Model risk in credit risk

Model Risk in Credit Risk

R.FONTANA
Department of Mathematical Sciences G. Lagrange,
Politecnico di Torino.
E. LUCIANO [1]
ESOMAS Department and Collegio Carlo Alberto, Università di Torino
P. SEMERARO[2]
Department of Mathematical Sciences G. Lagrange,
Politecnico di Torino.

July 8, 2020

[1]Elisa Luciano gratefully acknowledges financial support from the Italian Ministry of Education, University and Research (MIUR), "Dipartimenti di Eccellenza" grant 2018-2022.

[2]Roberto Fontana and Patrizia Semeraro gratefully acknowledge financial support from the Italian Ministry of Education, University and Research (MIUR), "Dipartimenti di Eccellenza" grant 2018-2022.

Abstract

We provide sharp analytical upper and lower bounds for Value-at-Risk and sharp bounds for Expected Shortfall of portfolios of any dimension subject to default risk. To do so, the main methodological contribution of the paper consists in analytically finding the convex hull generators for the class of exchangeable Bernoulli variables with given mean and for the class of exchangeable Bernoulli variables with given mean and correlation in any dimension. Using these analytical results we first describe all possible dependence structures for default, in the class of finite sequences of exchangeable Bernoulli random variables. We then measure how model risk affects Value-at-Risk and Expected Shortfall.

keywords: Exchangeable Bernoulli distribution; risk measures; model risk, credit risk, default risk.

1 Introduction

Models for default risk are prone to so-called model risk because of the difficulty of describing the causes of default or even of enumerating them [1]. The issue is particularly relevant when considering the joint default of specific obligors or particular categories of obligors, because beyond the model risk for marginal defaults there is also model risk in their joint distribution. We focus on joint modeling.

The issue of model risk in default modeling has been known ever since credit risk appears in the academic literature. Professionals are well aware of its importance. The first purpose of the current paper is to measure model risk by providing the range of portfolio losses across all possible dependence structures for defaults, in the class of exchangeable Bernoulli random variables. To this end, we use two popular risk measures, Value-at-Risk (VaR) and Expected Shortfall (ES).

Univariate models of default belong to two families: structural and reduced-form models. The structural models, initiated by [2], reconduct default to the fact that the so-called asset value of a firm goes below a given monetary threshold. Reduced-form models, on which seminal work is due to [3], estimate the intensity of default from interest rates on defaultable debt. Intensity of default is then interpreted as a fixed parameter or a stochastic process. For a survey of the approaches see for instance [4]. Multivariate models either make use of a copula to aggregate univariate default probabilities (see for instance [5] or [6]) or use a Bernoulli mixture model (see Chapter 8 in [7]).

The difficulties in choosing a model for univariate modeling and calibrating it have been shown to be considerable. For structural models, the asset value is unobservable. For reduced-form models, rates of return on bonds are thought to include also a liquidity spread, which is difficult to separate from the default spread.

The difficulties in choosing or calibrating a multivariate model are even greater (see the early recognition in [8]). Structural models can be calibrated, provided that the correlation matrix of asset values can be calibrated. Multivariate reduced-form models are usually calibrated using the corresponding structural dependence (see Chapter 10 in [6]).

The previous literature which assesses model risk in joint default usually takes as given the marginal probabilities of default, as we do: marginal default indicators are Bernoulli variables. Existing research tries to explore the range of joint default probabilities, or the possible distribution of the loss from credit risk, which is the weighted sum of the marginal Bernoulli variables, where the weights are the exposures of the creditor towards different obligors. To do that, the literature uses different copulas (see [9]). Here we use the fact that all joint distributions or distributions of sums are generated

[1]For a discussion of the conceptual basis for model risk, with particular attention to derivative pricing, see [1]

starting from a finite number of so-called ray densities. Differently from copulas, all the rays can be found, either numerically or analytically.

The main contribution of the current paper consists in analytically finding the convex hull generators for the class of exchangeable Bernoulli variables with given mean and for the class of exchangeable Bernoulli variables with given mean and correlation. The analytical solution allows us to work in any dimension. The second contribution consists in describing all the joint distributions of defaults, even for large portfolios, and/or the possible distributions of the loss. This is because the multivariate Bernoulli variables represent the default indicators of a portfolio of obligors, through the ray densities, that we can find analytically. Because we can represent all the possible distributions of the loss, we can compute the bounds for two synthetic risk measures used in the finance literature, Value-at-Risk (VaR) and Expected Shortfall (ES). The paper then provides a third contribution: we show that the sharp - or attainable - VaR bounds are reached on the ray densities and we find an analytical expression for them. We also explicitly find bounds for the ES. We then measure the consequence of using a specific model looking at the range of the possible VaR and ES. Thus, the mathematical contribution consists of the analytical description of the ray densities in high dimensions. We also contribute to mathematical finance by measuring the model risk using the multivariate distributions that incorporate all possible dependence structure for defaults. They obtain as linear convex combinations of generators that can be found analytically. This solution allows us to find bounds to measure model risk.

The paper proceeds as follows: Section 2 reviews the background literature. Section 3 introduces the mathematical framework. The notion and properties of rays for exchangeable Bernoulli variables are described in Section 4. Section 5 introduces the risk measures and provides analytical bounds for exchangeable Bernoulli variables. Model risk is discussed in Section 6. Sections 6.1 and 6.2 provide calibrated examples.

2 Background literature

The paper has two strands of background literature: a purely mathematical one and a mathematical finance one.

The main reference for the mathematical approach is [10], which develops a simple method to represent all the Bernoulli variables with some specified moments as a convex hull of densities belonging to the same class, the ray densities. Rays and extremal rays, which generate all the convex hull, appear in the theory of convex polyhedra. References for the theory of convex polyhedra are, e.g., [11] and [12].

The paper [10] provides an algorithm for finding the extreme rays of a given class of multivariate Bernoulli variables without restrictions either on the number of variables or on the specified moments. The only drawback of the method is the amount of computational effort required for the numerical solution. In the current paper we show that the exchangeability assumption allows us to find the extreme rays in an analytical way. That is crucial for the mathematical-finance application, in which the number of Bernoulli variables - representing the credit obligors - is usually in the order of several hundreds.

The background literature in mathematical finance produces risk-measures, e.g. VaR and their bounds, both in the absence and in the presence of constraints on the variance of the sum of the losses. The latter constraint is empirically relevant, since realizations of the aggregate loss from a portfolio of credits are often available, and its variance - i.e. the variance of the sum of the losses - can be estimated (see [13]). Actually, as long as the additional constraint improves the bounds, use of the estimated variance is welcome. In Section 5, we rephrase the variance constraint in terms of a fixed value of the equicorrelation of the Bernoulli variables, and provide our VaR bounds both in the absence and in the presence of that bound (see [13]).

Early results on VaR bounds were obtained from the Fréchet bounds, for given marginal distributions, in the absence of variance constraints on the sum and for the case in which the marginal distributions are more than two. This case is interesting for VaR applications and it is our reference in the current paper. The bounds were not sharp (see for instance [14]). Using duality theory, [15] provided the sharp lower VaR bound for the case in which the marginal random variables are more than two, identical, continuous and satisfy a monotonicity - namely non-decreasing or non-increasing - property. The bound is analytical. An extension for the case of identical, continuous margins was provided by [16]. The bounds obtained there are sharp, are derived under three different conditions - called mixing, attainment and ordering conditions - but have to be computed numerically. An effective method to compute them is provided. [17] instead extended the result of [15] to marginal distributions which are neither continuous, nor identical and monotone. The bounds are neither sharp nor analytical, but numerical results are provided for the case in which restrictions on the dependence structure of the marginal distributions - restrictions on the joint distributions of couples, triplets etc - are given. The numerical approach is shown to be highly simplified when the marginal distributions are identical. From this literature, the last result is the closest to ours, since our marginal distributions are discrete and identical. We do not consider the dependence structure of n-tuples while we assume exchangeability. In contrast with [17], our VaR bounds, provided for Bernoulli margins, are sharp and analytical. They are derived from a different mathematical approach, since we do not use duality, but rather, as mentioned above, the theory of polyhedral cones and their generation through ray densities.

A further improvement in the numerical approach of [17] is provided in [18]. This is useful for the case in which the marginal distributions are continuous, even non identical, and the number of marginal distributions is high, as in the credit risk case. The condition for applying it is that the marginal distributions can be divided into subgroups, each containing identically distributed (i.d) random variables. A fortiori the numerical approach applies to iid random variables. The algorithm provides sharp bounds, but requires continuity of the marginal distributions, so it does not apply to our set-up. A similar comment applies to [19], which uses the notion of positive dependence to improve on the VaR bounds, computed numerically. The easily-computable version of their non-sharp bounds can be adopted when the marginal distributions are identical and continuous with decreasing density and symmetric copula bounds. Because of the continuity requirement, the numerical approach they suggest cannot be adopted in the case examined in this paper. So, even if we wanted to compute the VaR bounds numerically instead of analytically, we could not adopt the approach of [18] or [19], because their assumptions on the marginal distributions and on the dependence structure are not consistent with our set-up. Their achievements are mentioned here to provide the reader with a possible comprehensive picture of the related literature on VaR bounds.

Results on VaR bounds, for given marginal distributions, in the presence of variance constraints on the sum are in [13]. The variance constraint is an upper bound on the variance value. The bounds are analytical, but are not sharp all the time. The paper provides a numerical algorithm to approximate the sharp bounds. Our bounds for this case are instead sharp and analytical.

3 Default indicators: mathematical set-up

We consider a credit portfolio P with d obligors.

