limit Theorem (Sum of iid rvs with finite expectation and finite variance).

Let X_1, X_2, ..., X_n *be iid rvs with* $E[X_i] = \mu_X$ *and* $\mathrm{Var}(X_i) = \sigma_X^2$ *for* $i = 1, 2, ..., n$.

We define the rv $S_n = \sum_{i=1}^{n} X_i$, *with* $E[S_n] = n\mu_X$ *and* $\mathrm{Var}(S_n) = n\sigma^2$.

Then, we have

$$\frac{S_n - E[S_n]}{\sqrt{\mathrm{Var}(S_n)}} \xrightarrow{d} Z,$$

when $n \to \infty$ *i.e.*

$$\lim_{n\to\infty} \Pr\left(\frac{S_n - E[S_n]}{\sqrt{\mathrm{Var}(S_n)}} \leq x\right) = F_Z(x) = \Phi(x)$$

for every real number x.

The following theorem provides an extension to the classical and well-known version of the Central Limit Theorem.

Theorem 49 Central Limit Theorem (Sum of independent rv with finite expectation and finite variance, and satisfying Lindeberg's condition.).

Let X_1, X_2, ..., X_n *be independent rvs with* $E[X_i] = \mu_{X_i}$ *and* $\mathrm{Var}(X_i) = \sigma^2_{X_i}$ *for* $i = 1, 2, ..., n$.

We define the rv $S_n = \sum_{i=1}^n X_i$, *with*

$$E[S_n] = \sum_{i=1}^n E[X_i] = \sum_{i=1}^n \mu_{X_i} = \mu_{S_n}$$

and

$$\mathrm{Var}(S_n) = \sum_{i=1}^n \mathrm{Var}(X_i) = \sum_{i=1}^n \sigma^2_{X_i} = \sigma^2_{S_n}.$$

(Lindeberg's condition) Suppose that for every $\varepsilon > 0$ *the condition*

$$\lim_{n\to\infty} \frac{1}{\sigma^2_{S_n}} \sum_{i=1}^n E\left[\left(X_i - \mu_{X_i}\right)^2 1_{\left\{\left|X_i - \mu_{X_i}\right| > \varepsilon\sigma^2_{S_n}\right\}}\right] = 0$$

is satisfied.

Then, we have

$$\frac{S_n - E[S_n]}{\sqrt{\mathsf{Var}(S_n)}} \stackrel{d}{\to} Z$$

when $n \to \infty$, i.e.

$$\lim_{n\to\infty} \Pr\left(\frac{S_n - E[S_n]}{\sqrt{\mathsf{Var}(S_n)}} \leq x\right) = F_Z(x) = \Phi(x)$$

for every real number x.

Proof. Voir e.g. [**?**]. ■

Remark 50 *Lindeberg's condition means that variances of the rvs X_i must be small relative to the variance of the sum S_n. See e.g. [**?**] for details.*

12.3.3 Be careful

The quality of the approximation by the standard normal distribution depends on the number of risks and the choice of distribution for the rvs $X_1, ..., X_n$.

It is a simple approximation to use, but, frequently in non-life actuarial science, it is not satisfactory, especially in the tail of the distribution of the rv S.

The normal approximation should be applied carefully, because it may lead to inappropriate assessment of the risk measures.

To remedy this problem, we use recursive methods and methods based on Monte Carlo simulation.

Other moment-based approximations could be also considered (see the Appendix) but there are ot of scope of the Math541.

Important : both central limit theorems assumed that the rvs X_1, ..., X_n. These two theorems cannot be used to justify to use the normal approximation in the context of a portfolio of dependent risks.

13 Monte Carlo Stochastic Simulation Methods - First Steps

13.1 Introduction

Monte Carlo stochastic simulation methods are a crucial tool actuarial and quantitative risk management.

For many risk models actuarial science, it is impossible to know the exact shape of the distribution of the aggregate claim amount for the portfolio of an insurance company.

The aim of this section is to present a brief introduction to Monte Carlo stochastic simulation methods.

13.2 Pseudo-random number generators

The practical implementation of stochastic simulation methods are based on the ability to produce systematically samples from the standard uniform distribution.

These samples of the standard uniform distribution are called "pseudo-random numbers" and they are produced using pseudo-random number generators (PRNG).

The goal of a good PRNG is to mimic as closely as possible the sample of a sequence of iid rvs standard uniform distribution.

