THE DISTRIBUTION FUNCTION OF A SUM OF DEPENDENT RISKS: A GEOMETRIC-COMBINATORIAL APPROACH

Marcello Galeotti[1], Emanuele Vannucci[2],

[1] *University of Florence, marcello.galeotti@dmd.unifi.it*
[2] *University of Pisa, emanuele.vannucci@unipi.it*

Abstract

The problem of evaluating the distribution of the sum of dependent risks has been treated recently through a new numerical approach proposed in [1], [2], [6], called AEP algorithm. We consider a theoretical result extending the proof of the convergence of this algorithm to any dimension and analyze its use for applications in finance and insurance.

Keywords: dependent risks, copulas, measure of simplexes, computation algorithm.

Introduction

The evaluation of the sum of dependent risks is a main issue for many applications in finance and insurance. One of the reasons for such an interest is the approach of the regulatory standard on bank capital adequacy introduced by the so called Basel II [3], which is based on a quantile (Value-at-Risk) of the distribution function of the potential losses of a bank, collecting several risky positions. There would be no need to recall that in insurance applications the evaluation of aggregate losses of many risks is a very common task, as is the case, for example, with stop-loss reinsurance treaties. Such a subject is addressed by a wide range of literature contributions exploiting the concept of copula, which permits to separate the analysis of the dependence structure from that of the marginal distributions. Whereas the theory based on copula functions allows to measure the effects of dependence among different risks, especially when stress-testing is required, another relevant problem is to assess bounds for the probability of the sum of dependent risks exceeding some threshold. To this purpose [1], [2], [6] introduced a new numerical procedure. In fact the authors proposed an algorithm whose implementation required a bounded density of the multivariate distribution and whose convergence was proven for dimension less or equal than 5. However, in a quite recent breakthrough, [5] has proven the convergence of the algorithm for any dimension. Building on that, we propose an improvement of the algorithm, which, exploiting a combinatorial approach, allows to consider a summation of a smaller number of terms. This way we aim at a considerably faster numerical procedure, which, avoiding some memory constraints, can make the algorithm manageable and competitive with Monte Carlo methods also for higher (i.e. larger than 5) dimensions. The layout of the paper is the following. In the first paragraph we summarize the AEP algorithm and its main results. Then we give a sketch of Galeotti's theorem, which extends the proof of the convergence of the AEP algorithm to any dimension. In the third paragraph some combinatorial details are fully developed in order to find some efficient ways for the numerical implementation of the AEP algorithm. Then we propose the application of the AEP to finance risk evaluation, as Value At Risk and Expected Shortfall evaluation, and insurance risk evaluation, in particular for the Stop-Loss reinsurance treaty, whose analityc description coincides with the Expected Shortfall one. We end by a short list of comments and open problems.

1. AEP algorithm

The evaluation of the sum of dependent risks is a main issue for many applications in finance and insurance (Value-at-Risk, Expected Shortfall, Stop-Loss Reinsurance, ...). The key problem for a positive s is to estimate

$$P[X_1 + \ldots + X_d \leq s]$$

where:
$X_1, \ldots, X_d$ are non-negative (equivalently, bounded from below) r.v.;
$H(x_1, \ldots, x_d)$ is the distribution function;

V_H is the relative probability measure.
This problem has been tackled, generally, by multivariate Monte Carlo procedures exploiting copula theory. In articles [1], [2], [6] the authors propose a new approach based on a numerical procedure utilizing combinatorial arguments derived from shuffles of copulas, concept defined in [4], and used in many successive works.
The main idea of the AEP algorithm is to approximate the H-measure of a d-dimensional simplex $S(\mathbf{0}, s)$ (or, after rescaling, $S(\mathbf{0}, 1)$)

$$V_H\left[\boldsymbol{S}(0,s))\right] = H(x_1, \dots, x_d)$$

by an algebraic sum of H-measures of hypercubes $Q(\boldsymbol{b}, h)$ (overlapping when $d > 2$).
A hypercube $Q(\boldsymbol{b}, h)$, where $\boldsymbol{b} = (b_1, \dots, b_d) \in \mathbb{R}^d$, $h \in \mathbb{R}$, is defined by

