I N S T I T U T D E S T A T I S T I Q U E

B I O S T A T I S T I Q U E E T

S C I E N C E S A C T U A R I E L L E S

(I S B A)

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

D I S C U S S I O N
P A P E R

2014/09

Measuring Portfolio Risk under Partial Dependence Information

BERNARD, C., DENUIT, M. AND S. VANDUFFEL

MEASURING PORTFOLIO RISK UNDER PARTIAL DEPENDENCE INFORMATION

Carole Bernard
Department of Statistics and Actuarial Science
University of Waterloo
Waterloo, Canada

Michel Denuit
Institut de Statistique, Biostatistique & Sciences Actuarielles
Université Catholique de Louvain
Louvain-la-Neuve, Belgium

Steven Vanduffel[1]
Faculty of Economics
Vrije Universiteit Brussel
Elsene, Belgium

March 8, 2014

[1]Corresponding author. Carole Bernard acknowledges support from NSERC, the CAE research grant in Waterloo and the Humboldt Research Foundation. The financial support of PARC "Stochastic Modelling of Dependence" 2012-17 awarded by the "Communauté française de Belgique" is gratefully acknowledged by Michel Denuit. Steven Vanduffel acknowledges support from BNP Paribas Fortis. We thank Jing Yao for excellent research assistance.

Abstract

There is a recent interest in finding bounds for risk measures of portfolios when the marginal distributions of its risky components are assumed to be known. This problem is well studied when the dependence among the risks is unknown, but the bounds obtained are too wide to be useful in practice. Unfortunately, additional dependence information, such as knowledge of some higher-order moments, makes the problem hard to deal with.

We motivate that replacing the knowledge of the marginal distributions by the knowledge of the mean of the portfolio sum does *not* result in significant loss of information, while making it possible to find explicit bounds if also higher-order moments as source of dependence information are available. Effectively, we propose a new and elementary derivation of bounds on various risk characteristics, including distribution functions and Value-at-Risk (VaR).

Our results make it possible for supervisory authorities to assess the robustness of risk models used in practice and to identify issues of incomparability of risk models across different institutions.

Key words: Stochastic dominance, Moment space, s-convex order, Value-at-Risk.

1 Motivation

Quantifying the risk of a sum of random variables has always been a central problem in various disciplines including engineering, finance, risk management and insurance. Clearly, portfolios of risks are at the core of insurance business as this activity is based on a risk exchange between the policyholders who want to get rid of their individual risk and the insurer who counts on diversification effects to keep the riskiness of the entire portfolio under control. But our study may also apply to engineering where sums of random variables naturally appear when studying computer networks design problems and traffic flow problems or in finance where they arise in the valuation of Asian or basket options, amongst others. For an insurance portfolio, the assumption of independence between the risky components is sometimes realistic in which case the insurer can resort to the Central Limit Theorem or more accurate methods such as Panjer's recursion or Monte Carlo simulation to quantify with precision the maximum loss (Value-at-Risk (VaR)) he can suffer in a given period of time at a certain probability level. In the majority of cases, however, the different risks are influenced by one or more common factors such as geography, inflation or economic environment, and it is hard to specify the joint distribution.

Ultimately, some assumptions will be made resulting in the choice of a multivariate model that makes it possible to measure the risk of the portfolio at hand while, in fact, there are many models that are consistent with the available information. It is then of great interest (and importance) for all stakeholders involved to assess the robustness of the model with respect to a change in its underlying assumptions: what is the maximum or minimum value for a certain risk measure that can be justified given a certain set of information? In particular, regulators, confronted with scandals such as the subprime crisis and the LIBOR fraud seem to care more and more about model uncertainty (and complexity). For example, in their discussion paper, the Basel Committee (2013) insists that a desired objective of a solvency framework concerns comparability: "Two banks with portfolios having identical risk profiles apply the framework's rules and arrive at the same amount of risk-weighted assets and two banks with different risk profiles should produce risk numbers that are different proportionally to the differences in risk". The financial news agency Bloomberg (2013) recently reported: "While firms submit their models to national regulators for validation, they don't have to disclose them publicly. Surveys by the Basel Committee have shown that risk-weightings for the same assets vary among banks, undermining their credibility".