The PRNG has to be fast and easy to implant.

There must be a good compromise between speed, good statistical properties and unpredictability.

The classical PRNG is the linear congruential generator.

Algorithm 51 Linear congruential generator (LCG). *The LCG is defined by the recursive relation*

$$x_n = (ax_{n-1}) \bmod m, \; n \in \mathbb{N}^+,$$

where a and m are positive integers carefully chosen and x_0 is the initial value, called the "seed".

The n-th sample of the rv $U \sim U(0,1)$ is obtained with $U^{(n)} = \frac{x_n}{m}$ for $n \in \mathbb{N}^+$.

Remark 52 *Recommended values of* (a, m)*:*

- $a = 41\ 358$ *and* $m = 2^{31} - 1 = 2\ 147\ 483\ 647$ *(see [**?**])*
- $a =$*950706376 and* $m = 2^{31} - 1 = 2\ 147\ 483\ 647$ *(see Lemieux (2009)).*

The LCG is illustrated in the following example.

Example 53 *Applying the LCG defined in 51 with* $x_0 = 343\ 463\ 463$*, we obtained the following values of* $x_1, ..., x_4$ *and* $U^{(1)}, ..., U^{(4)}$ *:*

j	1	2	3	4
x_j	1 505 061 496	1 519 843 273	831 737 044	587 608 106
$U^{(j)}$	0.700848874	0.707732175	0.387307743	0.273626347

□

Within this monograph, the algorithm 51 will be used in some examples and exercises.

A detailed treatment on uniform random number generation and PRNG is beyond the scope of this monograph.

See e.g. Lemieux (2009) for a detailed treatment.

However, it is strongly recommended to avoid to use LCG for practical implementations (whatever the contexts).

Each software has is own PRNG.

For example, the Mersenne-Twister generator is the default PRNG in R. Other PRNGs are also proposed in R.

R function:

- Generator of samples of standard uniform distribution: `runif()` ;
- Seed : `setseed()`.

13.3 Basic sampling methods

13.3.1 Context

Let X be a rv with cdf F_X.

We examine different methods to simulate samples $X^{(j)}$ $(j = 1, 2, ..., m)$ of the rv X.

13.3.2 Inverse Function Method

Assume that the rv X has a cdf F_X.

Furthermore, let us assume that we have a closed-form expression for F_X^{-1}.

The inverse function method is based on the Quantile Function Theorem, that we recall below.

Theorem 54 Quantile Function Theorem. *Let X be a rv with cdf F_X and quantile function F_X^{-1}. Let U be a rv such that $U \sim U(0,1)$. Then, the cdf of $F_X^{-1}(U)$ is F_X, i.e. $F_X^{-1}(U) \overset{d}{=} X$.*

In the following algorithm, we decribe the inverse function method to simulate samples $X^{(j)}$ $(j = 1, 2, ..., m)$ of the rv X with cdf $F_X(x)$.

Algorithm 55 Inverse Function Method.

1. **Step 1.** *Simulate a sample* $U^{(j)}$ *of* $U \sim U(0, 1)$ *from a PRNG.*
2. **Step 2**. *Simulate a sample* $X^{(j)}$ *of* X *with*

$$X^{(j)} = F_X^{-1}\left(U^{(j)}\right).$$

Repeat steps 1 and 2 for $j = 1, 2, ..., m$. $\square$

The Inverse Function Method is easy to implement when the quantile function is explicit.

Example 56 *Let* $X \sim Exp\left(\beta = \frac{1}{1000}\right)$ *with* $F_X^{-1}(u) = -\frac{1}{\beta}\ln(1-u)$. *With 5 samples of rv* $U \sim U(0,1)$, *we obtained the following réalisations of rv* X *:*

j	*1*	*2*	*3*	*4*	*5*
$U^{(j)}$	*0.09727801*	*0.30285445*	*0.01541586*	*0.98764877*	*0.53034863*
$X^{(j)}$	*102.34065*	*360.76106*	*15.53592*	*4393.99927*	*755.76463*

$\square$

In the following table, we provide a list of continuous distributions having an explicit quantile function with $u \in]0,1[$.