$$Q(\boldsymbol{b}, h) = \{ \begin{array}{l} x_k \in (b_k, b_k + h], \forall k = 1, \dots, d, \text{if } h > 0 \\ x_k \in (b_k + h, b_k], \forall k = 1, \dots, d, \text{if } h < 0 \end{array}$$

setting $Q(\boldsymbol{b}, 0) = \varnothing$.
Let $\boldsymbol{i}_0, \dots, \boldsymbol{i}_N$, $N = 2^d - 1$, be the 2^d vertexes of $\{0,1\}^d$ (e.g. $\boldsymbol{i}_0 = \mathbf{0} = (0, \dots, 0)$,$\boldsymbol{i}_N = \mathbf{1} = (1, \dots, 1)$) and let $\#\boldsymbol{i}$ denote the number of 1's in the vector $\boldsymbol{i}$. Then the H-measure of the hypercube $Q(\boldsymbol{b}, h)$ is given by

$$V_H[Q(\boldsymbol{b}, h)] = P[X_k \in (b_k, b_k + h], \forall k = 1, \dots, d] =$$

$$= \sum_{j=0}^{N} (-1)^{d - \#\boldsymbol{i}_j} H(\boldsymbol{b} + h\boldsymbol{i}_j)$$

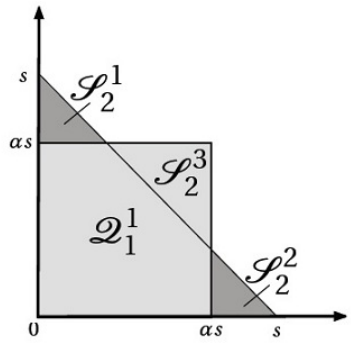
The first step of the AEP algorithm consists in approximating the H-measure of $S(\mathbf{0}, 1)$ by that of a hypercube

$$Q_1^1 = Q(\mathbf{0}, \alpha), \alpha \in [1/d, 1)$$

Then, posed $P_1 := V_H[Q_1^1]$, the procedure can be iterated, so that at the (n+1)-th iteration we have

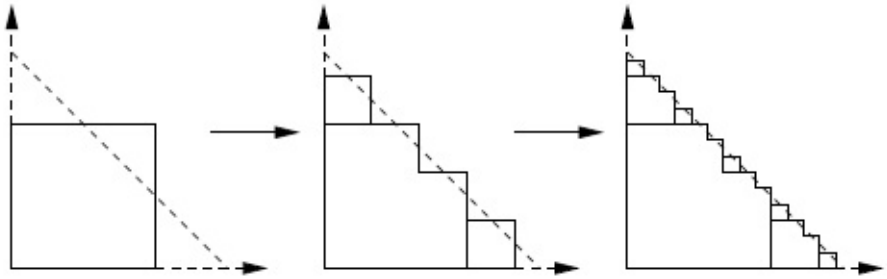
$$V_H[S(\mathbf{0}, 1)] = P_n + \sum_{k=1}^{N^n} \sigma_k^{n+1} \, V_H[S_k^{n+1}]$$

where $\sigma_k^{n+1} = \pm 1$ according if the simplex measure has to be added or subtracted.
Let us observe that when $d = 2$ the new simplexes generated at each step do not overlap.



$$S_1^1 = S(\mathbf{0}, s) = (Q_1^1 \cup S_2^1 \cup S_2^2) \setminus S_2^3, \ \forall \alpha \in [1/2, 1).$$

Actually the following figures show the construction of new hypercubes at each step of the algorithm in dimension two.

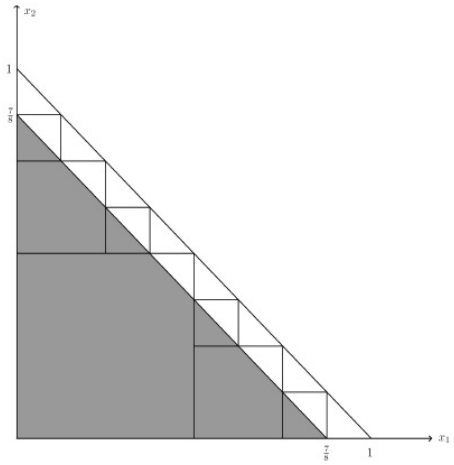


The AEP convergence was proved in [1] for $d \leq 5$ when $\alpha = \frac{2}{d+1}$. Moreover, with a method based on Richardson's extrapolation, the convergence proof was extended to $d \leq 8$ assuming that the joint distribution H had a density V_H with continuous first and second derivatives.