Some notation is needed. Let the random variable $\boldsymbol{X} = (X_1, \dots, X_d)$ be the default indicators for the portfolio P and let us assume that the indicator $\boldsymbol{X}$ is exchangeable, i.e. $\boldsymbol{X} \in \mathcal{E}_d$, where $\mathcal{E}_d$ is the class of d-dimensional exchangeable Bernoulli distributions. Let $\mathcal{E}_d(p)$ be the class of exchangeable Bernoulli distributions with the same Bernoulli marginal distributions $B(p)$, where p is the marginal default probability of each obligor. If $\boldsymbol{X} = (X_1, \dots, X_d)$ is a random vector with joint distribution in $\mathcal{E}_d(p)$, we denote

- its cumulative distribution function by F_p and its probability mass function (pmf) by f_p;
- the column vector which contains the values of f_p over $\mathcal{X}_d := \{0,1\}^d$, by $(f_p(\boldsymbol{x}) : \boldsymbol{x} \in \mathcal{X}_d)$ respectively; we make the non-restrictive hypothesis that the set $\mathcal{X}_d$ of 2^d binary vectors is ordered according to the reverse-lexicographical criterion. For example $\mathcal{X}_2 = \{00, 10, 01, 11\}$ and $\mathcal{X}_3 = \{000, 100, 010, 110, 001, 101, 011, 111\}$;
- we denote by $\mathcal{P}_d$ the set of permutations on $\{1, \dots, d\}$;

Recall that the expected value of X_i is p, $\mathrm{E}[X_i] = p$, $i = 1, \ldots, d$. We denote $q = 1 - p$. We assume that vectors are column vectors.

3.1 Exchangeable Bernoulli variables

Let us consider a pmf $f_p \in \mathcal{E}_d(p)$. Since $f_p(\boldsymbol{x}) = f_p(\sigma(\boldsymbol{x}))$ for any $\sigma \in \mathcal{P}_d$, any mass function f_p in $\mathcal{E}_d(p)$ defines $f_i := f_p(\boldsymbol{x})$ if $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathcal{X}_d$ and $\#\{x_j : x_j = 1\} = i$. Therefore we identify a mass function f_p in $\mathcal{E}_d(p)$ with the corresponding vector $\boldsymbol{f}_p := (f_0, \ldots, f_d)$. Furthermore, by exchangeability the moments depend only on their order, we therefore use μ_α to denote a moment of order $\alpha = \mathrm{ord}(\boldsymbol{\alpha}) = \sum_{i=1}^d \alpha_i$, where $\boldsymbol{\alpha} \in \mathcal{X}_d$. For example we have $\mu_1 = p$. We also observe that the correlation ρ between two Bernoulli variables $X_i \sim B(p)$ and $X_j \sim B(p)$ is related to the second-order moment $\mu_2 = \mathrm{E}[X_i X_j]$ as follows

$$\mu_2 = \rho p q + p^2. \tag{3.1}$$

3.2 Joint defaults, loss distribution and risk measures

To model the loss of a credit risk portfolio P of d obligors we consider the sum of the percentage individual losses

$$L = \sum_{i=1}^{d} w_i X_i,$$

where $w_i \in (0, 1]$ and $\sum_{1=1}^d w_i = 1$. In this paper we consider the case $w_i = \frac{1}{d}$, $i \in \{1, \ldots, d\}$. The extension to unequal weights can be done numerically. For equal weights, $L = \frac{S_d}{d}$, where

$$S_d = \sum_{i=1}^{d} X_i$$

represents the number of defaults. Therefore, the distribution of S_d represents the distribution of the loss. Since the vector of default indicators $\boldsymbol{X}$ is assumed to be exchangeable, there is a one-to-one correspondence between the distribution of the number of defaults and the joint distribution of $\boldsymbol{X}$. In fact, as said in the preliminaries, since $f_p(\boldsymbol{x}) = f_p(\sigma(\boldsymbol{x}))$ for any $\sigma \in \mathcal{P}_d$, any mass function f_p in $\mathcal{E}_d(p)$ defines $f_i := f_p(\boldsymbol{x})$ if $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathcal{D}_d$ and $\#\{x_j : x_j = 1\} = i$. We can define a one-to-one correspondence between $\mathcal{E}_d(p)$ and the class of the distributions on the number of defaults.

Let $\mathcal{S}_d(p)$ be the class of distributions p_S on $\{0, \ldots, d\}$ such that $S_d = \sum_{i=0}^d X_i$ with $\boldsymbol{X} \in \mathcal{E}_d(p)$. Let $p_S(j) = p_j = P(S_d = j)$ and $\boldsymbol{p}_S = (p_0, \ldots, p_d)$.

The map:

$$\begin{aligned} E : \mathcal{E}_d(p) &\to \mathcal{S}_d(p) \\ f_j \to p_j &= \binom{d}{j} f_j. \end{aligned} \tag{3.2}$$

is a one-to-one correspondence between $\mathcal{E}_d(p)$ and $\mathcal{S}_d(p)$. Therefore we have

$$\mathcal{E}_d(p) \leftrightarrow \mathcal{S}_d(p) \tag{3.3}$$

We now prove that the class of distributions $\mathcal{S}_d(p)$ coincides with the entire class of discrete distributions with mean dp, say $\mathcal{D}_d(dp)$. This fact is useful to simplify the search of the generators of $\mathcal{E}_d(p)$. The class $\mathcal{D}_d(dp)$ is not of special interest in this context, but it is introduced for technical reasons.

Proposition 3.1. *It holds $\mathcal{S}_d(p) = \mathcal{D}_d(dp)$.*

Proof. 1) $\mathcal{S}_d(p) \subseteq \mathcal{D}_d(dp)$. This is trivial.

2) $\mathcal{D}_d(dp) \subseteq \mathcal{S}_d(p)$. Let $\{p_0, \ldots, p_d\} \in \mathcal{D}_d(dp)$. Let us define $f_i = \frac{p_i}{\binom{d}{j}}$ and $p(x_1, \ldots, x_d) = f_i$ for all $(x_1, \ldots, x_d) \in \mathcal{X}_d$ such that $\sum_{j=1}^{d} x_j = i$. The mass function p is the mass function of a d-dimensional Bernoulli random vector, which is exchangeabe by contruction. By exchangeability $E[X_1] = \ldots = E[X_d]$. We have

$$\begin{aligned} E[X_1] = P(X_1 = 1) &= \sum_{(x_1,\ldots,x_d):x_1=1} p(x_1, \ldots, x_d) = \sum_{i=1}^{d} \sum_{\substack{(x_1,\ldots,x_d):x_1=1, \\ \sum_{i=1}^{d} x_i = i}} p(x_1, \ldots, x_d) \\ &= \sum_{i=1}^{d} \sum_{\substack{(x_1,\ldots,x_d):x_1=i, \\ \sum_{i=1}^{d} x_i = 1}} f_i = \sum_{i=1}^{d} \binom{d-1}{i-1} \frac{p_i}{\binom{d}{i}} \\ &= \sum_{i=1}^{d} \frac{(d-1)!}{(i-1)!(d-1-i+1)!} \frac{i!(d-i)!}{d!} p_i \\ &= \sum_{i=1}^{d} \frac{i}{d} p_i = \frac{1}{d} pd = p. \end{aligned}$$

Then $\boldsymbol{X} \in \mathcal{E}_d(p)$.

Now let $S_d := \sum_{i=1}^{d} X_i$. We have $P(S_d = j) = \binom{d}{j} f_j = p_j$ and $\{p_0, \ldots, , p_d\} \in \mathcal{S}_d(p)$. $\square$

Therefore the three classes $\mathcal{E}_d(p)$, $\mathcal{S}_d(p)$ and $\mathcal{D}_d(dp)$ are essentially the same class, i.e.

$$\mathcal{E}_d(p) \leftrightarrow \mathcal{S}_d(p) \equiv \mathcal{D}_d(dp) \tag{3.4}$$

Thanks to the above proposition to find the generators of $\mathcal{S}_d(p)$ we can look for the generators of $\mathcal{D}_d(dp)$. This simplifies the search. The generators we find are in one-to-one relationship with the generators of $\mathcal{E}_d(p)$.

4 Exchangeable Bernoulli generators

We build on the results in [10], where the authors represent the Fréchet class of multivariate d-dimensional Bernoulli distributions with given margins and/or pre-specified moments as the points of a convex hull. The generators of the convex hull are mass functions in the class and they can be explicitly found. We show here that under the condition of exchangeability we analytically find the ray densities. We focus on two classes: the class $\mathcal{E}_d(p)$ and the class $\mathcal{E}_d(p, \rho)$, i.e. the class of exchangeable Bernoulli vectors with given p and given correlation ρ. The one-to-one correspondence E between the distributions $\boldsymbol{p}_S \in \mathcal{S}_d(p)$ and $\boldsymbol{f}_p \in \mathcal{E}_d(p)$ is also a one-to-one correspondence between the distributions $\boldsymbol{p}_S \in \mathcal{S}_d(p, \rho)$ and $\boldsymbol{f}_p \in \mathcal{E}_d(p, \rho)$.

In Section 3.1 we represent the class $\mathcal{E}_d(p)$ as a convex hull of mass functions in the class, which we call ray densities, so that each mass function is a convex combination of ray densities belonging to $\mathcal{E}_d(p)$. We analytically find the ray densities and their number, that depends on the dimension d and the mean value p. The one-to-one map between $\mathcal{E}_d(p)$ and $\mathcal{S}_d(p)$ and Proposition 3.1 are crucial.

