Insurance: Mathematics and Economics 52 (2013) 599–605

Contents lists available at SciVerse ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

Choosing a random distribution with prescribed risks

Ignacio Cascos [a], Ilya Molchanov [b,*]

[a] Department of Statistics, Universidad Carlos III de Madrid, Av. Universidad 30, E-28911 Leganés (Madrid), Spain

[b] Department of Mathematical Statistics and Actuarial Science, University of Berne, Sidlerstr. 5, CH-3012 Berne, Switzerland

HIGHLIGHTS

- We survey several algorithms for the simulation of random probability distributions.
- We show how to simulate random distributions with given risk measures.
- We consider the case of vanilla risk measures: the Value at Risk and the Expected Shortfall.
- We suggest a new algorithm for simulating random distributions with given spectral risk measure.

ARTICLE INFO

Article history:
Received June 2011
Received in revised form
November 2012
Accepted 26 March 2013

JEL classification:
C15

Subject category:
IM01
IE43

Keywords:
Risk measure
Random probability distribution
Simulation
Lorenz curve

ABSTRACT

We describe several simulation algorithms that yield random probability distributions with given values of risk measures. In case of vanilla risk measures, the algorithms involve combining and transforming random cumulative distribution functions or random Lorenz curves obtained by simulating rather general random probability distributions on the unit interval. A new algorithm based on the simulation of a weighted barycentres array is suggested to generate random probability distributions with a given value of the spectral risk measure.

1. Introduction

Simulation methods for random variables and stochastic processes form an essential part of applied probability and statistics. The most common situation consists in simulating random variables with given distributions. In the financial context this would correspond to simulating, e.g. asset prices, gains or losses, if their distribution is known.

However, in many practical situations the distribution from which the data come is not specified. In these circumstances, it is necessary first to simulate the distribution and then sample random variables from it. A simple example is the stock price that follows a log-normal distribution with unknown volatility. Such distributions form a parametric family, and it is easy to simulate first the random volatility value and then sample from the corresponding log-normal distribution. In the Bayesian statistics framework, the distribution of the volatility could be termed the prior distribution; the observed data are then used to arrive at the posterior distribution of the unknown parameter.

In a non-parametric context, one needs a prior distribution on the family of probability measures without artificially restricting attention to a particular subfamily of probability distributions, e.g. the normal or log-normal ones. Motivated by this, Ferguson (1974) introduced random Dirichlet probability measures so that the values of the measure on any partition follow the Dirichlet distribution. It is known that these random probability measures are discrete with probability one.

Several known constructions of random continuous probability measures are mentioned in Section 2; see also Monticino (2001) for a comprehensive and instructive survey. Some of these constructions result in absolutely continuous probability distributions, while others always produce singular ones. It is essential that the constructed random probability distributions have full support in

* Corresponding author. Tel.: +41 31 631 8801; fax: +41 31 631 3870.
E-mail addresses: ignacio.cascos@uc3m.es (I. Cascos), ilya@stat.unibe.ch, ilya.molchanov@stat.unibe.ch (I. Molchanov).

0167-6687/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.insmatheco.2013.03.014

the space of all probability measures, meaning that each weakly open set of probability measures with positive probability contains a simulated random distribution.

A further body of work concerns constructions of random probability measures with given mean or the distribution of the mean; see Hill and Monticino (1998). Random measures with given mean and variance have been constructed in Bloomer and Hill (2002).

In this paper we address the problem of choosing at random a continuous probability measure on the real line which has particular values of risk measures associated with it, in particular the value at risk and the expected shortfall; see, e.g., Föllmer and Schied (2004) and the more recent study of Goovaerts et al. (2010), where also relationships to Actuarial Science are emphasised. These risk measures are recalled in Section 3 and their connection to Lorenz curves is mentioned. Several operations suitable to combine Lorenz curves are discussed in Section 4.

The construction of random probability distributions with prescribed risks described in Section 5 is based on the calculation of the generalised random Lorenz curves associated with simulated random cumulative distribution functions. The key tool consists in transforming random Lorenz curves and combining several of them.

Our simulation technique uses as input any construction of random probability measures (possibly with a given mean) so that it leads either to absolutely continuous, discrete, or singular measures depending on the choice of the underlying algorithm. In the case of continuous probability measures, it yields random distributions with the full support property. It means that the simulated random distribution with a positive probability belongs to each open set of probability measures intersected with the family of probability measures having given values of risk measures and given essential infimum and essential supremum.

A lot of attention in Actuarial Science is devoted to finding bounds of certain risk measures for random variables with some fixed characteristics, for example, moments or values of other risk measures; see Goovaerts et al. (2011) and Kaas et al. (2009). This problematic is also extremely important in the case of several risks, whose dependence structure is unknown or complicated, with the aim to bound the combined risk; see Embrechts et al. (2005); Embrechts and Puccetti (2006) and also Kaas et al. (2009). While the worst case bounds are extremely important, they may be substantially wider than those encountered in most realised cases. In view of this, the decision maker may be interested in average bounds that would be derived with the help of simulating a number of random distributions. This may be termed as Monte Carlo methods in the space of probability measures.

