ASTIN / AFIR-ERM COLLOQUIUM
PANAMA CITY, PANAMA
20 - 24 AUGUST, 2017

The distribution 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*

The problem

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

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

where
$X_1, ..., X_d$ non-negative (bounded from below) r.v.
$H(x_1, ..., x_d)$ distribution function
V_H probability measure

Outline

1. AEP algorithm
2. Convergence in any dimension: Galeotti extension
3. Application to VaR and ES
4. The convergence speed problem
5. Open problems
6. Main references

1. AEP algorithm

Arbenz P., Embrechts P., Puccetti G. (Bernoulli, 2011): The AEP algorithm for the fast computation of the distribution of the sum of dependent random variables.

AEP algorithm: approximates the H-measure of a d-dimensional simplex $S(\mathbf{0}, s)$ (or by rescaling $S(\mathbf{0}, 1)$)

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

by an algebraic sum of H-measures of hypercubes $Q(\mathbf{b}, h)$ (overlapping when $d > 2$)

1. AEP algorithm

Given $\mathbf{b} = (b_1, \ldots, b_d) \in \mathbb{R}^d$, $h \in \mathbb{R}$

$$Q(\mathbf{b}, h) = \begin{cases} x_k \in (b_k, b_k + h], \forall k = 1, ..., d, \text{if } h > 0 \\ x_k \in (b_k + h, b_k], \forall k = 1, ..., d, \text{if } h < 0 \end{cases}$$

It holds $Q(\mathbf{b}, 0) = \emptyset$.
$\mathbf{i}_0, \ldots, \mathbf{i}_N$, $N = 2^d - 1$, 2^d vectors of $\{0, 1\}^d$
(e.g. $\mathbf{i}_0 = \mathbf{0} = (0, \ldots, 0)$,$\mathbf{i}_N = \mathbf{1} = (1, \ldots, 1)$)
$\#\mathbf{i}$ the numbers of 1's in the vector $\mathbf{i}$

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

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

1. AEP algorithm

The first step: replacing $S(\mathbf{0}, 1)$ by a hypercube

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

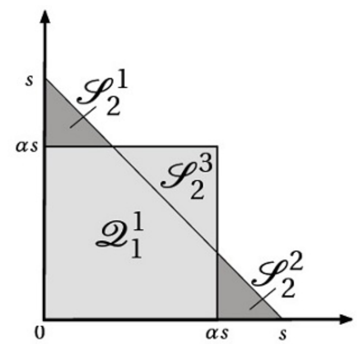
then from $P_1 := V_H[Q_1^1]$ at (n+1)-th iteration

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

where $\sigma_k^{n+1} = -1$ or 1 if the simplex k have to be respectively added or subtracted

1. AEP algorithm

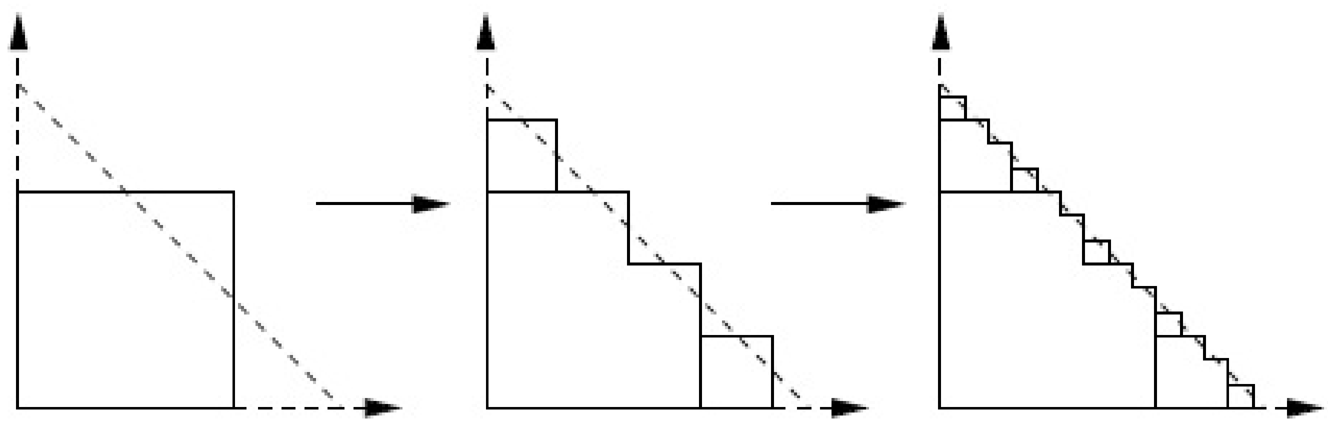
With $d = 2$ the new simplexes generated at each step do not overlap.