In Section 3.2 we represent the class $\mathcal{E}_d(p, \rho)$, as well as $\mathcal{S}_d(p, \rho)$, as a convex hull of ray densities. We analytically find them using the one-to-one correspondence between the class $\mathcal{E}_d(p)$ and the class $\mathcal{S}_d(p)$ and between the relative subclasses $\mathcal{S}_d(p, \rho)$ and $\mathcal{E}_d(p, \rho)$. We prove that the ray densities in $\mathcal{S}_d(p, \rho)$ have support on at most three points. By so doing, also in this case the dimension d is not an issue.

4.1 Generators for given marginal default probabilities

Using the equivalence $\mathcal{S}_d(p) \equiv \mathcal{D}_d(pd)$ stated in Proposition 3.1 a pmf in $\mathcal{S}_d(p)$ is a pmf on $\{0, \dots, d\}$ with mean pd. Thanks to the map E in Equation (3.4) this is also equivalent to finding a set of conditions that a pmf of a multivariate Bernoulli has to satisfy for being in $\mathcal{E}_d(p)$. This fact is crucial in the following proposition.

Proposition 4.1. *Let $\boldsymbol{Y}$ be a discrete random variable defined over $\{0, \dots, d\}$ and let p_Y be its pmf. Then*

$$Y \in \mathcal{S}_d(p) \iff \sum_{j=0}^{d} (j - pd) p_Y(j) = 0.$$

Proof. Let $\boldsymbol{Y}$ be a discrete random variable defined over $\{0, \dots, d\}$. By Proposition 3.1 $Y \in \mathcal{S}_d(p)$ iff $E[Y] = pd$. It holds

$$E[Y] = pd \Longleftrightarrow E[Y - pd] = 0 \Longleftrightarrow \sum_{j=0}^{d} (j - pd) p_Y(j) = 0.$$

$\square$

Using Proposition 4.1 we can find all generators of $\mathcal{S}_d(p)$. Thanks to the map E, that is equivalent to finding all the generators of $\mathcal{E}_d(p)$.

We have to find the solutions $\boldsymbol{p}_S = (p_0, \dots, p_d)$ of

$$\sum_{j=0}^{d} (j - pd) p_j = 0. \tag{4.1}$$

with the conditions $p_j \geq 0$, $j = 0, \dots, d$ and $\sum_{j=0}^{d} p_j = 1$. From the standard theory of linear equations we know that all the positive solutions of (4.1) are elements of the convex cone

$$\mathcal{C}_p = \{\boldsymbol{z} \in \mathbb{R}^{d+1} : \sum_{j=0}^{d} a_j z_j = 0, \, \boldsymbol{I}\boldsymbol{z} \geq 0\}, \tag{4.2}$$

where $a_j = j - pd$ and I is the $(d+1) \times (d+1)$ identity matrix. We recall that a subset $\mathcal{C}$ of $\mathbb{R}^d$ is called a cone if $0 \in \mathcal{C}$ and $\boldsymbol{x} \in \mathcal{C}$ implies $\lambda \boldsymbol{x} \in \mathcal{C}$ for every non-negative real scalar λ. The particular cones consisting of a non-zero vector $\boldsymbol{x}$ and all its multiples $\lambda \boldsymbol{x}, \lambda \geq 0$ are rays. A cone which contains at least one non-zero vector is therefore just the union of the rays it contains.

Definition 4.1. *A ray $\boldsymbol{x}$ of a convex cone $\mathcal{C}$ is an extreme ray of $\mathcal{C}$ if $\boldsymbol{x}$ is not a positive linear combination of two linearly independent vectors of $\mathcal{C}$.*

Classical result in the theory of convex polyhedra is that every convex cone can be expressed as a convex combination of extreme rays (see [11] and [12]). Therefore the solutions of (4.2) can be generated as convex combinations of a set of generators which are referred to as extremal rays of the linear system. A first step to analytically find the extremal rays is to find their support. In this particular case we can deduce that the support has at most two points because the binding constraints are $d + 2$, i.e. $p_j \geq 0$, $j = 0, \cdots, d$ and equation (4.1). However, in the following proposition we consider the support of the extremal rays of a cone defined by a general m-dimensional homogeneous linear system. We give the general result because it applies also to the case, discussed in the next section, of given default correlations and can also be applied to find ray densities for any set of given moments.

Proposition 4.2. *Let us consider the linear system*

$$A\boldsymbol{z} = 0, \quad \boldsymbol{z} \in \mathbb{R}_+^{d+1} \tag{4.3}$$

where A *is a* $m \times (d+1)$ *matrix,* $m \leq d$ *and* $\mathrm{rank}(A) = m$. *The extremal rays of the system* (4.3) *have at most* $m+1$ *non-zero components.*

Proof. Let $\mathcal{C}_A = \{\boldsymbol{z} \in \mathbb{R}^{d+1} : A\boldsymbol{z} = 0, I\boldsymbol{z} \geq 0\}$ be the convex cone of all the positive solutions of (4.3). From Lemma 2.3 in [20] it follows that a solution $\boldsymbol{r}$ of (4.1) is an extremal ray of $\mathcal{C}_A$ iff $I^*\boldsymbol{z} = 0$ for a submatrix $n_{I^*} \times (d+1)$, I^* of I and

$$\mathrm{rank} \begin{bmatrix} A \\ I^* \end{bmatrix} = d.$$

Therefore it must be $\mathrm{rank}(I^*) \geq d-m$ and then $\boldsymbol{r}$ has at most $(d+1)-(d-m) = m+1$ non-zero components. □

Corollary 4.1. *The extremal rays of the convex cone* $\mathcal{C}_p$ *in* (4.2) *have at most two non-zero components.*

Proof. Let $a_j = j - pd$, $j = 0, \ldots, d$. The matrix $A = [a_0, \ldots, a_d]$ is the row vector of the coefficients. Since $\mathrm{rank}\, A = m = 1$ then an extremal ray $\boldsymbol{r}$ has at most two non-zero components. □

Proposition 4.3. *The extremal rays of the convex cone* $\mathcal{C}_p$ *in* (4.2) *are*

$$p_{j_1,j_2}(y) = \begin{cases} \frac{j_2-pd}{j_2-j_1} & y = j_1 \\ \frac{pd-j_1}{j_2-j_1} & y = j_2 \\ 0 & \textit{otherwise} \end{cases}, \tag{4.4}$$

with $j_1 = 0, 1, \ldots, j_1^M$, $j_2 = j_2^m, j_2^m + 1, \ldots, d$, j_1^M *is the largest integer less than* pd *and* j_2^m *is the smallest integer greater than* pd.

If pd *is integer the extremal rays contain also*

$$p_{pd}(y) = \begin{cases} 1 & y = pd \\ 0 & \textit{otherwise} \end{cases}. \tag{4.5}$$

Proof. Let $a_j = j - pd$ and consider the case pd not integer. Equation (4.1) becomes

$$\sum_{j=0}^{d} a_j p_j = 0.$$

By Corollary 4.1 the extremal rays have at most two non zero components, say j_1, j_2. Therefore the extremal rays can be found considering the equations

$$a_{j_1}p_{j_1} + a_{j_2}p_{j_2} = 0,$$

where we make the non restrictive assumption $j_1 < j_2$. The equation (4.1) has positive solutions only if $a_{j_1}a_{j_2} < 0$. We observe that $a_{j_1} < 0$ for $0 \leq j_1 \leq j_1^M$ and $a_{j_2} > 0$ for $j_2^m \leq j_2 \leq d$. In this case we have $j_2^m = j_1^M + 1$. It follows that for $0 \leq j_1 \leq j_1^M$ and $j_2^m \leq j_2 \leq d$ we have $a_{j_1}a_{j_2} < 0$. A positive solution of Equation (4.1) is

$$\begin{cases} \tilde{p}_y(j_1) = x_{j_1} = j_2 - pd \\ \tilde{p}_y(j_2) = -x_{j_2} = pd - j_1 \end{cases} .$$

We have $\tilde{p}_y(j_1) + \tilde{p}_y(j_2) = j_2 - pd + pd - j_1 = j_2 - j_1$ and then the normalized extremal rays corresponding to j_1 and j_2 are given by (4.4). If pd is integer we have $a_{pd} = 0$. It follows that (4.5) is also an extremal solution.

□

We denote by $R_{(j_1,j_2)}$ and R_{pd} the random variables whose pmf are $\boldsymbol{r}_{(j_1,j_2)}$ and $\boldsymbol{r}_{pf}$ respectively. We will refer to $\boldsymbol{r}_{(j_1,j_2)}$ and $\boldsymbol{r}_{pf}$ as extremal ray densities and $R_{(j_1,j_2)}$ and R_{pd} as extremal ray random variables. We will omit extrema for the sake of simplicity. Notice that $\boldsymbol{r}_{(0,d)} = (1-p, 0, \ldots, 0, p)$.

Corollary 4.2. *If pd is not integer there are* $n_p = (j_1^M + 1)(d - j_1^M)$ *ray densities.*

If pd is integer there are $n_p = d^2p(1-p) + 1$ *ray densities.*

Note that there are roughly $d^2p(1-p)$ extreme rays for large d. This is numerically manageable in mathematical-finance applications, where large investment banks typically deal with several hundreds of default indicators. For example with $d = 1000$ and $p = 0.1$ we obtain 90000 extreme rays. Nevertheless, we will provide analytical expressions for VaR and this avoids numerical search, and makes any dimension d possible.

We have proved the following.

Theorem 4.1. *The following holds.* $S_d \in \mathcal{S}_d(p)$ *iff there exist* $\lambda_1, \ldots, \lambda_{n_p} \geq 0$ *summing up to 1 such that*

$$\boldsymbol{p}_S = \sum_{i=1}^{n_p} \lambda_i \boldsymbol{r}_i,$$

where $\boldsymbol{r}_i$ *are the ray densities of* $\mathcal{C}_p$ *as defined in* (4.2) *and* n_p *is their number.*

4.1.1 Second order moments

Let $\boldsymbol{X} \in \mathcal{E}_d(p)$ and let $\mu_2 = E[X_iX_j]$ its second order cross moment.