Distribution	Symbol	Cdf F_X	Quantile function F_X^{-1}
Exponential	$X \sim Exp\,(\beta)$	$1 - e^{-\beta x}$	$-\frac{1}{\beta} \ln(1-u)$
Pareto	$X \sim Pa\,(\alpha, \lambda)$	$1 - \left(\frac{\lambda}{\lambda+x}\right)^{\alpha}$	$\lambda \left((1-u)^{-\frac{1}{\alpha}} - 1\right)$
Weibull	$X \sim We\,(\tau, \beta)$	$1 - e^{-(\beta x)^{\tau}}$	$\frac{1}{\beta}(-\ln(1-u))^{\frac{1}{\tau}}$
Burr	$X \sim Burr\,(\alpha, \lambda, \tau)$	$1 - \left(\frac{\lambda}{\lambda+x^{\tau}}\right)^{\alpha}$	$(\lambda \left\{(1-u)^{-1/\alpha} - 1\right\})^{1/\tau}$
Log-Logistic	$X \sim LL\,(\lambda, \tau)$	$\frac{1}{1+\left(\frac{x}{\lambda}\right)^{\tau}}$	$\lambda \left(u^{-1} - 1\right)^{-1/\tau}$
Beta$(\alpha, 1)$	$X \sim Beta(\alpha, 1)$	x^{α} $(x \in [0,1])$	$u^{\frac{1}{\alpha}}$
Beta$(1, \beta)$	$X \sim Beta(1, \beta)$	$1 - (1-x)^{\beta}$ $(x \in [0,1])$	$1 - (1-u)^{\frac{1}{\beta}}$

The software R provide quantile functions for several discrete and continuous functions.

13.3.3 Simulation of samples of a rv defined as a function of a finite number of rvs

Let Z be a rv defined as a function of a vector of (independent or dependent) rvs $(X_1, ..., X_n)$ i.e.

$$Z = \phi(X_1, ..., X_n).$$

Assume that there is no closed-form for F_Z and F_Z^{-1}.

Algorithm 57 *Simulate samples of a rv defined as a function of a finite number of rvs*

1. **Step 1**. *Simulate a sample* $\left(X_1^{(j)}, ..., X_n^{(j)}\right)$ *of the vector of rvs* $(X_1, ..., X_n)$.

2. **Step 2**. *Simulate a sample $Z^{(j)}$ of the rv Z with*

$$Z^{(j)} = \phi\left(X_1^{(j)}, ..., X_n^{(j)}\right).$$

Repeat steps 1 and 2 for $j = 1, 2, ..., m$.

An interesting example of function ϕ for applications in actuarial science is

$$\phi(x_1, ..., x_n) = \sum_{i=1}^{n} x_i.$$

13.3.4 Simulation of samples of S, a finite sum of rvs

Recall our context: we consider a portfolio of risks represented by the rvs $X_1, ..., X_n$.

The rvs $X_1, ..., X_n$ may be independent or dependent.

The aggregate claim amount for the entire portfolio is represented by the rv S where

$$S = \sum_{i=1}^{n} X_i.$$

Assume that there is no closed-form for F_S and F_S^{-1}.

Algorithm 58 *Simulate samples of a rv defined as a function of a finite number of rvs*

1. **Step 1**. *Simulate a sample* $\left(X_1^{(j)}, ..., X_n^{(j)}\right)$ *of the vector of rvs* $(X_1, ..., X_n)$.
2. **Step 2**. *Simulate a sample* $S^{(j)}$ *of the rv* S *with*

$$S^{(j)} = X_1^{(j)} + ... + X_n^{(j)}.$$

Repeat steps 1 and 2 for $j = 1, 2, ..., m$.

In this chapter, we assume that the rvs $X_1, ..., X_n$ are independent with cdf F_{X_i} and quantile function $F_{X_i}^{-1}$, $i = 1, 2, ..., n$.

Algorithm 59 *Simulate samples of a vector of independent rvs*

1. **Step 1**. *Simulate a sample* $\left(U_1^{(j)}, ..., U_n^{(j)}\right)$ *of the vector of iid rvs* $(U_1, ..., U_n)$, *where* $U_i \sim Unif\,(0,1)\,,\ i = 1,2,...,n$.
2. **Step 2**. *Simulate a sample* $\left(X_1^{(j)}, ..., X_n^{(j)}\right)$ *of the vector of rvs* $(X_1, ..., X_n)$ *using the inverse method in the following way :*

$$X_1^{(j)} = F_{X_1}^{-1}\left(U_1^{(j)}\right) \quad ... \quad X_n^{(j)} = F_{X_n}^{-1}\left(U_n^{(j)}\right).$$

Repeat steps 1 and 2 for $j = 1,2,...,m$.