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

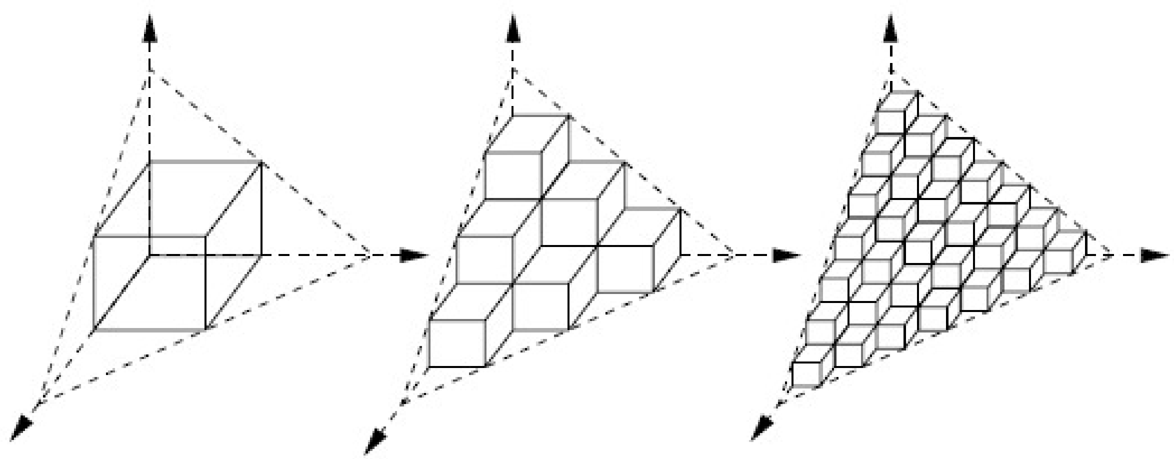
1. AEP algorithm: graphical intuition with $d = 2$

First three steps of AEP



1. AEP algorithm: graphical intuition with $d = 3$

First three steps of AEP



1. AEP algorithm:
convergence in dimension 5 extendible to 8

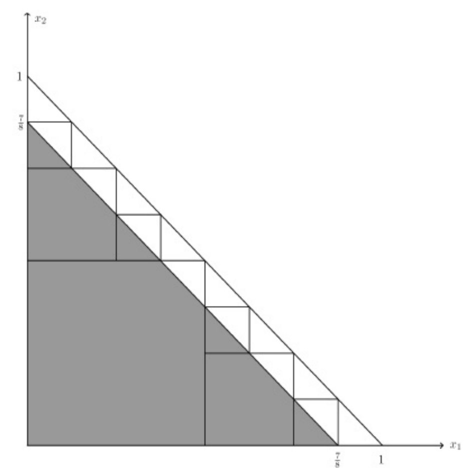
AEP: convergence for $d \leq 5$, when $\alpha = \frac{2}{d+1}$.
With a method based on Richardson's extrapolation technique, AEP convergence is extended to $d \leq 8$ if the joint distribution H has a density V_H with continuous first and second derivatives.

2. Convergence in any dimension: Galeotti extension

Hypothesis: H has a density V_H bounded in a neighborhood of the simplex diagonal.
Idea: geometrical approximation of the simplex (disregarding probability aspects) with hypercubes.
At any step of the algorithm, a corresponding subsimplex of $S(\mathbf{0},1)$ 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.
The simplex is geometrically approximated, the convergence follows from the assumed boundedness of the density in a neighborhood of the simplex diagonal.