Assessing model uncertainty is not straightforward to do. There is a rich literature devoted to bounds on the distribution function or, equivalently, the Value-at-Risk (VaR) of a risky portfolio when the marginal distributions of the components (describing the stand-alone risks) are assumed to be known but not their interdependence; see amongst others Rüschendorf (1982), Denuit et al. (1999), Embrechts and Puccetti (2006a), Puccetti and Rüschendorf (2012a) and Embrechts et al. (2013). Specifically, when the portfolio is homogeneous (all marginal distributions are identical) one can often obtain sharp (attainable, best-possible) bounds explicitly. However, the analysis for inhomogeneous portfolios is fairly complicated and theoretical results are scarce. Recently, Puccetti and Rüschendorf (2012a) have introduced the rearrangement algorithm (RA) as a practical way to approximate sharp bounds on the distribution of a portfolio sum. The RA was further used by Embrechts et al. (2013) to approximate sharp upper and lower bounds for the VaR of a portfolio. While their numerical examples provide evidence that the RA is indeed able to approximate the sharp bounds accurately, they also show that the gap between worst-case and best-case VaR numbers is usually very high. Furthermore, the upper bound on VaR is always larger than the VaR one would obtain in the case that all risks have maximum correlation (comonotonic), a situation that is hard to accept by practitioners. For example, Embrechts et al. (2013) show in their Figure 5 that for a portfolio of Pareto(2) distributed risks the upper bound on the VaR is about 2 times larger than the comonotonic VaR (i.e. when the marginal risks are assumed to be comonotonic). More generally, Puccetti and Rüschendorf (2012d) show that under some mild conditions the worst Value-at-Risk behaves asymptotically as the worst Tail Value-at-Risk (TVaR). The intuition behind this result is as follows. The VaR (measured at some probability level p) of a comonotonic sum is of course just a particular point of the quantile function of this sum. Now, by changing the comonotonic dependence in the *upper* tail of the marginal supports (from level p onwards), one is able to adjust the upper quantiles of the sum (from level p onwards). As the quantile function is non-decreasing, it is then clear that the highest VaR will be obtained if one can change the dependence such that the quantile function of the sum becomes a constant on $(p, 1)$. The constant value is then the maximum VaR and is equal to the comonotonic TVaR (Bernard et al. (2013)).

The situation described above stresses that information on the dependence is crucial if one wants to build models that provide risk numbers that are trustworthy in the sense that upper and lower bounds for these numbers stay in some reasonable range. For example, one may want to assume that the risks are positively dependent, the variance of the aggregate risk can be estimated accurately from a statistical analysis of observed losses or some information on the copula function might be available. In the context of additional dependence information, the results in the literature are more limited and of an ad-hoc nature. Rüschendorf (1991), Embrechts and Puccetti (2009), Embrechts et al. (2013) consider the situation in which some of the bivariate distributions are known, Denuit et al. (1999) study VaR bounds assuming that the joint distribution of the risks is bounded by some distribution, and Cheung and Vanduffel (2013) study bounds for sums of random variables when the marginals and the variance of the sum are known. However, the bounds that are proposed in these papers are often hard to deal with, especially for high-dimensional and inhomogeneous portfolios, and they do not necessarily sharpen the unconstrained bounds in a significant way; see also Chernih et al. (2010) for an illustration in the context of credit risk portfolio modelling. We quote from Embrechts et al. (2013): "Additional positive dependence information added on top of the marginal distributions does not improve the VaR bounds substantially".

To some extent, these observations contrast with the findings of Bernard et al. (2013). They consider the presence of a variance constraint on the portfolio sum as a source of dependence information and show that doing so can significantly tighten the (unconstrained) VaR bounds. Indeed, the basic insight in Bernard et al. (2013) is that, in presence of a variance constraint, an upper bound for the VaR can be obtained by changing the comonotonic dependence on the *entire* support of the marginal distributions, in such a way that the resulting portfolio sum (the quantile function) becomes atomic and takes two values only, say a and b ($a < b$), that have maximum difference while satisfying the global variance constraint. The constrained upper bound for the VaR, i.e. b, will usually be lower than the unconstrained bound B and the difference between both bounds can be substantial. In other words, in the above statement of Embrechts et al. (2013) one may change the word "can" into "may" and add to it "especially, in the situation in which one is concerned with performing VaR calculations at very high probability levels and the average correlation is high"; see also Bernard et al. (2013) for more information and examples that demonstrate the impact of a variance constraint on the (unconstrained) VaR bounds.