Proposition 4.4. *Let $\boldsymbol{X} \in \mathcal{E}_d(p)$. It holds*

$$\mu_2 = \sum_{k=0}^{d} \frac{k(k-1)}{d(d-1)} p_k. \tag{4.6}$$

Proof. By exchangeability we can fix any pair $i, j \in \{1, \dots, d\}$. It holds

$$\begin{aligned}\mu_2 = P(X_i = 1, X_j = 1) &= \sum_{k=0}^{d} P(X_i = 1, X_j = 1 | S_d = k) P(S_d = k) \\ &= \sum_{k=2}^{d} \frac{\binom{d-2}{k-2}}{\binom{d}{k}} p_k = \sum_{k=2}^{d} \frac{k(k-1)}{d(d-1)} p_k \\ &= \sum_{k=0}^{d} \frac{k(k-1)}{d(d-1)} p_k,\end{aligned}$$

□

We straightforward obtain the formula for the moment of order α,

$$\mu_\alpha = \sum_{k=\alpha}^{d} \frac{\binom{d-\alpha}{k-\alpha}}{\binom{d}{k}} p_k.$$

Thanks to the one-to-one map E we can find the bounds for the second order moments of $\mathcal{E}_d(p)$ using the second order moments of S_d. We have

$$E[S_d^2] = E[(X_1 + \ldots + X_d)^2] = pd + d(d-1)\mu_2. \tag{4.7}$$

Proposition 4.5. *Let $\boldsymbol{X} \in \mathcal{E}_d(p)$. Then if pd is not integer*

$$\frac{1}{d(d-1)} [-j_1^M (j_1^M + 1) + 2 j_1^M pd] \leq \mu_2 \leq p.$$

If pd is integer

$$\frac{p(pd-1)}{(d-1)} \leq \mu_2 \leq p. \tag{4.8}$$

Proof. From (4.7) we have $\mu_2 = \frac{1}{d(d-1)}[E[S_d^2] - pd]$. Since $S_d \in \mathcal{S}_d(p)$, its density is a convex linear combinations of the ray densities. In [10] it is proved that the moments of S_d are linear combinations of the moments of the ray variables. We obtain

$$E[R^2_{(j_1, j_2)}] = j_1^2 \frac{j_2 - pd}{j_2 - j_1} + j_2^2 \frac{pd - j_1}{j_2 - j_1} = -j_1 j_2 + (j_1 + j_2) pd, \tag{4.9}$$

and if pd is integer

$$E[R^2_{pd}] = (pd)^2. \tag{4.10}$$

To maximize μ_2 we have to maximize $E[S_d^2]$. From (4.9) and (4.10) we easily get that the ray variable for which the second order moment is maximum is $R_{(0,d)}$ and we have $E[R_{(0,d)}^2] = pd^2$. Then, after some computations, the maximum second moment is $\mu_2^M = p$.

To minimize μ_2 we have to minimize $E[S_d^2]$. We consider two cases. If pd is not integer, from (4.9) we have that the ray variable for which the second order moment is minimum is $R_{(j_1^M, j_2^m)} = R_{(j_1^M, j_1^M+1)}$, for which we have $E[R^2_{(j_1^M, j_1^M+1)}] = -j_1^M(j_1^M+1) + (2j_1^M+1)pd$ and the assert follows. If pd is integer the ray variable for which the second order moment is minimum is R_{pd}. Since $E[R_{pd}^2] = (pd)^2$, (4.8) follows.

$\square$

Notice that the lower bound for μ_2 goes to p^2 for large d, which is a likely practical case in the credit risk framework. In fact $\frac{p(pd-1)}{(d-1)} \to p^2$ for $d \to \infty$. The case $\mu_2 = p^2$ corresponds to incorrelation, thus for d large we essentially have positive correlation. Thanks to equation (3.1), the next corollary to the above proposition provides bounds for the correlation coefficient.

Corollary 4.3. *Let $\boldsymbol{X} \in \mathcal{E}_d(p)$. Then if pd is not integer*

$$\frac{\frac{1}{d(d-1)}[-j_1^M(j_1^M+1) + 2j_1^M pd] - p^2}{p(1-p)} \leq \rho \leq 1.$$

If pd is integer

$$-\frac{1}{d-1} \leq \rho \leq 1. \tag{4.11}$$

Remark 1. *The above corollary improves previous results (see e.g. [21]) for the lower bound of correlation for non-integer pd. It is indeed known that for a finite sequence of exchangeable random variables $\rho \geq -\frac{1}{d-1}$, therefore we show that*

$$\rho_{min} = \frac{\frac{1}{d(d-1)}[-j_1^M(j_1^M+1) + 2j_1^M pd] - p^2}{p(1-p)} \geq -\frac{1}{d-1}$$

. Since j_1^M is the largest integer less than pd, we have $j_1^M = pd - \epsilon$, with $\epsilon \in (0,1)$. It holds:

$$\begin{aligned}
\rho_{min} &= \frac{\frac{1}{d(d-1)}[-j_1^M(j_1^M+1) + 2j_1^M pd] - p^2}{p(1-p)} \\
&= \frac{\frac{1}{d(d-1)}[-(pd-\epsilon)(pd-\epsilon+1) + 2(pd-\epsilon)pd] - p^2}{p(1-p)} \\
&= \frac{\frac{1}{d(d-1)}[p^2d^2 - pd + \epsilon - \epsilon^2] - p^2}{p(1-p)} = \frac{-pd + p^2 d + \epsilon - \epsilon^2}{d(d-1)} \frac{1}{p(1-p)} \\
&= -\frac{1}{d-1} + \frac{\epsilon - \epsilon^2}{d(d-1)p(1-p)} \geq -\frac{1}{d-1},
\end{aligned}$$

in fact $0 < \epsilon \leq 1$ *implies* $\epsilon - \epsilon^2 < 1$. *We notice that, if* pd *is integer, then* $j_1^M = pd - 1$, *i.e.* $\epsilon = 1$ *and we find the lower bound in* (4.11).

4.2 Generators for given marginal default probabilities and default correlations

In this section we consider the class of multivariate exchangeable Bernoulli mass functions with given margins p and given correlation ρ, i.e. the class $\mathcal{E}_d(p, \rho)$. We now find the generators of $\mathcal{S}_d(p, \rho)$.

Since $S_d \in \mathcal{S}_d(p, \rho)$ iff $E[S_d] = pd$ and $E[S_d^2] = dp + d(d-1)\mu_2$ (see equation (4.7)), we can define an homogeneous linear system whose solutions are the pmf in $\mathcal{S}_d(p, \rho)$. Similarly to the procedure which we adopted to obtain Theorem 4.1 we can prove the following result.

Theorem 4.2. $S_d \in \mathcal{S}_d(p, \rho)$ *iff there exist* $\lambda_1, \ldots, \lambda_{n_p} \geq 0$ *summing up to 1 such that*

$$\boldsymbol{p}_S = \sum_{i=1}^{n_p} \lambda_i \boldsymbol{r}_{\rho,i},$$

where $\boldsymbol{r}_{\rho,i}$ *are the normalized extremal rays of the cone* $\mathcal{C}_{p,\rho}$ *defined by the linear system:*

$$\begin{cases} \sum_{j=0}^{d} [j - pd] p_j = 0 \\ \sum_{j=0}^{d} [j^2 - (pd + d(d-1)\mu_2)] p_j = 0. \end{cases} \tag{4.12}$$

The following corollary of Proposition 4.2 characterizes the ray densities of $\mathcal{S}_d(p, \rho)$.

Corollary 4.4. *The extremal rays of* $\mathcal{S}_d(p, \rho)$ *have support on at most three points.*

Proposition 4.6. *Let* $\alpha_j := j - pd$ *and* $\beta_j := j^2 - (pd + d(d-1)\mu_2)$ *and let* $A_{ij} = \det \begin{bmatrix} \alpha_i & \alpha_j \\ \beta_i & \beta_j \end{bmatrix}$. *If*

$$\begin{cases} A_{jk} \geq 0 \\ A_{ik} \leq 0 \\ A_{ij} \geq 0. \end{cases} \tag{4.13}$$

the extremal rays of (4.1) *are* $\boldsymbol{r}_\rho = (p_0, \ldots, p_d)$, *where* $p_l = 0, l \neq i, j, k$,

$$\begin{aligned} p_i &= \frac{jk - (j + k - 1)dp + d(d-1)\mu_2}{(k-i)(j-i)} \\ p_j &= -\frac{ik - (i + k - 1)dp + d(d-1)\mu_2}{(k-j)(j-i)} \\ p_k &= \frac{ij - (i + j - 1)dp + d(d-1)\mu_2}{(k-j)(k-i)}, \end{aligned} \tag{4.14}$$

with $i < j < k$.