In forthcoming chapters, we will present simulation methods for vectors of dependent rvs. Also, when the rv X_i is defined as a random sum of independent rvs or when the distribution of the rv X_i is defined as a mixture of distribution, simulations methods are explained in Chapter 2.

13.4 Monte Carlo simulation method

13.4.1 General idea

Let X be a rv with cdf F_X and quantile function F_X^{-1}.

Recall that the expectation of the rv X can be represented as

$$E[X] = E\left[F_X^{-1}(U)\right] = \int_0^1 F_X^{-1}(u)\,du.$$

Then, we can approximate $\theta = E[X]$ by the empiral mean i.e.

$$\theta \simeq \widehat{\theta}_{(m)} = \frac{1}{m}\sum_{j=1}^{m} X^{(j)} = \frac{1}{m}\sum_{j=1}^{m} F_X^{-1}\left(U^{(j)}\right).$$

By the Law of Large Numbers, the approximation $\widehat{\theta}_{(m)}$ converges toward θ with probability 1 when the number m of samples of X tends toward ∞.

Now, we want to compute $\theta = E\left[g\left(X\right)\right]$ (assuming that the expectation exists).

Assume that we are able to simulate samples $X^{(1)}$, ..., $X^{(m)}$ of the rv X.

The objective the Monte Carlo simulation method is to approximate using the samples $X^{(1)}$, ..., $X^{(m)}$.

The approximation of $\theta = E\left[g\left(X\right)\right]$ is

$$\theta \simeq \widehat{\theta}_{(m)} = \frac{1}{m}\sum_{j=1}^{m} g\left(X^{(j)}\right).$$

Again, due to the Law of Large Numbers, the approximation $\widehat{\theta}$ converges toward θ with probability 1 when the number m of samples of X tends toward ∞.

13.4.2 Variance of the approximation

The variance of $\widehat{\theta}_{(m)}$ is

$$\mathsf{Var}\left(\widehat{\theta}_{(m)}\right) = \frac{1}{m}\mathsf{Var}\left(g\left(X\right)\right).$$

Given the Central Limit Theorem, the error $\left(\widehat{\theta}_{(m)} - \theta\right)$ is asymptotically normal with mean 0 and standard deviation $\frac{\sqrt{\mathsf{Var}(g(X))}}{\sqrt{m}}$.

It implies that the quality of the approximation improves at a rate of order $\sqrt{m}$.

The term $\frac{\sqrt{\mathsf{Var}(g(X))}}{\sqrt{m}}$ is usually called the standard error.

Often, it is impossible to compute $\mathsf{Var}\left(g\left(X\right)\right)$.

Then, we estimate its value using the classical sampling estimator

$$\widehat{\mathsf{Var}}\left(g\left(X\right)\right) = \frac{1}{m-1}\sum_{j=1}^{m}\left(g\left(X^{(j)}\right) - \widehat{\theta}_{(m)}\right)^{2}. \quad (24)$$

Based on the asymptotic distribution of the error $\left(\widehat{\theta}_{(m)} - \theta\right)$, we can build a confidence interval of level α (e.g. 95 %) for $\widehat{\theta}_{(m)}$ for which the lower and the upper bounds are

$$\widehat{\theta}_{(m)} \pm \frac{\sqrt{\mathsf{Var}\left(g\left(X\right)\right)}}{\sqrt{m}}\Phi^{-1}\left(1 - \frac{(1-\alpha)}{2}\right).$$

Generally, since $\sqrt{\mathrm{Var}\left(g\left(X\right)\right)}$ is not known, we use (24) to compute the bounds of the confidence interval

$$\widehat{\theta}_{(m)} \pm \frac{\sqrt{\widehat{\mathrm{Var}}\left(g\left(X\right)\right)}}{\sqrt{m}}\Phi^{-1}\left(1-\frac{(1-\alpha)}{2}\right).$$

13.4.3 Example #1

Let $X_1 \sim LN\left(6, 1^2\right)$ and $X_2 \sim Ga\,(0.6, 1/1000)$ be pair of independent rvs.

We define $S = X_1 + X_2$.

The samples from the uniform distribution are simulated with the PNRG defined in algorithm 51 assuming $x_0 = 2012$ (seed).

We simulate $m = 1\ 000\ 000$ samples of the pair of rvs (X_1, X_2) in order to simulate m samples of S.