Bernard et al. (2013) also show that the constrained VaR bound b is closely related to the classical Cantelli bound (involving mean and standard deviation) for distributions. Their construction draws an interesting parallel between, on the one hand, the problem of finding VaR bounds when the marginal distributions are given and some higher-order moments, and, on the other hand, the problem of finding bounds when the mean of the portfolio sum (but *not* the marginal distributions) and the higher-order moments are given. Indeed, it is well-known that solutions to the latter problem are in most cases given by atomic random variables, which is consistent with their findings.

This observation is important, as the first problem (i.e., the inclusion of the knowledge of the marginal distributions in the set-up of the problem) is clearly relevant but hard to deal with and few explicit results are available. In contrast, the second problem (only some moments of the portfolio are available) is substantially easier to deal with, and the loss of information by changing "the knowledge of all marginal distributions" by "the knowledge of the portfolio mean" appears to be limited. In fact, by skipping the information on the marginal distributions and adding the information on the higher-order moments instead, one effectively has the opportunity to exploit the available dependence information. Moreover, in practice loss statistics may only be available at the portfolio level, or, may not be sufficiently rich enough, to derive the marginal distributions of the risks involved. In contrast, the setting of the moment bounds problem is exactly adapted to the situation in which one has portfolio statistics available. Note however that the higher the order of the moments the more vulnerable their estimators become to outliers, and more observations will be needed to estimate them with sufficient confidence. A possible way to deal with this issue consists in using robust estimators; see e.g. Huber and Ronchetti (2009). Alternatively, one may consider the statistical uncertainty on the moments explicitly and our framework allows to deal with it.

In summary, it is of crucial importance to consider all available dependence information when finding the VaR of a portfolio and casting the problem as a higher moment problem (hereby no longer using the marginal distributions) provides a way to deal with this problem efficiently.

In this paper, we provide a new derivation of moment bounds on various risk characteristics, including distribution functions and Value-at-Risk. This problem has already been studied in the literature by various authors including Hürlimann (1998), Hürlimann (2002), De Schepper and Heijnen (2010), Kaas and Goovaerts (1986a), (1986b). Given a risk X, Kaas and Goovaerts (1986a), (1986b) derive lower and upper bounds on quantities of the form $E[g(X)]$, for some given function g, when X belongs to a class of random variables satisfying certain moment conditions (see also the references in their paper). Hürlimann (2002) finds analytical bounds for the VaR that are based on the mean, variance, skewness and kurtosis. The standard approaches in the literature are based on linear programming or special polynomials and their solutions were, in most cases, given by atomic random variables with similar structure.

The present paper exploits this remarkable feature and derives such bounds in an elementary way using well-known crossing properties of distribution functions and some stochastic dominance concepts. Our approach is straightforward to implement and allows to deal efficiently with several problems of interest in insurance and finance. We illustrate this feature by discussing model uncertainty on credit risk portfolio modelling and on the collective risk model (the latter is a standard model used in the insurance industry). We show that adding dependence information can reduce model error significantly. However, model risk is a concern that cannot be readily ignored. We argue that supervisory authorities can use the techniques developed in this paper to assess how robust (VaR based) capital requirements are with respect to model missspecification. In practice, the moments are often not known with certainty but need to be estimated from data. To account for this feature, we extend our framework to also include statistical uncertainty on the moment constraints.

The remainder of this paper is organized as follows. In Section 2, sets of distributions with prescribed moment sequence, the so-called moment spaces, are discussed. Then, bounds on distribution functions with given moments are derived in an elementary, probabilistic way in Section 3. Section 4 is devoted to the corresponding moment bounds on quantiles, or Value-at-Risk (VaR). Section 4.2 extends the results in the presence of statistical uncertainty and offers an application to the collective risk model. More applications are developped in Section 5. The final Section 6 concludes and provides some policy implications.