Proof. The extremal rays of (4.1) can be found as follows. Let $\alpha_j := j - pd$ and $\beta_j := j^2 - (pd + d(d-1)\mu_2)$, we can write system (4.12) as follows:

$$\begin{cases} \sum_{j=0}^{d} \alpha_j p_j = 0 \\ \sum_{j=0}^{d} \beta_j p_j = 0, \end{cases} \tag{4.15}$$

Let now $A = \begin{bmatrix} \alpha_0 & \dots & \alpha_d \\ \beta_0 & \dots & \beta_d \end{bmatrix}$. From Corollary 4.4 we have to find the positive solutions (x_j, x_j, x_k), for $i < j < k$, of

$$\begin{cases} \alpha_i x_i + \alpha_j x_j + \alpha_k x_k = 0 \\ \beta_i x_i + \beta_j x_j + \beta_k x_k = 0, \end{cases} \tag{4.16}$$

Then, from a positive solution, we find $p_l = \frac{x_l}{x_i + x_j + x_k}$, $l \in \{i, j, k\}$. Letting $x_k = 1$ the system 4.16 becomes

$$\begin{cases} \alpha_i x_i + \alpha_j x_j = -\alpha_k \\ \beta_i x_i + \beta_j x_j = -\beta_k, \end{cases}$$

and its solution can be determined by standard computation using Cramer's formula. We have

$$\begin{cases} x_i = \frac{A_{jk}}{A_{ij}} \\ x_j = -\frac{A_{ik}}{A_{ij}} \\ x_k = 1. \end{cases} \tag{4.17}$$

It immediately follows that the positive solutions can be obtained if conditions (4.13) hold. By standard computation (4.14) follows from (4.17). We observe that if $rank(A) = 1$ then the extremal rays have support on two or one point. $\square$

Corollary 4.5. *Conditions* (4.13) *are equivalent to*

$$\begin{cases} \mu_2 \geq \max\{\mu_2^{i,j}, \mu_2^{j,k}\} \\ \mu_2 \leq \mu_2^{i,k} \end{cases} \tag{4.18}$$

where $\mu_2^{j,k}$ is the second order moment of the extreme ray density $\boldsymbol{r}_{(j,k)}$.

Proof. The expression for second order moment of the ray density $\boldsymbol{r}_{(j,k)}$ is

$$d(d-1)\mu_2^{jk} = (j + k - 1)dp - jk. \tag{4.19}$$

As a consequence $A_{ij} = d(d-1)(j-1)[\mu_2 - \mu_2^{j,k}]$ and conditions (4.18) follow. $\square$

We conclude this section with the following proposition that gives necessary and sufficient conditions for a ray density in $\mathcal{S}_d(p)$ to be also a ray density in $\mathcal{S}_d(p, \rho)$.

Proposition 4.7. *A ray density $\boldsymbol{r} \in \mathcal{S}_d(p, \rho)$ has support on two points iff it is a ray density in $\mathcal{S}_d(p)$ and $\mu_2^{\boldsymbol{r}} = \mu_2$, where $\mu_2^{\boldsymbol{r}}$ is the second order cross moment of $\boldsymbol{r}$.*

Proof. If $\boldsymbol{r}$ is a solution of (4.12) it is also a solution of (4.1) and since it has support on two points by assumption, it is an extremal solution. Thus $\boldsymbol{r} \in \mathcal{S}_d(p)$ is a ray density. Viceversa if $\boldsymbol{r} \in \mathcal{S}_d(p)$ it satisfies the first equation of (4.12) by definition and if $\mu_2^{\boldsymbol{r}} = \mu_2$ it satisfies the second equation by construction. Since it has mass on two points it is an extreme solution of (4.12). □

5 Financial risk measures and their bounds

As measures of portfolio risk we consider the Value-at-Risk (VaR) and the Expected Shortfall (ES) of S_d. We recall their definition for a general random variable Y. See [22] for the definition of Expected Shortfall for discrete variables.

Definition 5.1. *Let Y be a random variable representing a loss with finite mean. Then the VaR_α at level α is defined by*

$$VaR_\alpha(Y) = \inf\{y \in \mathbb{R} : P(Y \leq y) \geq \alpha\}$$

and the expected shortfall at level α is defined by

$$ES_\alpha(Y) = \frac{1}{1-\alpha}(E[Y; Y \geq VaR_\alpha(Y)] + VaR_\alpha(Y)(1 - \alpha - P(Y \geq VaR_\alpha(Y)))).$$

The following proposition provides the bounds for the VaR_α and ES_α of S_d, for S_d in a given class.

Proposition 5.1. *1. Let $S_d \in \mathcal{S}_d(p)[\mathcal{S}_d(p,\rho)]$ and let $VaR_\alpha(S_d)$ be its value at risk. Then*

$$\min_R VaR_\alpha(R) \leq VaR_\alpha(S_d) \leq \max_R VaR_\alpha(R),$$

where R are the ray densities of $\mathcal{S}_d(p)[\mathcal{S}_d(p,\rho)]$.

2. Let $S_d \in \mathcal{S}_d(p)[\mathcal{S}_d(p,\rho)]$ and let $ES_\alpha(S_d)$ be its expected shortfall. Then

$$\min_R VaR_\alpha(R) \leq ES_\alpha(S_d) \leq d,$$

where R are the ray densities of $\mathcal{S}_d(p)[\mathcal{S}_d(p,\rho)]$.

Proof. 1. Let $\tau_S = \mathrm{VaR}_\alpha(S_d) = inf\{y \in \mathbb{R} : P(S_d \leq y) \geq \alpha\}$. Let $\tau_i = \mathrm{VaR}_\alpha(R_i)$, $\tau_M = \max_i \tau_i$ and $\tau_m = \min_i \tau_i$. It holds

$$\begin{aligned} P(S_d \leq \tau_M) &= \sum_{y \leq \tau_M} p_S(y) = \sum_{y \leq \tau_M} \sum_{i=1}^{n_p} \lambda_i p_{R_i}(y) \\ &= \sum_{i=1}^{n_p} \lambda_i \sum_{y \leq \tau_M} p_{R_i}(y) \geq \sum_{i=1}^{n_p} \lambda_i \alpha = \alpha, \end{aligned}$$

thus $\tau_S \leq \tau_M$. It holds

$$P(S_d \leq \tau_m) = \sum_{i=1}^{n_p} \lambda_i \sum_{y \leq \tau_m} p_{R_i}(y) = \sum_{i=1}^{n_p} \lambda_i \beta_i,$$

with $\beta_i \leq \alpha$ therefore we have $P(S_d \leq \tau_m) \leq \alpha$. Thus $\tau_S \geq \tau_m$ and $\tau_m \leq \tau_S \leq \tau_M$.

2. $ES_\alpha \geq \tau_m$ and $ES_\alpha \leq d$ are trivial.

□

The above proposition shows that, in $\mathcal{S}_d(p)$ $[\mathcal{S}_d(p, \rho)]$, VaR_α reaches the maximum and minimum values on the ray densities and therefore we are able to explicitly find them.

Remark 2. *The bounds for ES_α are weak and trivial. Nevertheless, at least in some cases, they cannot be improved. In fact, consider the ray density $\boldsymbol{r} = (1-p, 0, \ldots, 0, p) \in \mathcal{E}_d(p)$. If $1 - p \leq \alpha$ then $ES_\alpha = d$. As a consequence for marginal default probabilities higher then $1 - \alpha$ the bound is reached.*

Thanks to Proposition 4.3 that gives the analytical expression of the ray densities of $\mathcal{S}_d(p)$, the following proposition provides the analytical sharp bounds for VaR_α in $\mathcal{S}_d(p)$.

Proposition 5.2. *Let us consider the class $\mathcal{S}_d(p)$. Let j_1^M be the largest integer smaller than pd, j_2^m be the smallest integer greater than pd and $j_1^p = \frac{(p-(1-\alpha))d}{\alpha}$.*

1. *If $p < 1 - \alpha$, $\min VaR_\alpha(R_{(j_1,j_2)}) = 0$ and $\max VaR_\alpha(R_{(j_1,j_2)}) = [\frac{pd}{1-\alpha}]$ if $\frac{pd}{1-\alpha}$ is not integer and $\max VaR_\alpha(R_{(j_1,j_2)}) = \frac{pd}{1-\alpha} - 1$ if it is integer.*

2. *If $1 - \alpha \leq p \leq 1 - \alpha + \frac{\alpha}{d} j_1^M$, $\min VaR_\alpha(R_{(j_1,j_2)}) = j_1^*$, where j_1^* is the smallest integer greater or equal to j_1^p and $\max VaR_\alpha(R_{(j_1,j_2)}) = d$.*

3. *If $p > 1 - \alpha + \frac{\alpha}{d} j_1^M$, $\min VaR_\alpha(R_{(j_1,j_2)}) = j_2^m = j_1^M + 1$ and $\max VaR_\alpha(R_{(j_1,j_2)}) = d$. In this case, if pd is integer $j_1^M + 1 = pd$.*

Proof. Let us consider first the case pd not integer. The ray densities are given in (4.4) with $0 \leq j_1 \leq j_1^M$ and $j_1^M + 1 \leq j_2 \leq d$. From the definition of VaR_α we have

$$\mathrm{VaR}_\alpha(R_{(j_1,j_2)}) = j_1 \Longleftrightarrow \boldsymbol{r}_{(j_1,j_2)}(j_1) \geq \alpha.$$

It follows

$$\boldsymbol{r}_{(j_1,j_2)}(j_1) = \frac{j_2 - pd}{j_2 - j_1} \geq \alpha,$$

then

$$j_2 \geq -\frac{\alpha}{1-\alpha} j_1 + \frac{pd}{1-\alpha}.$$

We also know that $j_2 \leq d$, so let us determine the point P of intersection of $j_2 = -\frac{\alpha}{1-\alpha} j_1 + \frac{pd}{1-\alpha}$ and $j_2 = d$. The solution of

$$\begin{cases} j_2 = -\frac{\alpha}{1-\alpha} j_1 + \frac{pd}{1-\alpha} \\ j_2 = d \end{cases}$$

is $P = (j_1^P, j_2^P) = (\frac{(p-(1-\alpha))d}{\alpha}, d)$. We distinguish three cases, depending on j_1^P.