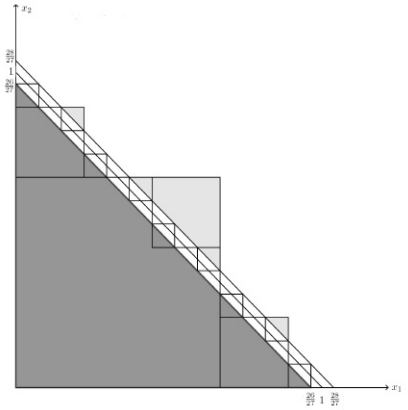
2. Convergence in any dimension: Galeotti extension

In a recent paper [5] the convergence of the AEP in any dimension has been proved. The only required hypothesis is that H has a density bounded in a neighborhood of the simplex diagonal.
The main idea of Galeotti's proof consists in a geometrical approximation of the simplex with hypercubes, initially disregarding probability aspects. This way Galeotti is able to define, at any step of the algorithm, a corresponding sub-simplex of $S(\mathbf{0}, 1)$ which is exactly filled up, by summing positive and negative

hypercubes, while in a suitably chosen strip outside the simplex positive and negative hypercubes exactly compensate.
Then the convergence follows from the assumed boundedness of the density in a neighborhood of the simplex diagonal.
Let us observe that with $d = 2$ and $\alpha = 1/2$, the squares (2-dimensions hypercubes) are disjoint, so that the sub-simplex is exactly filled by the hypercubes generated at the n-th step. The following figure shows the idea of this geometrical approximation



The further step in Galeotti's proof consists in exploiting the self-similarity of the construction in order to show that the exact filling holds also when at most one vertex of the hypercubes generated at the n-th step lies outside the simplex. For $n = 2$ and $\alpha \in \left[\frac{1}{2}, \frac{2}{3}\right]$ the figure below gives an insight of such a result.



Finally Galeotti exploits a Lemma, stating that the AEP algorithm converges for the Lebesgue measure when $d \geq 2$ and $\alpha \in \left[\frac{1}{d}, \frac{1}{\sqrt[d]{d!}}\right]$, plus some analytical arguments whose details are shown in [5]

3. Some combinatorial arguments

In order to focus on the combinatorial arguments used in the geometrical approach, let us start with the case $d = 3$. The volume of the simplex $S(\mathbf{0}, s)$, i.e.

$$V = vol\left[(0,0,0),(s,0,0),(0,s,0),(0,0,s)\right] = \frac{s^3}{6}$$

can be evaluated by adding and subtracting cubes. Let us observe that the cube with vertexes

$$[(0,0,0),(as,0,0),(0,as,0),(as,as,0),$$

$$(0,0,as),(as,0,as),(0,as,as),(as,as,as)]$$

has volume $V(a) = (as)^3$

If $\frac{1}{3} \leq a \leq \frac{1}{2}$ there is no overlapping and we have $V(\frac{1}{2}) = \frac{s^3}{8}$, $V(\frac{1}{3}) = \frac{s^3}{27}$

If we aim at a better approximation of the volume of the simplex we have to consider 7 simplexes:

1 of type A to subtract

$$A = [(sa, sa, (1-2a)s), (sa, s(1-2a), as),$$

$$(s(1-2a), sa, as), (sa, sa, sa)]$$

with volume V(A)=

$$\frac{1}{6}\det\begin{pmatrix} s(3a-1) & 0 & 0 \\ 0 & s(3a-1) & 0 \\ 0 & 0 & s(3a-1) \end{pmatrix} = \frac{s^3(3a-1)^3}{6}$$

3 of type B to add

$$B = [(s,0,0),(as,0,0),(as,s(1-a),0),(sa,0,(1-a)s)]$$

with volume V(B)=

$$\frac{1}{6}\det\begin{pmatrix} (1-a)s & (1-a)s & (1-a)s \\ 0 & s(1-a) & 0 \\ 0 & 0 & s(1-a) \end{pmatrix} = \frac{s^3(1-a)^3}{6}$$

and 3 of type C to subtract

$$C = [(sa, sa, 0), (s(1-a), sa, 0), (as, s(1-a), 0), (sa, sa, (1-2a)s)]$$

with volume V(C)=

$$\frac{1}{6}\det\begin{pmatrix} 0 & (2a-1)s & 0 \\ 0 & 0 & (2a-1)s \\ s(1-2a) & s(1-2a) & s(1-2a) \end{pmatrix} = \frac{s^3(1-2a)^3}{6}$$