2 Moment spaces

2.1 Description

We denote by $\mathcal{M}_s([a,b];\mu_1,\mu_2,\ldots,\mu_{s-1})$ the class of all the random variables X valued in the interval $[a,b]$ ($a,b\in\mathbb{R}$) and with prescribed first $s-1$ moments $EX^k=\mu_k$, $k=1,2,\ldots,s-1$. The restriction to finite supports may appear as a limitation of the setting but it is actually not. Indeed, the upper limit to the financial loss for which the insurance company underwrites is generally fixed by the contract or determined through reinsurance techniques. Note also that in any case insurers have limited liability (up to their capital) and finite resources to meet claims. In the remainder of the paper, we discuss separately two cases according to the parity of s.

- When s is even, then $s=2m$ for some positive integer m and an odd number of moments $\mu_1,\ldots,\mu_{2m-1}$ are known.

- When s is odd, then $s = 2m + 1$ for some non-negative integer m and an even number of moments $\mu_1, \mu_2, \ldots, \mu_{2m}$ are known.

It is clear that not every vector $(\mu_1, \mu_2, \ldots, \mu_{s-1}) \in \mathbb{R}^{s-1}$ corresponds to a sequence of moments, i.e. we may have that $\mathcal{M}_s([a, b]; \mu_1, \mu_2, \ldots, \mu_{s-1}) = \emptyset$. Furthermore, other sequences of moments may correspond to a unique probability distribution. For example, if $\xi \in [a, b]$ then the moment sequence $(\xi, \xi^2, \ldots, \xi^{s-1})$ corresponds to the degenerated random variable $X = \xi$ a.s. for any $s \geqslant 3$. These situations are clearly without further interest and we would like to impose conditions on moment sequences to ensure that the moment space at hand contains more than one element[1]. It is possible to describe these conditions precisely in terms of some moment matrices (see Appendix A). In what follows we always tacitly assume that these conditions are satisfied.

A special role in a given moment space is then played by two elements with a minimum number of support points. These variables correspond to the so-called s-convex minimum resp. maximum in the moment space $\mathcal{M}_s([a, b]; \mu_1, \mu_2, \ldots, \mu_{s-1})$ and their properties will be crucial in the construction of bounds on various risk quantities of interest. We recall the concept of $s-$convex order first and next explain to construct the distributions of the $s-$convex extrema explicitly; see also Denuit et al. (1999) for more details.

2.2 s-convex order

Let X and Y be two random variables in $\mathcal{M}_s([a, b]; \mu_1, \mu_2, \ldots, \mu_{s-1})$. If the inequality $E[g(X)] \leqslant E[g(Y)]$ holds for every function g with a non-negative $s-$th order derivative such that the expectations exist, then X is said to be smaller than Y in the s-convex order sense (denoted $X \preceq_{s-cx} Y$). Note that the 1-convex order is the usual first-degree stochastic dominance and the 2-convex order is the usual convex order or second-order stochastic dominance. Many properties of the s-convex stochastic order can be found in Denuit et al. (1998).

It is clear that when $s = 2$ the distribution functions F_X and F_Y must cross at least once (as they share the same mean). It is also well-known that if they cross exactly once then X and Y are ordered in the sense of 2-convex order. Interestingly, these crossing properties can be generalized to any value of s. To this end, we recall that for a real-valued function g defined on a subset I of the real line, the number of changes of sign of g on this subset I is

$$S^-(g) = \sup S^-[g(\xi_1), g(\xi_2), \ldots, g(\xi_n)], \tag{2.1}$$

where the supremum is extended over all sets $\xi_1 < \xi_2 < \cdots < \xi_n \in I$, n is arbitrary but finite and $S^-[\theta_1, \theta_2, \ldots, \theta_n]$ is the number of sign changes of the indicated sequence $\{\theta_1, \theta_2, \ldots, \theta_n\}$, zero terms being discarded (see, for example, Karlin (1968)). The functions g_1 and g_2 are said to cross each other k times if $S^-(g_1 - g_2) = k$. As the sign of the $s-$th derivative of the test function g is controlled, the s-convex order is closely related to the higher-degree increases in risk introduced by Ekern (1980).

[1] and thus many more elements as the moment space is closed under mixing.

We can now formulate the following property; see also Theorems 4.2 and 4.3 in Denuit et al. (1998) and Proposition 3.2 in Denuit et al. (1999).