1. $j_1^p < 0$. In this case it will follow that $\mathrm{VaR}_\alpha(R_{(0,j_2)}) = 0$ for all $\frac{pd}{1-\alpha} < j_2 \leq d$ and then the minimum value of $\mathrm{VaR}_\alpha(R_{(j_1,j_2)}) = 0$. With respect to the maximum value of $\mathrm{VaR}_\alpha(R_{(j_1,j_2)})$ it will be obtained by $\mathrm{VaR}_\alpha(R_{(0,j_2^*)})$, where j_2^* is the largest integer smaller than $\frac{pd}{1-\alpha}$.

2. $0 \leq j_1^p \leq j_1^M$. Let us define j_1^* as the smallest integer greater or equal to j_1^P. It follows that $\mathrm{VaR}_\alpha(R_{(j_1^*,j_2)}) = j_1^*$, $j_2^* < j_2 \leq d$ with j_2^* is the smallest integer greater or equal to $-\frac{\alpha}{1-\alpha} j_1^* + \frac{pd}{1-\alpha}$. Then the minimum value of $VaR_\alpha(R_{(j_1,j_2)}) = j_1^*$. The maximum value of $\mathrm{VaR}_\alpha(R_{(j_1,j_2)})$ is d.

3. $j_1^p > j_1^M$. In this case $\mathrm{VaR}_\alpha(R_{(j_1,j_2)}) = j_2$. Then the minimum value of $\mathrm{VaR}_\alpha(R_{(j_1,j_2)}) = j_2^m = j_1^M + 1$ and the maximum value of $\mathrm{VaR}_\alpha(R_{(j_1,j_2)}) = d$. If pd is integer we also have $j_1^M + 1 = pd$.

□

Using a different approach and considering Bernoulli mixture models, the problem of finding bounds for VaR_α under variance constraints is also addressed in [13]. Differently from [13], we prove that in our case the bounds are attained, thus they are sharp bounds. Also note that for small marginal default probabilities ($p < 1-\alpha$) the minimum VaR_α is zero and the maximum VaR_α is lower than $\frac{pd}{1-\alpha}$.If $p \geq 1-\alpha$ the maximum VaR_α is d, corresponding to the case where all names default together.

We do not have analytical bounds for the problem with given marginal and correlation, but we can explicitly find the bounds in $\mathcal{S}_d(p,\rho)$ by browsing over all the extremal rays, described in Proposition 4.6.

6 Model risk analysis

The theory developed so far allows us to perform model risk analysis for large portfolios.

Consistently with it, let us suppose we have a credit portfolio P with 100 obligors. Let the random vector $\boldsymbol{X} = (X_1, \ldots, X_{100})$ collect the default indicators for the portfolio P and assume $\boldsymbol{X} \in \mathcal{E}$, where $\mathcal{E} := \mathcal{E}_{100}$. The variable $S := S_{100}$ represents the number of defaults and the distribution of S represents the distribution of the loss. We analytically find bounds of VaR_α and ES_α, for $\alpha = 0.90$, $\alpha = 0.95$ and $\alpha = 0.99$ for two classes of multivariate exchangeable Bernoulli variables - $\mathcal{E}(p)$ and $\mathcal{E}(p, \rho)$ - and different levels of the parameters p and ρ.

The analysis of these two classes of models allows us to study the model risk for given margins and for given margins and correlation. For a specific correlation coefficient we perform a sensitivity analysis of their behavior when ρ changes. For each correlation, we also consider the VaR_α and ES_α associated to a specific joint model (the Bernoulli mixture one), to show how the method can be used to assess the risk of a specific model, considering how far its VaR_α and ES_α are from the bounds.

To complete the picture, for any p we provide the range of admissible moments including correlation, for the hundred Bernoulli variables.

In all cases we consider three scenarios corresponding to three marginal default probabilities $p = 0.3\%$, $p = 1.7\%$ and $p = 26.6\%$, which are the 1-year marginal default probabilities for the rating classes A, BBB and B of Moody's, as resulting from [23] table 13 page 40. We do not investigate the correctness of the marginal default probability, which would be the case if we were investigating marginal model risk.

6.1 Model risk for given margins

Here we deal with $\mathcal{E}(p)$ in the three scenarios $p = 0.3\%, p = 1.7\%$ and $p = 26.6\%$. All the results in this section are analytical, both for the moments and the risk measure. The VaR_α bounds are given by Proposition 5.2 and the ES_α bounds are given by Proposition 5.1. Indeed, bounds for the all moments of the distributions in the class are reached on the ray densities, as proved in [10]. The bounds for the second order moment and correlation are analytical too, and their expressions are given in Section 4.1.1.

6.1.1 Scenario 1: $p = 0.3\%$

Before computing VaR_α and ES_α for the class $\mathcal{S}(0.3\%)$, corresponding to Moody's A rating, let us describe it. As proved in Corollary 4.2, since $pd = 0.3$ is not integer, the number of rays is $n_p = (j_1^M + 1)(d - j_1^M) = 100$, being $j_1^M = 0$ and $d = 100$. The second order moment and correlation are given in Table 1. Obviously, the first moment coincides with p and its range is a singleton. Notice that all positive correlations and some negative are allowed. This is possible since we consider finite sequences of Exchangeable Bernoulli variables and not only the mixing models, i.e. the De Finetti's sequences. So, per se, independently of any model, a hundred Bernoulli default indicators with $p = 0.3\%$ and equicorrelation cannot span all negative dependence, but are able to span any level of positive dependence, zero correlation and negative down to -0.3%.

Order	Min moment	Max moment
2	0	0.003
ρ	-0.003	1

Table 1: Bounds of the moments for the $\mathcal{E}(0.3\%)$ class of multivariate Bernoulli

Table 2 shows the bounds for the VaR_α for the three levels $\alpha = 0.90$, $\alpha = 0.95$ and $\alpha = 0.99$. They are obtained applying Proposition 5.2, case 1.

α	Min VaR_α	Max VaR_α
0.9	0	2
0.95	0	5
0.99	0	29

Table 2: VaR_α of the number of defaults for the $\mathcal{E}(0.3\%)$ class of multivariate Bernoulli

Table 3 shows the bounds for the ES_α on the ray densities for the three levels $\alpha = 0.90$, $\alpha = 0.95$ and $\alpha = 0.99$.These bounds are found by browsing over the extremal rays. The bounds for the entire class, following Proposition 5.1, are broader: 0 and 100 for all the levels of confidence under exam.

α	Min ES_α	Max ES_α
0.9	1	3
0.95	1	6
0.99	1	30

Table 3: ES_α of the number of defaults for the $\mathcal{E}(0.3\%)$ extremal rays

6.1.2 Scenario 2

Let us assume $p = 1.7\%$. Based on Corollary 4.2 the class $\mathcal{S}(1.7\%)$ has $n_p = 2 \cdot 99 = 198$ ray distributions. By computing the correlations of each ray density we notice that we have 198 different correlations. Table 4 provides the bounds of the second order moment and correlation also for this class.

Table 5 shows the bounds for the VaR_α for the three levels $\alpha = 0.90$; $\alpha = 0.95$ and $\alpha = 0.99$. For the first two levels of confidence Proposition 5.2, subcase 1 applies. When $\alpha = 0.99$ Proposition 5.2, subcase 2 applies, with $j_1^p = 0.701$ and $j_1^* = 1$.

Table 6 shows the bounds for the ES on the ray densities for the three levels $\alpha = 0.90$; $\alpha = 0.95$ and $\alpha = 0.99$. The overall bounds, according to Proposition 5.1, are 0 and

Order	Min moment	Max moment
2	$1414{\cdot}10^{-4}$	0.017
ρ	-0.009	1

Table 4: Bounds of the moments for the $\mathcal{E}(1.7\%)$ class of multivariate Bernoulli

α	Min VaR_α	Max VaR_α
0.9	0	16
0.95	0	33
0.99	1	100

Table 5: VaR_α of the number of defaults for the $\mathcal{E}(1.7\%)$ class of multivariate Bernoulli

100 for $\alpha = 0.9$ and $\alpha = 0.95$, 1 and 100 for $\alpha = 0.99$. Since $1.7\% \geq 1\%$ we have $\mathrm{ES}_{0.99} = 100$, as noticed in Remark 2.

α	Min ES_α	Max ES_α
0.9	2	17
0.95	2	34
0.99	2	100

Table 6: ES_α of the number of defaults for the $\mathcal{E}(1.7\%)$ extremal rays

6.1.3 Scenario 3

We consider the class $\mathcal{E}(26.6\%)$. The number of ray densities is much higher relative to the other two classes considered since it is $n_p = 27 \times 74 = 1998$. Table 7 shows that the range of the second order moment and correlation of this class is wider than for the other classes.

Order	Min moment	Max moment
2	0.069	0.266
ρ	-0.01	1

Table 7: Bounds of the moments for the $\mathcal{E}(26.6\%)$ class of multivariate Bernoulli

Table 8 shows the bounds for the VaR_α for the three levels $\alpha = 0.90$; $\alpha = 0.95$ and $\alpha = 0.99$. They obtain from Proposition 5.2 subcase 2, with $j_1^p = 18.44, 22.74$ and 25.86 respectively, which gives $j^* = 19, 23, 26$.

The following Table 9 shows the bounds for the ES_α on the ray densities for the three levels $\alpha = 0.90$; $\alpha = 0.95$ and $\alpha = 0.99$. The overall bounds would be $(19, 100)$, $(23, 100)$ and $(26, 100)$. As one can see the maximum ES_α is d=100 for each α, since $26.6\% > 1 - \alpha$.