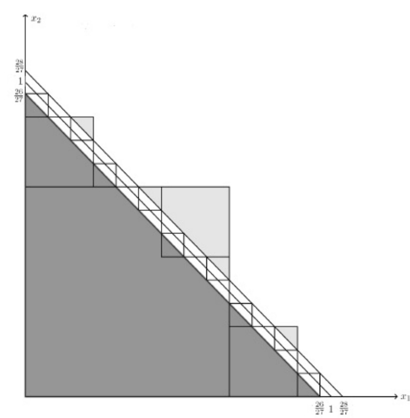
2. Convergence in any dimension: the case $d = 2$ as intuition

With $d = 2$ and $\alpha = 1/2$, the squares (2-dimensions hypercubes) are disjointed then the sub-simplex is exactly filled by the hypercubes generated at n-th step.



2. Convergence in any dimension: the case of $d = 2$ as intuition

By self-similarity we have an exact filling also when at most one vertex of the hypercubes generated at n-th step lies outside the simplex (combinatorial arguments), for $n \geq 1$ and $\alpha \in \left[\frac{1}{2}, \frac{2}{3}\right]$.



2. Convergence in any dimension: the proof

Lemma: AEP algorithm converges for the Lebesgue measure when $d \geq 2$ and $\alpha \in \left[\frac{1}{d}, \frac{1}{\sqrt[d]{d!}}\right]$.

The proof is divided in 5 "natural" steps (full technicalities in 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.

2. Convergence in any dimension: the proof

So a good choice is $\alpha = \frac{2}{d+1}$, since it is
$\frac{2}{d+1} < \sqrt[d]{\frac{2}{d!}}$ $\forall d = 1, 2, 3, ...$ (induction argument)

For maximum convergence, through some combinatorial details we have for α
$a = \sqrt[d]{\frac{1}{d!}}$
$d = 2$ is $\alpha = \sqrt{\frac{1}{2}} = 0.70711$
$d = 3$ is $\alpha = \sqrt[3]{\frac{1}{3!}} = 0.55032$
$d = 4$ is $\alpha = \sqrt[4]{\frac{1}{4!}} = 0.4518$
$d = 10$ is $\alpha = \sqrt[10]{\frac{1}{10!}} = 0.22081$
$d \to +\infty$ is (Stirling formula) $\alpha = \frac{e}{d}$

3. Application to VaR and ES

Let $\mathbf{X} = (X_1, ..., X_d)$, the vector of r.v. which describes d random losses with generic marginals $F_1, ..., F_d$ and joint distribution function H.
VaR at level α is defined

$$VaR(\alpha) = inf\,\{K \in \mathbb{R} | P(X_1 + ... + X_d \leq K) = 1 - \alpha\}$$

For the random Shortfall S with threshold K, we have

$$S(\alpha) \propto (X_1 + ... + X_d) | (X_1 + ... + X_d) > K, K = VaR(\alpha)$$

with the generic r-th moment given by

$$E_H(S^r) =\propto \int_{x_1+...+x_d \geq K} (x_1 + ... + x_d)^r \, dV_H$$

where V_H is the probability measure.

3. Application to VaR and ES

Let consider the subspace

VaR

$\{x_1, ..., x_d | x_1 + ... + x_d \leq K\} \subset \mathbb{R}^d$

Expected Shortfall

$\{x_1, ..., x_d | x_1 + ... + x_d > K\} \subset \mathbb{R}^d$

divided into "stripes" defined by the extreme values s_i and s_{i+1}, s.t. $s_{i+1} > x_1 + ... + x_d \geq s_i$.

Each value s_i identifies a simplex $S(\mathbf{0}, s_i)$ which can be evaluated as stated before.

3. Application to VaR (easy)

An estimate of VaR is given by K such that

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

$s_N = K$, $s_0 = 0$

Fixed $\epsilon \in \mathbb{R}$ arbitrarily small, VaR estimation converges to the correct value K for each succession s_i, $i = 0, 1, 2, ...$ such that

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

3. Application to ES (difficult)

Bounds for the estimation of the generic h-th moment of the random Shortfall $E_H(S^r)$ are given by (since the positive skewness in each interval (s_i,s_{i+1}))

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

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

$s_N = L$, $L >> K$, $s_0 = K$.

The two bounds converge to the correct value if $L \to +\infty$, $s_{i+1} - s_i \to 0$ and then $N \to +\infty$ with an obvious trade-off between estimation accuracy and computation time.