Property 2.1 (Crossing conditions). *Let X and Y be two random variables in the set $\mathcal{M}_s([a,b];\mu_1,\mu_2,\ldots,\mu_{s-1})$ with different distributions F_X and F_Y, then, $S^-(F_X-F_Y)\geqslant s-1$. Moreover, $S^-(F_X-F_Y)=s-1$ implies that $X\preceq_{s-cx}Y$ provided the last sign of F_X-F_Y is a "+". By symmetry of the argument, $S^-(F_X-F_Y)=s-1$ implies that $Y\preceq_{s-cx}X$ provided the last sign of F_X-F_Y is a "-".*

2.3 Identification of s-convex extrema

Let us now construct the distributions of the s-convex minimum and maximum in the moment space $\mathcal{M}_s([a,b];\mu_1,\mu_2,\ldots,\mu_{s-1})$. In doing so, we rely on a more general result of De Vijlder (1996) that allows finding a discrete distribution in a given moment space when some points of the support are fixed. Specifically, if $\alpha_1,\ldots,\alpha_i$ are fixed support points in $[a,b]$ and $\theta_1,\ldots,\theta_j$ are unknown support points in (a,b), De Vijlder (1996) derives in Section 8.2.2, Theorem 4, the unique Y in $\mathcal{M}_{i+2j}([a,b];\mu_1,\mu_2,\ldots,\mu_{i+2j-1})$ having $\alpha_1,\ldots,\alpha_i$ and $\theta_1,\ldots,\theta_j$ as support points. This construction can be explained as follows. *First*, the unknown support points $\theta_1,\ldots,\theta_j$ are the roots of the following equation based on a Vandermonde determinant

$$\begin{vmatrix} 1 & x & x^2 & \cdots & x^j \\ \mu_{i;\alpha_1,\ldots,\alpha_i} & \mu_{i+1;\alpha_1,\ldots,\alpha_i} & \mu_{i+2;\alpha_1,\ldots,\alpha_i} & \cdots & \mu_{i+j;\alpha_1,\ldots,\alpha_i} \\ \mu_{i+1;\alpha_1,\ldots,\alpha_i} & \mu_{i+2;\alpha_1,\ldots,\alpha_i} & \mu_{i+3;\alpha_1,\ldots,\alpha_i} & \cdots & \mu_{i+j+1;\alpha_1,\ldots,\alpha_i} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{i+j-1;\alpha_1,\ldots,\alpha_i} & \mu_{i+j;\alpha_1,\ldots,\alpha_i} & \mu_{i+j+1;\alpha_1,\ldots,\alpha_i} & \cdots & \mu_{i+2j-1;\alpha_1,\ldots,\alpha_i} \end{vmatrix} = 0, \tag{2.2}$$

with the elements of the matrix defined recursively as follows $\mu_{k;\alpha_1,\ldots,\alpha_l}=\mu_{k;\alpha_1,\ldots,\alpha_{l-1}}-\alpha_l\mu_{k-1;\alpha_1,\ldots,\alpha_{l-1}}$ starting from $\mu_{k;\alpha_1}=\mu_k-\alpha_1\mu_{k-1}(\mu_1=1,\ \mu_{-1}=0)$. This fundamental result is easily derived as follows. First, notice that for any $k\geqslant\ell$,

$$\mu_{k;\alpha_1,\ldots,\alpha_\ell}=E\left[Y^{k-\ell}\prod_{n=1}^{\ell}(Y-\alpha_n)\right].$$

Denote $r_k:=\mathrm{P}[Y=\theta_k]$. Equation (2.2) is now a direct consequence of the matrix factorization