We provide in the following table the two first samples of the pair of rvs (X_1, X_2) and of the rv S :

j	$X_1^{(j)}$	$X_2^{(j)}$	$S^{(j)}$
1	69.0344	416.8423	485.8767
2	418.63450	128.2339	546.8684

We know that $E[S] = 1265.1416$ and $\mathrm{Var}(S) = 1\ 360\ 191$.

Based on the m samples, the approximations of $E[S]$ and $\mathrm{Var}(S)$ are 1266.251 and 1 365 455.

One important quantity in actuarial science is the stop-loss premium.

The *stop-loss* $\pi_S(d)$ function corresponds to expectation of a function $g(x) = \max(x-d; 0)$ such that

$$\pi_S(d) = E[\max(S-d; 0)], \tag{25}$$

where $d \in \mathbb{R}$.

We provide in the table below the approximations for the stop-loss premium

$$\pi_S(2000k) = E[\max(S - 2000 \times k; 0)],$$

for $k = 1, 2$.

We also provide the lower and the upper bounds of the confidence interval (level of confidence = 95 %) for those approximations :

k	lower bd	approx. $\pi_S(2000k)$	upper bd	standard error
1	210.8036	212.2769	213.7501	0.7517
2	45.6727	46.5121	47.3516	0.4283

13.5 Computation of risk measures

13.5.1 Method

Let X be rv with cdf F_X and quantile function F_X^{-1}.

We have simulate m sampled values $X^{(1)}$, ..., $X^{(m)}$ of the rv X.

An approximation of the cdf F_X is obtained with the empirical function defined by

$$F^{(m)}(x) \simeq \frac{1}{m}\sum_{j=1}^{m} 1_{\{X^{(j)}\leq x\}}. \tag{26}$$

The approximation to the $VaR_{\kappa}(X)$ is derived from $F^{(m)}$ using the basic definition of a quantile i.e.

$$VaR_{\kappa}(X) \simeq F^{(m)-1}(\kappa) = \inf\left\{X^{(j)}, i = 1, 2, ..., m \;; F^{(m)}\left(X^{(j)}\right) \geq \kappa\right\}.$$

Let $\left(X^{[1]}, ..., X^{[m]}\right)$ be the sorted sampled values (in increasing order) of the rv X.

We fix j_0 such that $F^{(m)-1}(\kappa) = X^{[j_0]}$.

Using (14) and (15) for the TVaR jointly with (26), the approximation of

$TVaR_{\kappa}(X)$ is given by

$$
\begin{aligned}
& TVaR_{\kappa}(X) \\
\simeq\ & \frac{1}{1-\kappa}\left(\frac{1}{m}\sum_{j=1}^{m} X^{(j)} \times 1_{\left\{X^{(j)}>F^{(m)-1}(\kappa)\right\}}\right) \\
& +\frac{1}{1-\kappa}\left(F^{(m)-1}(\kappa)\left(F^{(m)}\left(F^{(m)-1}(\kappa)\right)-\kappa\right)\right) \\
\simeq\ & \frac{1}{1-\kappa}\left(\frac{1}{m}\sum_{j=j_0+1}^{m} X^{[j]} + X^{[j_0]}\left(F^{(m)}\left(X^{[j_0]}\right)-\kappa\right)\right). \qquad (27)
\end{aligned}
$$

Assume that $m \times \kappa$ is an integer (which is the case in practice).

Then, this integer is j_0.

It follows that $\left(F^{(m)}\left(X^{[j_0]}\right)-\kappa\right)=0$ and (27) becomes

$$\begin{aligned} TVaR_\kappa(X) &\simeq \frac{1}{m(1-\kappa)}\sum_{j=1}^{m} X^{(j)}\times 1_{\left\{X^{(j)}>X^{[j_0]}\right\}} \\ &= \frac{1}{m-j_0}\sum_{j=j_0+1}^{m} X^{[j]}, \end{aligned}$$

which corresponds to the empiral mean of the $m-j_0$ largest sampled values of the rv X.

Note that (27) and (23) are similar.

We illustrate the method with this very simple example.