3. Application to VaR and ES: standard scenario

Pareto marginals

$F_i(x) = 1 - (1+x)^{-\theta_i}$, $x \geq 0$, $\theta \geq 0$.
To have $E[X] = \frac{1}{\theta - 1}$ it musts be $\theta \geq 1$

Clayton copula

$C^{\delta}_{\mathrm{Cl}}(u_1, ..., u_d) = (u_1^{-\delta} + ... + u_d^{-\delta} - d + 1)^{-\frac{1}{\delta}}$
$u_i \in [0, 1]$, $i = 1, ..., d$ with $\delta \in (0, \infty)$
if $\delta \to 0$ then copula tends to independence,
if $\delta \to \infty$ then copula tends to comonotonicity

Model parameters

$\delta = 5$, $d = 2$

Contract parameter

$K = 20$

Algorithm parameters

$s_{i+1} = s_i + \gamma$
$n = 4$, $\alpha = \frac{1}{\sqrt{2}}$

3. Application to VaR and ES: numerical results

Case $\theta = 2$ (then $E[X] = 1$)

	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

Case $\theta = 1.2$ (then $E[X] = 5$)

	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

3. Application to VaR and ES: numerical results

With $\theta = 2$ $VaR(99\%) = 15.96$
It is $V_H\left[S(\mathbf{0}, 15.96)\right] = 0.990001$

With $\theta = 1.2$ $VaR(99\%) = 90.73$
It is $V_H\left[S(\mathbf{0}, 90.73)\right] = 0.9900005$

Obs.
The algorithm is stopped at a relatively small value.

4. The convergence speed problem

Considering the case $\theta = 1.2$, the estimation for Expected Shortfall is not stable even if $L = 10000$.
How high has to be L to reach "stability"?
With $L = 20000$, $E_{med} = 4.176$.
With $L = 40000$, $E_{med} = 4.392$.
And ... what about $E_{med} = 4.821$ with $L = 200000$???

The level of threshold L at which the algorithm will be stopped seems to be absolutely crucial for the quality of estimation.

4. The convergence speed problem

Fixed ϵ the parameter that affects the convergence speed is n.
Obs. Given d, if n increases of 1 then the number of hypercubes to be considered in the estimation is d times higher.

An alternative to speed up the algorithm is to randomize the hypercubes considered in the estimation.
Let $\lambda \in [0, 1]$ the share of the hypercubes considered in the estimation.
Obs. If n decreases of 1, then the same reduction of time is obtained considering $\lambda = 1/d$.

4. The convergence speed problem: standard scenario

Pareto marginals

$F_i(x) = 1 - (1+x)^{-\theta_i}$, $x \geq 0$, $\theta \geq 0$.

To have $E[X] = \frac{1}{\theta - 1}$ it musts be $\theta \geq 1$

Clayton copula

$C^{\delta}_{\text{Cl}}(u_1, \dots, u_d) = (u_1^{-\delta} + \dots + u_d^{-\delta} - d + 1)^{-\frac{1}{\delta}}$

$u_i \in [0,1]$, $i = 1, \dots, d$ with $\delta \in (0, \infty)$

if $\delta \to 0$ then copula tends to independence,

if $\delta \to \infty$ then copula tends to comonotonicity

Model parameters

$d = 2$, (then $\alpha = \frac{1}{\sqrt{2}}$)

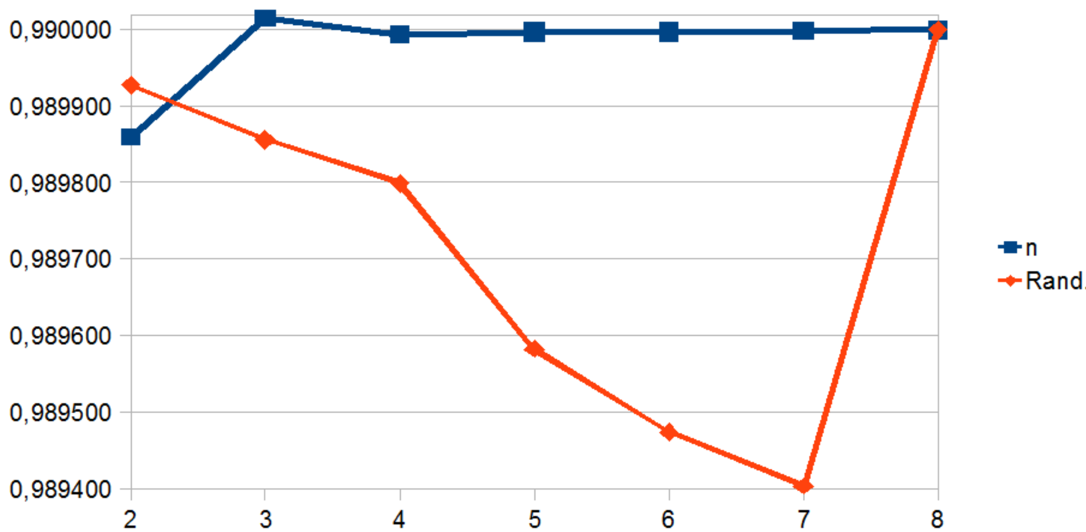
$\theta = 1.5$ (then $E[X] = 2$), $\delta = 5$.