$$\begin{pmatrix} 1 & x & x^2 & \cdots & x^j \\ \mu_{i;\alpha_1,\ldots,\alpha_i} & \mu_{i+1;\alpha_1,\ldots,\alpha_i} & \mu_{i+2;\alpha_1,\ldots,\alpha_i} & \cdots & \mu_{i+j;\alpha_1,\ldots,\alpha_i} \\ \mu_{i+1;\alpha_1,\ldots,\alpha_i} & \mu_{i+2;\alpha_1,\ldots,\alpha_i} & \mu_{i+3;\alpha_1,\ldots,\alpha_i} & \cdots & \mu_{i+j+1;\alpha_1,\ldots,\alpha_i} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{i+j-1;\alpha_1,\ldots,\alpha_i} & \mu_{i+j;\alpha_1,\ldots,\alpha_i} & \mu_{i+j+1;\alpha_1,\ldots,\alpha_i} & \cdots & \mu_{i+2j-1;\alpha_1,\ldots,\alpha_i} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ r_1 \prod_{k=1}^{i} (\theta_1 - \alpha_k) & r_2 \prod_{k=1}^{i} (\theta_2 - \alpha_k) & \cdots & r_j \prod_{k=1}^{i} (\theta_j - \alpha_k) & 0 \\ r_1\theta_1 \prod_{k=1}^{i} (\theta_1 - \alpha_k) & r_2\theta_2 \prod_{k=1}^{i} (\theta_2 - \alpha_k) & \cdots & r_j\theta_j \prod_{k=1}^{i} (\theta_j - \alpha_k) & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ r_1\theta_1^{j-1} \prod_{k=1}^{i} (\theta_1 - \alpha_k) & r_2\theta_2^{j-1} \prod_{k=1}^{i} (\theta_2 - \alpha_k) & \cdots & r_j\theta_j^{j-1} \prod_{k=1}^{i} (\theta_j - \alpha_k) & 0 \end{pmatrix} \times \begin{pmatrix} 1 & \theta_1 & \theta_1^2 & \cdots & \theta_1^j \\ 1 & \theta_2 & \theta_2^2 & \cdots & \theta_2^j \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \theta_j & \theta_j^2 & \cdots & \theta_j^j \\ 1 & x & x^2 & \cdots & x^j \end{pmatrix}.$$

Second, the probability masses $\mathrm{P}[Y = \alpha_\ell]$ $(\ell = 1, 2, ..., i)$ and $\mathrm{P}[Y = \theta_\ell]$ $(\ell = 1, 2, ..., j)$ can be obtained from solving the Vandermonde system of equations

$$\sum_{l=1}^{i} \mathrm{P}[Y = \alpha_\ell]\alpha_\ell^k + \sum_{l=1}^{j} \mathrm{P}[Y = \theta_\ell]\theta_\ell^k = \mu_k, \qquad k = 0, 1, \ldots, i + j - 1 \tag{2.3}$$

Using equations (2.2) and (2.3) we are able to construct some particular distributions that have a minimum number of support points as follows (see also Denuit et al. (1998) and Denuit et al. (1999)).

Even number of moments fixed (s is odd, s=2m+1): We consider two particular discrete random variables in $\mathcal{M}_{2m+1}([a, b]; \mu_1, \mu_2, \ldots, \mu_{2m})$ having $(m + 1)$ atoms with one of the atoms given as either a or b. The first one puts positive masses $p_1, p_2, \ldots, p_{m+1}$ on the atoms $a, x_2, \ldots, x_{m+1}$, with $a < x_2 < \cdots < x_{m+1} < b$. Here, the $x_2, \ldots, x_{m+1}$ are the s distinct roots of equation (2.2) with $i = 1$, $\alpha_1 = a$ and $j = m$. From Property 2.1 we know that the corresponding distribution has at least $2m - 1$ crossing points with any other distribution function of a variable in $\mathcal{M}_{2m+1}([a, b]; \mu_1, \mu_2, \ldots, \mu_{2m})$, and as per construction it must then have exactly $2m$ such crossings. This random variable is the smallest for the $(2m + 1)$-convex order among all random variables in $\mathcal{M}_{2m+1}([a, b]; \mu_1, \mu_2, \ldots, \mu_{2m})$. We denote it by $X_{min}^{(2m+1)}$. The second one puts positive masses $q_1, q_2, \ldots, q_{m+1}$ on the atoms $y_1, y_2, \ldots, y_m$, b, with $a < y_1 < y_2 < \cdots < y_m < b$, where $y_1, y_2, \ldots, y_m$ are the m distinct roots of (2.2) with $i = 1$, $\alpha_1 = b$ and $j = m$. From Property 2.1 it readily follows that this random variable is the largest in the sense of the $(2m + 1)$-convex order among all random variables in $\mathcal{M}_{2m+1}([a, b]; \mu_1, \mu_2, \ldots, \mu_{2m})$. We denote it by $X_{max}^{(2m+1)}$.