α	Min VaR_α	Max VaR_α
0.9	19	100
0.95	23	100
0.99	26	100

Table 8: VaR_α of the number of defaults for the $\mathcal{E}(26.6\%)$ class of multivariate Bernoulli

α	Min ES_α	Max ES_α
0.9	27	100
0.95	27	100
0.99	27	100

Table 9: ES_α of the number of defaults for the $\mathcal{E}(26.6\%)$ extremal rays

6.1.4 Cross scenario comparisons

The bounds for the second order moment are increasing from the first to the third, according to Proposition 4.5, given that pd is never an integer, and d is equal in the three scenarios. This happens as a combined effect of the behaviour of p and j_1^M. The lower bound of correlation, because of Corollary 4.3, is decreasing, while the upper bound is at the maximum, 1. The reader can appreciate how model risk increases, when the marginal probability does, and when the risk measure is VaR_α, by comparing the VaR_α range in Tables 2, 5, 8, or by looking at Figure 1, which represents directly the ranges. The VaR_α and ES_α ranges increase with the marginal default probability, and not only with the level of confidence, which is the standard result. Also, both the minimum and the maximum are non decreasing with p, there is an exception, namely $\alpha = 99\%$ when p goes from 1.7 to 26.6%.

From the point of view of finance, in all cases model risk, whether it is measured by a coherent measure (ES_α) or by a regulatory one (VaR_α), increases exactly when it matters more, namely when going to worse rating classes, whose marginal default is more likely. Obviously, the higher the marginal default probabilities, the higher the maximum VaR_α, for any p and α, but also the more it coincides with the worst possible loss, which has been normalized to 100.

6.2 Model risk for given margins and correlations

In this section we examine the behavior of the loss under the three scenarios above for the marginal default probability, when, on top of it, a specific value of the equicorrelation has been selected. We deal with $\mathcal{E}(p, \rho)$ in the three scenarios $p = 0.3\%, p = 1.7\%$ and $p = 26.6\%$ and provide bounds for VaR_α for three levels of correlation: $\rho = \frac{1}{6}; \frac{1}{2}; \frac{5}{6}$. Here, the ray densities are analytical as well as their VaR_α. The bounds are found by computationally searching the maximum and minimum VaR_α among the ray densities.

We also report the VaR_α corresponding to an exchangeable Bernoulli mixing model from the credit risk literature (for a complete overview see [7]) to give a sense of how a specific model can enter the bounds and whether it maintains a closedness to the lower or upper bound independently of the marginal probabilities or other parameters. We estimate the β-mixing model of each scenario and compute its VaR_α. Let us recall that, if S_β is the number of defaults of the β-mixing models, we have:

$$p_\beta(j) = \binom{d}{j} \int_0^1 p^k (1-p)^{d-k} d\Psi(p),$$

where $\Psi \sim \beta(a, b)$ is the mixing variable. We have

$$\begin{aligned} p &= E[\Psi] \\ \mu_2 &= E[\Psi^2]. \end{aligned}$$

Therefore we estimate the β parameters a and b by solving the equations

$$\begin{aligned} p &= \frac{a}{a+b} \\ \mu_2 &= \frac{a(a+1)}{(a+b)(a+b+1)}. \end{aligned}$$

Notice that for this model $\rho = 0$ is not admissible.

6.2.1 Scenario 1

Within the first scenario, $p = 0.3\%$ we consider three levels of correlation. Table 10 provides the bounds of VaR_α when correlation is assumed or calibrated to be $\frac{1}{6}$ and the single VaR_α for the β-mixing model.

Quantile	minVaR_α	max VaR_α	β -VaR_α
0.9	0	2	0
0.95	0	5	0
0.99	1	22	9

Table 10: VaR_α of the number of defaults for the β and the $\mathcal{E}(0.3\%, \frac{1}{6})$ class of multivariate Bernoulli

Table 11 provides the bounds of VaR_α when correlation is $\frac{1}{2}$ and the single VaR_α for the β-mixing model.

Table 12 provides the bounds of VaR_α when correlation is $\frac{5}{6}$ and the single VaR_α for the β-mixing model.

Within this scenario, the min and max VaR_α are weakly increasing with the confidence level, for given ρ, and weakly decreasing with ρ, for any given confidence level. The range is weakly increasing with the confidence level for given ρ and weakly decreasing with ρ for any confidence level.

Quantile	min VaR_α	max VaR_α	β -VaR_α
0.9	0	1	0
0.95	0	3	0
0.99	0	21	4

Table 11: VaR_α of the number of defaults for the β and the $\mathcal{E}(0.3\%, \frac{1}{2})$ class of multivariate Bernoulli

Quantile	min VaR_α	max VaR_α	β -VaR_α
0.9	0	0	0
0.95	0	1	0
0.99	0	7	0

Table 12: VaR_α of the number of defaults for the β and the $\mathcal{E}(0.3\%, \frac{5}{6})$ class of multivariate Bernoulli

6.2.2 Scenario 2

Within the second scenario, $p = 1.7\%$, we consider the same three levels of correlation. Table 13 provides the bounds of VaR_α when correlation is $\frac{1}{6}$ and the β-mixing model VaR_α.

Quantile	min VaR_α	max VaR_α	β -VaR_α
0.9	0	16	5
0.95	1	25	11
0.99	2	55	29

Table 13: VaR_α of the number of defaults for the β and the $\mathcal{E}(1.7\%, \frac{1}{6})$ class of multivariate Bernoulli

Table 14 provides the bounds of VaR_α when correlation is $\frac{1}{2}$ and the single VaR_α for the β-mixing model.

Quantile	minVaR_α	max VaR_α	β -VaR_α
0.9	0	9	0
0.95	0	25	5
0.99	1	93	57

Table 14: VaR_α of the number of defaults for the β and the $\mathcal{E}(1.7\%, \frac{1}{2})$ class of multivariate Bernoulli

Table 15 provides the bounds of VaR_α when correlation is $\frac{5}{6}$ and the single VaR_α for the β-mixing model.

Comments similar to the ones in Scenario 1 apply.

Quantile	min VaR_α	max VaR_α	β -VaR_α
0.9	0	3	0
0.95	0	8	0
0.99	61	100	94

Table 15: VaR_α of the number of defaults for the β and the $\mathcal{E}(1.7\%, \frac{5}{6})$ class of multivariate Bernoulli

6.2.3 Scenario 3

Within the third scenario, $p = 26.6\%$, we consider the same three levels of correlation. Table 16 provides the bounds of VaR_α when correlation is $\frac{1}{6}$ and the unique VaR_α for the β-mixing model. With the purpose of showing that the number of generators of a class vary significantly, and can be very huge, we provide here the number of ray densities of this class. In this case we do not have a closed formula to find the number of rays, but since we find them analytically, we enumerate them. They are 32.372.

Quantile	min VaR_α	max VaR_α	β -VaR_α
0.9	21	82	53
0.95	26	100	62
0.99	38	100	76

Table 16: VaR_α of the number of defaults for the β and the $\mathcal{E}(26.6\%, \frac{1}{6})$ class of multivariate Bernoulli

Table 17 provide the bounds of VaR_α when correlation is known and it is $\frac{1}{2}$ and the single VaR_α for the β-mixing model.

Quantile	min VaR_α	max VaR_α	β -VaR_α R
0.9	42	100	82
0.95	56	100	93
0.99	63	100	100

Table 17: VaR_α of the number of defaults forthe β and the $\mathcal{E}(26.6\%, \frac{1}{2})$ class of multivariate Bernoulli

Table 18 provide the bounds of VaR_α when correlation is known and it is $\frac{5}{6}$ and the single VaR_α for the β-mixing model.

Comments similar to scenarios 1 and 2 apply.

6.2.4 Cross scenario comparisons

From the tables above we note that all others equal, the bound range is smaller than when correlation is not specified (Section 6.1). For a given level of confidence and ρ,

Quantile	min VaR_α	max VaR_α	β -VaR_α
0.9	81	100	100
0.95	86	100	100
0.99	88	100	100

Table 18: VaR_α of the number of defaults for the β and the $\mathcal{E}(26.6\%, \frac{5}{6})$ class of multivariate Bernoulli

as the marginal default probability increases (from scenario 1 to 3) the minimum and maximum VaR_α are non decreasing, and the maximum VaR_α reaches the maximum possible loss when the default is as high as 26.6%, as it happened without an assigned value of ρ. For given level of confidence and ρ, the VaR_α range is decreasing in p.

The choice of a specific model, the popular mixing β one, obviously captures only one value of VaR_α within the range, with no regularity on whether this is closer to the minimum or maximum. So, even for such a specific model, there is no way to conclude that it is pessimistic or optimistic in the quantile-based measurement of losses. This is particularly relevant for applications since, as we know, Value-at-Risk is currently included in the regulatory provisions for banks and insurance companies. The fact that a specific model, like the β-mixing one, does not ensure to be pessimistic or optimistic, even after selection of correlation in evaluating VaR_α, is therefore of the utmost importance in regulatory-based choices of credit-risk models.

To make these conclusions stronger, figures 2, 3 and 4 plot the bounds for VaR_α when ρ takes equi-spaced values in the range $[0, \frac{11}{12}]$. The reader can appreciate how calibration risk as measured by the range of VaR_α weakly increases, for given ρ, when the marginal probability or the level of confidence does (so, from top to bottom across figures and from left to right in any given figure). For given level of confidence and p the maximum VaR_α tends to increase with ρ. Nevertheless, for small values of correlation it happens that the minimum VaR_α is decreasing with correlation. This is consistent with the theory, even if we are spanning positive correlation only, since quantiles are not linear functions of the variance. To give an example consider the following two ray densities r_m^1 and r_m^2 in $S_d(1.7\%)$ with correlations $\rho_1 = \frac{1}{6}$ and $\rho_2 = \frac{1}{2}$ respectively, such that $\mathrm{VaR}_\alpha(r_m^1) \geq \mathrm{VaR}_\alpha(r_m^2)$, $\alpha = 0.99$.