So, it is easy to check that

$V = V(a) - V(A) + 3V(B) - 3V(C) = = (as)^3 - \frac{s^3(3a-1)^3}{6} + 3 \cdot \frac{s^3(1-a)^3}{6} - 3 \cdot \frac{s^3(1-2a)^3}{6} = \frac{s^3}{6}$

Extending these result to a generic dimension d we get

$$(as)^d - \binom{d}{d}\frac{s^d(da-1)^d}{d!} + \binom{d}{1}\cdot\frac{s^d(1-a)^d}{d!} - \binom{d}{2}\cdot\frac{s^d(1-2a)^d}{d!} +$$
$$+\binom{d}{3}\cdot\frac{s^d(1-3a)^d}{d!} - \ldots + (-1)^d\binom{d}{d-1}\cdot\frac{s^d(1-(d-1)a)^d}{d!} = \frac{s^d}{d!}$$

The basic combinatorial formula to apply is

$$(d!)\,a^d + \sum_{h=0}^{d}(-1)^{h+1}\binom{d}{h}(1-ha)^d = 0$$

from which

$$\frac{s^d}{d!} = (as)^d + \frac{s^d}{d!}\sum_{h=1}^{d}(-1)^{h+1}\binom{d}{h}(1-ha)^d \equiv$$

$$\equiv (as)^d + s^d M = (as)^d\left(1+\frac{1}{a^d}M\right)$$

with $M = \frac{1}{d!} - a^d$

From the same formula, replacing s by $1-ha$, it follows

$$\frac{(1-ha)^d}{d!} = (a\,(1-ha))^d + (1-ha)^d\,M$$

from which

$$\frac{s^d}{d!} = \ldots = (as)^d\left(\frac{1-(d!M)^{n+1}}{1-d!M} + \frac{(d!)^n}{a^d}M^{n+1}\right)$$

Then, letting $n\to\infty$, we get

$$(as)^d\,\frac{1}{1-d!\left(\frac{1}{d!}-a^d\right)} = \frac{s^d}{d!}$$

if

$$\left|d!\left(\frac{1}{d!}-a^d\right)\right| = |M| < 1$$

with $-1 < 1-d!a^d < 1$, $-2 < -d!a^d < 0$, $0 < d!a^d < 2$, $0 < a < \sqrt[d]{\frac{2}{d!}}$.

So, by the previous results, it clearly emerges that a good choice is in fact $a = \frac{2}{d+1}$, since, by an induction argument, it is easy to prove that $\frac{2}{d+1} < \sqrt[d]{\frac{1}{d!}}$ $\forall d = 1, 2, 3, \ldots$.

Furthermore the above results allow to obtain the value of a for the maximum convergence in any dimension, that is

$a = \sqrt[d]{\frac{1}{d!}}$

$d = 2$ is $a = \sqrt{\frac{1}{2}} = 0.70711$

$d = 3$ is $a = \sqrt[3]{\frac{1}{3!}} = 0.55032$

$d = 4$ is $a = \sqrt[4]{\frac{1}{4!}} = 0.4518$

$d = 10$ is $a = \sqrt[10]{\frac{1}{10!}} = 0.22081$

Hence, when $d \to +\infty$, by Stirling formula $\boldsymbol{a \simeq \frac{e}{d}}$.

4. Finance and insurance applications

In this paragraph we describe the application of the AEP to finance risk evaluation, by computations of Value At Risk and Expected Shortfall, and to insurance risk evaluation, since in particular the analytic description of the Stop-Loss reinsurance treaty coincides with that of the Expected Shortfall. So in the following the ES evaluation can be identified with that of the ceded risk in Stop Loss reinsurance.