Odd number of moments fixed (s is even, s=2m): We define two particular elements in $\mathcal{M}_{2m}([a, b]; \mu_1, \mu_2, \ldots, \mu_{2m-1})$. The first one has the minimal number of support points, i.e. m support points strictly in (a, b). It puts positive masses $p_1, p_2, \ldots, p_m$ on the atoms $x_1, x_2, \ldots, x_m$, with $a < x_1 < x_2 < \cdots < x_m < b$. The atoms $x_1, x_2, \ldots, x_m$ are the m distinct roots of equation (2.2) with $i = 0$ and $j = m$. From Property 2.1 it follows that this random variable is the smallest in the sense of the $2m$-convex order among all random variables in $\mathcal{M}_{2m}([a, b]; \mu_1, \mu_2, \ldots, \mu_{2m-1})$. We denote it by $X_{min}^{(2m)}$. The second element in $\mathcal{M}_{2m}([a, b]; \mu_1, \mu_2, \ldots, \mu_{2m-1})$ of interest has $m+1$ support points: a, b and $m-1$ strictly in (a, b). It puts positive masses $q_1, q_2, \ldots, q_{m+1}$ on the atoms $a, y_2, \ldots, y_m, b$, with $a < y_2 < \cdots < y_m < b$. The atoms $y_2, \ldots, y_m$ are the $m - 1$ distinct roots of equation (2.2) with $i = 2$, $\alpha_1 = a$, $\alpha_2 = b$ and $j = m - 1$. Finally, Property 2.1 implies that this random variable is the largest in the sense of the $2m$-convex order among all the random variables in $\mathcal{M}_{2m}([a,b];\mu_1,\mu_2,\ldots,\mu_{2m-1})$. It will be denoted below by $X_{max}^{(2m)}$.

We end this sub-section with two properties that are used repeatedly in the remainder of this paper. From Property 2.1 and the structure of $X_{min}^{(s)}$ and $X_{max}^{(s)}$ it follows that for any $X \in \mathcal{M}_s([a,b];\mu_1,\mu_2,\ldots,\mu_{s-1})$ we must have

$$S^-(F_X - F_{X_{min}^{(s)}}) = S^-(F_X - F_{X_{max}^{(s)}}) = s-1 \tag{2.4}$$

which brings a set of conditions on the position of the graph of F_X. This argument is used repeatedly in the remainder of this paper (and was at the core of the construction of $X_{min}^{(s)}$ and $X_{max}^{(s)}$ that we outlined before). Also, as $X_{min}^{(s)}$ and $X_{max}^{(s)}$ both belong to $\mathcal{M}_s([a,b];\mu_1,\mu_2,\ldots,\mu_{s-1})$, we must have

$$S^-(F_{X_{min}^{(s)}} - F_{X_{max}^{(s)}}) = s-1 \tag{2.5}$$

so that the support points of $X_{min}^{(s)}$ and $X_{max}^{(s)}$ alternate in $[a,b]$. See (3.1) and (3.2) below.

3 Moment bounds on distribution functions

For any $t \in (a,b)$, we want to find a random variable $Z \in \mathcal{M}_s([a,b];\mu_1,\mu_2,\ldots,\mu_{s-1})$ with distribution function F_Z such that for all $X \in \mathcal{M}_s([a,b];\mu_1,\mu_2,\ldots,\mu_{s-1})$. It holds that,

$$F_Z(t-) = \mathrm{P}[Z < t] \leqslant F_X(t) \leqslant F_Z(t).$$

Notice that the random variable Z depends on the point t considered and that the bounds on $F_X(t)$ are sharp (best-possible). The random variable Z providing bounds on $F_X(t)$ can now be built as follows, depending on the number of moments known.

Odd number of moments fixed (s is even, s=2m): Assume that s is even, i.e. $s = 2m$ for some positive integer m. Consider a discrete variable Z in $\mathcal{M}_s([a,b];\mu_1,\mu_2,\ldots,\mu_{s-1})$ having atoms $z_1,\ldots,z_{m+1} \in [a,b]$, $z_1 < \ldots < z_{m+1}$. For any other random variable X in $\mathcal{M}_s([a,b];\mu_1,\mu_2,\ldots,\mu_{s-1})$, $X \neq X_{\min}^{(s)}$, $X \neq X_{\max}^{(s)}$, we know from Property 2.1 that $S^-(F_X - F_Z) \geqslant s-1$. By selecting appropriate locations for the atoms of Z, we can then ensure that $S^-(F_X - F_Z) = s$ always hold and we show that such Z allows to bound $F_X(t)$.