Simulating random distributions with given values of risk measures (reflecting prior beliefs and limited knowledge) may be helpful to compare various possible scenarios and try various strategies in order to find one that performs better on average or in the prescribed majority of cases. Simulating random distributions is useful to explore the variability of other features of probability distributions with given risk measures and also for training purposes for prospective decision makers.

2. Continuous random distribution functions

A random cumulative distribution function (cdf) is a function sampled at random from the family of non-decreasing right-continuous functions on the real line such that their limit at $-\infty$ is 0 and at $+\infty$ is 1. By mapping the real line into $[0, 1]$, it is possible to reduce this simulation problem to choosing a probability distribution on $[0, 1]$. The following well-known algorithms yield random distributions with support equal to $[0, 1]$.

Dubins and Freedman (1967). Algorithm DF starts with selecting a point at random in the unit square $S = [0, 1]^2$ according to a probability distribution μ, deleting the upper left and lower right rectangles defined by this point and selecting one point at random according to the rescaled μ in each of the remaining rectangles. The procedure is repeated, i.e. at each step a new point is selected in each of the remaining rectangles so that 2^{n-1} points are selected at step n and 2^n rectangles obtained. The union of the remaining rectangles converges to the graph of a random cumulative distribution function. The obtained distribution is continuous if and only if μ assigns probability zero to the vertical edges of S and a positive probability to the interior of S. Furthermore, if μ does not assign probability 1 to the main diagonal of S (i.e. the DF algorithm does not generate the uniform distribution on $[0, 1]$), then the generated random distributions are almost surely singular.

Graf et al. (1986). This special case of the DF algorithm arises if the reference measure μ is supported by $\{1/2\} \times [0, 1]$. The first random point in $[0, 1]^2$ has x-coordinate $1/2$, while its y-coordinate is sampled from μ and becomes the value $F(1/2)$ of the random cdf F. The next points are sampled from $[0, F(1/2)]$ and $[F(1/2), 1]$ using the rescaled μ and their values are assigned to $F(1/4)$ and $F(3/4)$ respectively, etc. In this way, F acquires random values on all dyadic rationals and then is extended by continuity.

The construction of Graf et al. (1986) extends the work by Kraft (1964), who showed how to choose the base measure μ in order to arrive at an almost surely absolutely continuous distribution F. Mauldin and Monticino (1995) suggested a variant of this algorithm allowing possibly different base measures μ at its various steps.

Hill and Monticino (1998) suggested an algorithm to produce a probability distribution at random with a fixed mean (or with a given distribution of the mean) such that the probability distribution is either discrete or continuous almost surely. The algorithm is based on the notion of sequential barycentre array that characterises a probability distribution. The nth row of the barycentre array consists of $2^n + 1$ elements. Those elements $m_{n,k}$, $k = 1, \ldots, 2^n - 1$, occupying even positions are replicated from the previous row ($m_{n,2j} = m_{n-1,j}$) and those at odd positions are conditional expectations (barycentres)

$$m_{n,2j-1} = \mathbf{E}[X | m_{n-1,j-1} < X < m_{n-1,j}].$$

Note that $m_{n,0} = 0$ and $m_{n,2^n} = 1$ for $n \geq 0$, or, in general, are the essential infimum and the essential supremum of the distribution.

Algorithm HM produces a random sequential barycentre array, and thus yields a random cdf. The element $m_{1,1}$ occupies the middle position in each row of the barycentre array and is the mean of the distribution, which can be fixed or simulated at the start of the algorithm. Each barycentre in the nth row occupying an even position is copied from the previous row, while the barycentres occupying an odd position are sampled from an interval formed by two adjacent barycentres from the $(n-1)$st row. For this, one takes independent copies of a random variable U distributed on $[0, 1)$ and calculates $m_{n,4j-1}$ as $m_{n-1,2j-1} + U(m_{n-1,2j} - m_{n-1,2j-1})$, while $m_{n,4j-3}$ is calculated as $m_{n-1,2j-1} - U(m_{n-1,2j-1} - m_{n-1,2j-2})$. If $U = \frac{1}{2}$ a.s., then all subsequent barycentres are taken to be the midpoints of the corresponding intervals and if the first barycentre is located at $\frac{1}{2}$, the resulting distribution is uniform; see Hill and Monticino (1998, Example 3.3). The key issue is to choose the distribution of U which has full support in $[0, 1)$, but allocates more mass to the mid-part of the unit interval. If U is likely to take values close to 0 or 1, they yield short intervals and the resulting random cdf has nearly vertical parts resembling atoms. We use the HM algorithm in our numerical examples.

Random distributions with arbitrary essential infimum and supremum. Random probability distributions with prescribed essential