4. The convergence speed problem: randomization

With $\epsilon = 0.000001$ the "benchmark" estimates of Var(β) for $\beta = 0.99, 0.995$, obtained with $n = 8$ are $K(0.99) = 38.33$ and $K(0.995) = 59.22$.
We use these estimates to calculate the probability at the same thresholds using different values for parameter n or using the randomization procedure.
Obs. The results obtained with the same lapse of time are reported in the same column of the following tables (e.g. the case $n = 4$ needs the same lapse of time of randomization with $\lambda = 1/16$).
Obs. We remark that in our procedure the randomization is not applied at the first step of AEP algorithm, since the "first" hypercube affects the estimation in a crucial way.

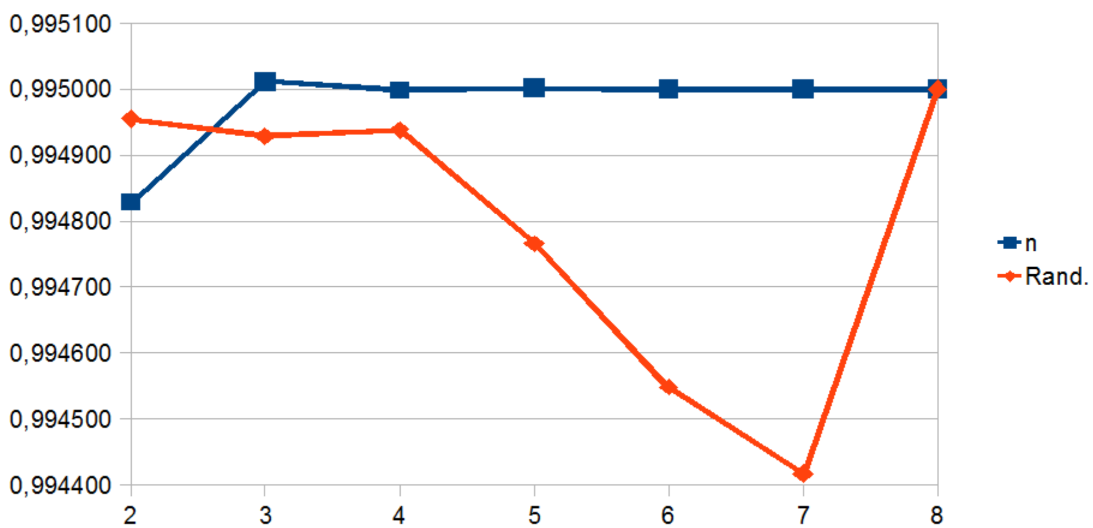
4. The convergence speed problem: numerical results

n	2	3	4	5	6	7
An.	.989860	.990015	.989994	.989997	.989997	.989998
Rd.	.989928	.989856	.989799	.989582	.989474	.989403



4. The convergence speed problem: numerical results

n	2	3	4	5	6	7
An.	.994829	.995012	.994999	.995003	.995000	.995000
Rd.	.994956	.994929	.994938	.994766	.994548	.994416



4. The convergence speed problem: comments

It is quite surprisingly that using the randomization procedure better estimates are obtained with a lower number of hypercubes!

For value of n near the "benchmark" the analytical estimation is fully satisfying, but for low levels of such parameter, the randomization seems to be a good opportunity.

5. Open problems

In order to speed-up AEP algorithm our future research seems to have two main features:
1) extension of AEP numerical procedures to higher dimensions;
2) verify the randomization procedure for higher dimensions;
3) considering some semi-random procedures instead of pure-random one considered in this work.

6. Main 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.