Example 60 *Let X_1 and X_2 be independent rvs with $X_1\sim Exp\left(\frac{1}{100}\right)$ and $X_2\sim Pa(4,300)$. We define $S=X_1+X_2$ where $E[S]=200$ and*

$\mathrm{Var}(S) = 30\ 000$. *In the following table, we provide the sampled values* $U_1 \sim U(0,1)$ *and* $U_2 \sim U(0,1)$. *From those values, we simulate the sampled values of the rvs* X_1 *and* X_2 *:*

j	$U_1^{(j)}$	$U_2^{(j)}$	$X_1^{(j)}$	$X_2^{(j)}$	$S^{(j)}$
1	0.260	0.018	30.111	1.365	31.476
2	0.003	0.536	0.300	63.489	63.789
3	0.627	0.550	98.618	66.284	164.902
4	0.965	0.238	335.241	21.094	356.335
5	0.832	0.909	178.379	246.212	424.591
6	0.590	0.931	89.160	285.341	374.501

The approximations of $F_S(400)$ *and* $E[\max(S-350;0)]$ *are respectively* $\frac{5}{6}$ *and* 17.571. *Then, the approximation of* $VaR_{0.75}(S)$ *is 374.501 and the approximation of* $TVaR_{0.75}(S)$ *is* 407.894. $\square$

13.5.2 Practical considerations

Let $\kappa = 0.99$.

If $m = 100$, then the approximations of $VaR_{0.99}(X)$ and $TVaR_{0.99}(X)$ are $X^{[99]}$ and $X^{[100]}$ (i.e. the approximation of $TVaR_{0.99}(X)$ involves the empiral mean with only one term).

If $m = 10\ 000$, then $VaR_{0.99}(X)$ and $TVaR_{0.99}(X)$ are approximated respectively $X^{[9900]}$ and

$$\frac{\sum_{j=9901}^{10\ 000} X^{[j]}}{100}.$$

If $m = 1\ 000\ 000$, then $VaR_{0.99}(X)$ and $TVaR_{0.99}(X)$ are approximated respectively $X^{[9900]}$ and

$$\frac{\sum_{j=990\ 001}^{1\ 000\ 000} X^{[j]}}{10\ 000}.$$

It means that one needs to simulate a large number m of samples of the rv X in order to obtain suitable approximations of $VaR_{\kappa}(X)$ and $TVaR_{\kappa}(X)$.

The number m of samples increases with the value of κ.

A simple solution to this challenge is to simulate more and more samples, which may difficult in some contexts. Another one is to use more performant sampling methods. However, these are out of scope of this monograph.

13.6 Confidence intervals for the VaR

The idea is to build a confidence interval for a quantile.

We consider the approach based on the binomial distributubion.

Let $\left(X^{[1]}, ..., X^{[m]}\right)$ be the sorted sampled values (in increasing order) of the rv X.

We fix j_0 such that $F^{(m)-1}(\kappa) = X^{[j_0]}$.

The number $R_{m,\kappa}$ of samples which are smaller or equal to $X^{[j_0]}$ follows a binomial distribution with parameters m and κ (i.e. $R_{m,\kappa} \sim Bin(m, \kappa)$).

The confidence interval with a confidence level α is $\left(X^{[j_1]}, X^{[j_2]}\right)$.

The lower bound $X^{[j_1]}$ and the upper bound $X^{[j_2]}$ are detetmined such that the probability

$$\Pr\left(X^{[j_1]} < VaR_\kappa(X) \leq X^{[j_2]}\right) = \Pr(j_1 < R_{m,\kappa} \leq j_2)$$

is as closed as possible to α.

To determine the values of j_1 and j_2 (and consequently the values of $X^{[j_1]}$ and $X^{[j_2]}$), we proceed as follows.

Assume that the number m of samples is large. Then, we approximate the binomial distribution by the normal distribution. We compute j_{diff} as the rounded value of $\left(\sqrt{\mathrm{Var}(R_{m,\kappa})} \times \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right)$, where $\mathrm{Var}(R_{m,\kappa}) = m \times \kappa \times (1 - \kappa)$.

Then, we fix $j_1 = j_0 - j_{diff}$ and $j_2 = j_0 + j_{diff}$.

It follows that the lower bound and the upper bounds of the confidence inter- valvare $X^{[j_0 - j_{diff}]}$ et $X^{[j_0 + j_{diff}]}$.

14 Challenges for research in actuarial science

- A lot of research is going about the risk measures.
- The choice of the risk measure has a significant impact.
- Statiscal methods and risk measures.
- Aggregation methods for independent and dependent risks
- Advanced simulation methods.
- Etc.