Let $\boldsymbol{X} = (X_1, \dots, X_d)$ be the vector of r.v. describing d random losses with generic marginals $F_1, \dots, F_d$ and joint distribution function H. The Expected Shortfall S_C with threshold level K is defined as

$$S_C = \{ \begin{array}{l} 0 \text{ if } X_1 + \ldots + X_d < K \\ X_1 + \ldots + X_d - K \text{ if } X_1 + \ldots + X_d \geq K \end{array}$$

with the generic r-th moment given by

$$E_H(S_C^r) = \int_{x_1+\ldots+x_d \geq K} (x_1 + \ldots + x_d - K)^r \, dV_H$$

where V_H is the probability measure.

Let us consider the subspace
$\{x_1, \ldots, x_d | x_1 + \ldots + x_d \geq K\} \subset \mathbb{R}^d$
divided in "strips" whose extreme values are denoted by s_i and s_{i+1}, s.t. $s_{i+1} > x_1 + \ldots + x_d \geq s_i$.
Each value s_i identifies a simplex $S(\mathbf{0}, s_i)$ whose probability measure can be evaluated as above.
An estimate of $E_H(S_C^r)$ is given by

$$E_H^N = \sum_{i=0}^{N-1} \left(V_H \left[S(\mathbf{0}, s_{i+1}) \right] - V_H \left[S(\mathbf{0}, s_i) \right] \right) \left[\frac{s_{i+1} + s_i}{2} - K \right]^r$$

$s_N = L$, $L >> K$, $s_0 = K$.
The estimate converges to the correct value if, for $L \to +\infty$, $s_{i+1} - s_i \to 0$. Precisely, consider an iterating procedure through steps $n = 1, 2, 3 \ldots$. Then $L(n) \to +\infty$ and $s_{i+1}^n - s_i^n \to 0$ implies $\frac{L(n)}{N(n)} \to 0$. As an example, choosing a suitable $l > 0$, set $L(n) = K + 2^{n-1} l$ and $s_{i+1}^n - s_i^n = 2^{-2^{n-1}+n-1} l$, so that $N(n) = 2^{2^{n-1}}$. In this case each partition is a refinement of the previous one plus the addition of an equal number of sub-intervals (as $\frac{s_{N(n+1)} - K}{s_{N(n)} - K} = 2$). Obviously there is a trade-off between estimation accuracy and computation

time. In general it will be necessary to define the most convenient sequence $s_{i+1} = g(s_i)$, starting from $s_0 = K$.

Let us observe that the convergence problem for the VaR is "*easier*", since an estimate is given by s such that

$$V_H\left[S(\mathbf{0}, s)\right] = \sum_{i=0}^{N-1} \left(V_H\left[S(\mathbf{0}, s_{i+1})\right] - V_H\left[S(\mathbf{0}, s_i)\right]\right) = 1 - \alpha$$

$s_N = s$, $s_0 = 0$
So, fixed $\epsilon \in \mathbb{R}$ arbitrarily small, the VaR estimation converges to the correct value s for each sequence s_i, $i = 0, 1, 2, \ldots$ such that

$$\exists i^* : |s_{i^*} - s| < \epsilon$$

The convergence of the estimation for Expected Shortfall is a harder problem. We can construct two bounds for the estimation of the generic r-th moment $E_H(S_C^r)$, considering the positive skewness in each interval (s_i, s_{i+1}), that is

$$E_{min} = \sum_{i=0}^{N-1} \left(V_H\left[S(\mathbf{0}, s_{i+1})\right] - V_H\left[S(\mathbf{0}, s_i)\right]\right)\left[s_i - K\right]^r$$

$$E_{med} = \sum_{i=0}^{N-1} \left(V_H\left[S(\mathbf{0}, s_{i+1})\right] - V_H\left[S(\mathbf{0}, s_i)\right]\right)\left[\frac{s_{i+1} + s_i}{2} - K\right]^r$$

$s_N = L$, $L >> K$, $s_0 = K$.
The two bounds converge to the correct value if, following the above observations, $\lim_{n\to+\infty} E_{med} - E_{min} = 0$, with an obvious trade-off between estimation accuracy and computation time.