Let

$$r_m^1(y) = \begin{cases} 0.4112205 & y = 0 \\ 0.5789222 & y = 2 \\ 0.0098574 & y = 55 \\ 0 & otherwise \end{cases},$$

$$r_m^2(y) = \begin{cases} 0.2104314 & y = 0 \\ 0.779779 & y = 1 \\ 0.0097896 & y = 94 \\ 0 & \text{otherwise} \end{cases} .$$

Easy calculations shows that they have mean 1.7 and therefore they are in $S_d(1.7\%)$. Since they have mass on three points they are ray densities (they are solutions of the linear system (4.12)). To verify correlations we first compute $E[(r_m^1)^2] = 32.13425$ and $E[(r_m^2)^2] = 87.28055$. Then $\mu_2^1 = 0.0030742$ and $\mu_2^2 = 0.0086445$ by using (4.7). Finally $\rho_1 = \frac{1}{6}$ and $\rho_2 = \frac{1}{2}$ are obtained by inverting (3.1). Looking at the densities it is evident that $\mathrm{VaR}_{0.99}(r_m^1) = 2$ and $\mathrm{VaR}_{0.99}(r_m^2) = 1$.

It also emerges from the figures that the VaR_α of the β-mixing model sometimes reaches the bound and depending on p and ρ its values with respect to the bounds significantly change. In particular for low (high) p the VaR_α of the β-mixing model coincides with the minimum (maximum) VaR_α. Figures 2, 3 and 4 show that, even if the β-mixing model is calibrated to match the moments of the Bernoulli, it tends to produce a VaR_α close to the minimum one for low p, and close to the maximum for high p. In any case, the width of the band between the minimum and the maximum, together with the specific location of the β VaR_α within it, give a sense of how stringent is the choice of a specific multivariate distribution.

Acknowledgements

The authors wish to thank Fabrizio Durante for his helpful comments. The authors also thank the Associate Editor and the referee for their valuable suggestions.

References

[1] R. Cont, “Model uncertainty and its impact on the pricing of derivative instruments,” *Mathematical finance*, vol. 16, no. 3, pp. 519–547, 2006.

[2] R. C. Merton, “On the pricing of corporate debt: The risk structure of interest rates,” *The Journal of Finance*, vol. 29, no. 2, pp. 449–470, 1974.

[3] R. A. Jarrow and S. M. Turnbull, “Credit risk: Drawing the analogy,” *Risk Magazine*, vol. 5, no. 9, 1992.

[4] T. Bielecki, M. Jeanblanc, and M. Rutkowski, “Credit risk modeling.,” *CSFI Lecture Notes Series*, vol. 2, 2009.

[5] U. Cherubini, E. Luciano, and W. Vecchiato, *Copula methods in finance*. John Wiley & Sons, 2004.

[6] D. Duffie and K. Singleton, *Credit risk: Pricing, Measurement, and Menagement*. Princeton University Press, 2003.

[7] A. J. McNeil, R. Frey, and P. Embrechts, *Quantitative risk management*, vol. 3. Princeton University Press, 2005.

[8] P. Embrechts, A. McNeil, and D. Straumann, “Correlation and dependence in risk management: properties and pitfalls,” *Risk management: value at risk and beyond*, vol. 1, pp. 176–223, 2002.

[9] P. Embrechts, A. Höing, and A. Juri, “Using copulae to bound the value-at-risk for functions of dependent risks,” *Finance and Stochastics*, vol. 7, no. 2, pp. 145–167, 2003.

[10] R. Fontana and P. Semeraro, “Representation of multivariate bernoulli distributions with a given set of specified moments,” *Journal of Multivariate Analysis*, vol. 168, pp. 290–303, 2018.

[11] W. Fenchel and D. W. Blackett, *Convex cones, sets, and functions*. Princeton University, Department of Mathematics, Logistics Research Project, 1953.

[12] D. Avis, “On the extreme rays of the metric cone,” *Canadian Journal of Mathematics*, vol. 32, no. 1, pp. 126–144, 1980.

[13] C. Bernard, L. Rüschendorf, and S. Vanduffel, “Value-at-risk bounds with variance constraints,” *Journal of Risk and Insurance*, vol. 84, no. 3, pp. 923–959, 2017.

[14] P. Embrechts and G. Puccetti, “Bounds for functions of dependent risks,” *Finance and Stochastics*, vol. 10, no. 3, pp. 341–352, 2006.

[15] B. Wang and R. Wang, "The complete mixability and convex minimization problems with monotone marginal densities," *Journal of Multivariate Analysis*, vol. 102, no. 10, pp. 1344–1360, 2011.

[16] G. Puccetti and L. Rüschendorf, "Sharp bounds for sums of dependent risks," *Journal of Applied Probability*, vol. 50, no. 1, pp. 42–53, 2013.

[17] G. Puccetti and L. Rüschendorf, "Bounds for joint portfolios of dependent risks," *Statistics & Risk Modeling with Applications in Finance and Insurance*, vol. 29, no. 2, pp. 107–132, 2012.

[18] P. Embrechts, G. Puccetti, and L. Rüschendorf, "Model uncertainty and var aggregation," *Journal of Banking & Finance*, vol. 37, no. 8, pp. 2750–2764, 2013.

[19] G. Puccetti, L. Rüschendorf, and D. Manko, "Var bounds for joint portfolios with dependence constraints," *Dependence Modeling*, vol. 4, no. 1, pp. 368–381, 2016.

[20] M. Terzer, *Large scale methods to enumerate extreme rays and elementary modes.* PhD thesis, ETH Zurich, 2009.

[21] N. Kolev and D. Paiva, "Multinomial model for random sums," *Insurance: Mathematics and Economics*, vol. 37, no. 3, pp. 494–504, 2005.

[22] C. Acerbi and D. Tasche, "On the coherence of expected shortfall," *Journal of Banking & Finance*, vol. 26, no. 7, pp. 1487–1503, 2002.

[23] S&P, "Default, transition, and recovery: 2017 annual global corporate default study and rating transitions," 2018.

Figure 1: VAR ranges for $p = 0.03\%; 1.7\%; 26.6\%$

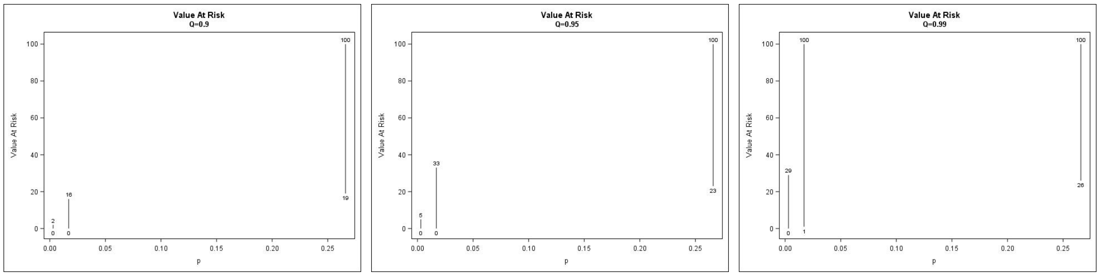

Figure 2: VAR bounds and β-mixing model VAR for $p = 0.03\%$ and different ρs

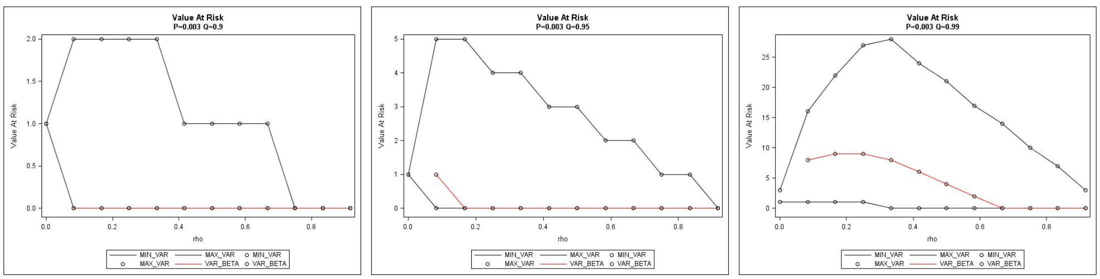

Figure 3: VAR bounds and β-mixing model VAR for $p = 1.7\%$ and different ρs

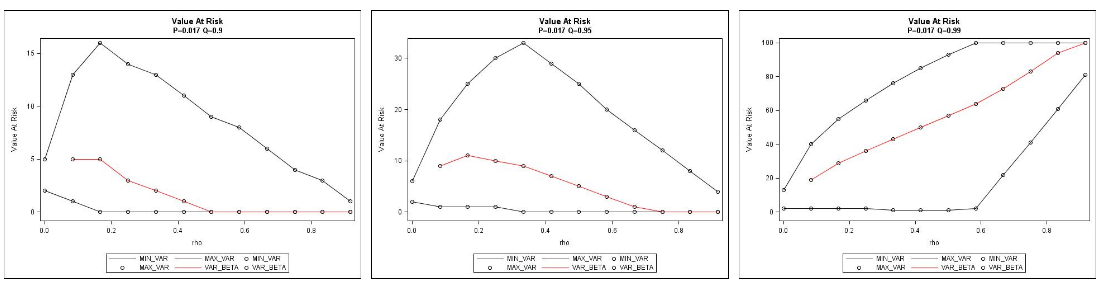

Figure 4: VAR bounds and β-mixing model VAR for $p = 26.6\%$ and different ρs