5. Numerical results: standard scenario

We consider a scenario with Pareto marginals,
$F_i(x) = 1 - (1 + x)^{-\theta_i}$, $x \geq 0$, $\theta \geq 0$, so that $E[X] = \frac{1}{\theta - 1}$ when $\theta \geq 1$,
and a dependence structure given by a Clayton copula , $i = 1, \ldots, d$ with $\delta \in (0, \infty)$.
Hence, if $\delta \to 0$ the copula structure tends to independence, while if $\delta \to \infty$ the copula structure tends to comonotonicity.
As to the distribution parameters we set $\delta = 5$, $d = 2$ and for the threshold level of Expected Shortfall we assume $K = 20$.
For the implementation of the algorithm we define
$s_{i+1} = s_i + \gamma$
$n = 4$. Moreover, as shown in the previous paragraph, in order to have the maximum rate of convergence with $d = 2$, we choose $\alpha = \frac{1}{\sqrt{2}}$.
If we assume $\theta = 2$ (then $E[X] = 1$) we obtain the following sequence

	γ, L	1000	2000	4000	10000
E_{med}	1	0.109	0.111	0.112	0.113
E_{min}	1	0.106	0.108	0.109	0.110
E_{med}	2	0.109	0.112	0.113	0.113
E_{min}	2	0.103	0.105	0.107	0.109

which seems to be quite stable as L increases. However, if we consider a higher expected value of the single loss, for example assuming $\theta = 1.2$ (implying $E[X] = 5$), we obtain

	γ, L	1000	2000	4000	10000
E_{med}	1	2.805	3.198	3.541	3.927
E_{min}	1	2.773	3.167	3.509	3.894
E_{med}	2	2.806	3.199	3.542	3.928
E_{min}	2	2.742	3.135	3.478	3.863

Then the level of L seems to be not sufficiently high in order to show some convergence of the estimation. Let us observe that, concerning the VaR computation, we get, with $\theta = 2$, $VaR(99\%) = 15.96$, since $V_H\left[S(\mathbf{0}, 15.96)\right] = 0.990001$, and, with $\theta = 1.2$, $VaR(99\%) = 90.73$, since $V_H\left[S(\mathbf{0}, 90.73)\right] = 0.9900005$.

It must be noted that in the case of VaR the algorithm can be stopped at a relatively small value of L, with an obvious advantage in computation time. Developing the estimation of Expected Shortfall in the case of $\theta = 1.2$, we seek a level of L allowing some "stability".
Let us observe that $L = 20000$ yields $E_{med} = 4.176$, whereas considering the larger value $L = 40000$ the estimation increases to $E_{med} = 4.392$, and so on, until $E_{med} = 4.821$ with $L = 200000$. There is a clear evidence that fixing the level of the threshold L at which the algorithm can be stopped seems to be absolutely crucial for the quality of estimation.

5. Comments and open problems

The AEP algorithm may be an alternative to Montecarlo methods for multivariate stochastic models; the convergence of the estimation is an *easier* problem for VaR and a less easy one for Expected Shortfall and Stop Loss reinsurance.
The future research should aim at two goals:
1) extending the AEP numerical implementation to higher dimensions;
2) speeding up the algorithm.
One possibility for the speed-up may consist in a randomization of the hypercubes which have to be added and subtracted at each step of the algorithm.
Then we could try to show a possible convergence to the correct value of $V_H[S(\mathbf{0}, s)]$, by extrapolating a sequence of estimates constructed through an increasing number $m_1 < m_2 < \ldots$ of hypercubes (semi-randomly?) selected from a total of N^n at the $n - th$ step.

6. References

[1] Arbenz P., Embrechts P., Puccetti G., (2011) The AEP algorithm for the fast computation of the distribution of the sum of dependent random variables. Bernoulli 17(2), 562-591.

[2] Arbenz, P., Embrechts, P., and G. Puccetti (2012), The GAEP algorithm for the fast computation of the distribution of a function of dependent random variables. Stochastics 84(5-6), 569-597

[3] Basel Committee on Banking Supervision (2006), International Convergence of Capital Measurement and Capital Standards.

[4] Durante F., Sarkoci P., Sempi C. (2009), Shuffles of copulas. Journal of Mathematical Analysis and Applications 352(2), 914-921

[5] Galeotti M. (2015), Computing the distribution of the sum of dependent random variables via overlapping hypercubes. Decisions in Economics and Finance 38(2), 231-255.

[6] Puccetti, G. and L. Rüschendorf (2015), Computation of sharp bounds on the expected value of a supermodular function of risks with given marginals. Commun. Stat. Simulat. 44(3), 705-718.