For $S^-(F_X - F_Z) = s$ to hold, we must have that either $z_1 = a$ or $z_{m+1} = b$ and the other m support points must be taken in (a,b). More precisely, if one has $z_1 = a < z_2 < \ldots < z_{m+1} < b$ then we have a total of $2m = s$ crossings indeed, namely

- one crossing over each flat part above (z_j, z_{j+1}), $j = 1,\ldots,m$ (i.e. a total of m crossings);
- one crossing at each atom $z_2,\ldots,z_{m+1}$ (i.e. a total of m crossings).

Next, if $a < z_1 < \ldots < z_m < z_{m+1} = b$ then we also have $2m = s$ crossings, which can be described as follows:

- one crossing over each flat part above (z_j, z_{j+1}), $j = 1, \ldots, m$ (i.e. a total of m crossings);
- one crossing at each atom $z_1, \ldots, z_m$ (i.e. a total of m crossings).

This means that each time F_Z jumps, i.e. at each $z_j \in (a, b)$, we must have

$$F_Z(z_j-) < F_X(z_j) < F_Z(z_j).$$

Now, as one support point of Z is given (a or b), we have $2m + 1 = s + 1$ unknown support points and probability masses. But we face s constraints ($s - 1$ moments and the probabilities summing up to 1). Thus, there remains a single degree of freedom. We are therefore free to locate one support point of Z at t in order to bound $F_X(t)$. Hence, we only have to build the unique random variable Z having the structure described above in which we take t as a particular atom.

In the reasoning that we just outlined, it appears that there are two possibilities for choosing Z. We now explain that, depending on the location of t with respect to the support points of the extrema $X^{(s)}_{\min}$ and $X^{(s)}_{\max}$, one can determine whether either a or b is to be considered as a support point of Z. Indeed, when $s = 2m$, the support points of $X^{(s)}_{\min}$ are $a < x_1 < x_2 < \ldots < x_m < b$ and the support points of $X^{(s)}_{\max}$ are $a = y_1 < y_2 < \ldots < y_m < y_{m+1} = b$. The x_j and y_j satisfy

$$a = y_1 < x_1 < y_2 < x_2 < \ldots < y_m < x_m < y_{m+1} = b. \tag{3.1}$$

At each atom z_j of Z, the distribution function F_Z must cross both $F_{X^{(s)}_{\min}}$ and $F_{X^{(s)}_{\max}}$ (see also (2.4)). This implies that,

- if $t \in (y_j, x_j)$ then each atom of Z must belong to an interval (y_ℓ, x_ℓ), $\ell = 1, \ldots, m$ and b must belong to the support of Z. The support of Z thus consists in

  $$a < z_1 < z_2 < \ldots < z_m < z_{m+1} = b$$

  with each $z_k \in (y_k, x_k)$ and $z_j = t$.
- if $t \in (x_j, y_{j+1})$ then each atom of Z must belong to an interval $(x_\ell, y_{\ell+1})$, $\ell = 1, \ldots, m$ and a must belong to the support of Z. The support of Z thus consists in

  $$a = z_1 < z_2 < \ldots < z_{m+1} < b$$

  with each $z_k \in (x_k, y_{k+1})$ and $z_j = t$.

Hence, the location of t with respect to the partition of $[a, b]$ created by the union of the supports of $X_{\min}^{(s)}$ and $X_{\max}^{(s)}$ satisfying (3.1) determines whether a or b belongs to the support of Z.

From the above reasoning we conclude that the support points of the optimal Z (allowing to bound $F_X(t)$) can thus be found by solving equation (2.2) with $i = 2$ and either $\alpha_1 = a$ and $\alpha_2 = t$ (when $t \in (x_j, y_{j+1})$) or $\alpha_1 = t$ and $\alpha_2 = b$ (when $t \in (y_j, x_j)$).

We illustrate this construction and the reasoning with a detailed worked-out example.

Example 3.1 (Case $s = 2$, $m = 1$). In $\mathcal{M}_2([a, b]; \mu_1)$ one has that $X_{min}^{(2)} = \mu_1$ almost surely, and $X_{max}^{(2)}$ is given as

$$X_{max}^{(2)} = \begin{cases} a & \text{with probability } \frac{b-\mu_1}{b-a}, \\ b & \text{with probability } \frac{\mu_1-a}{b-a}. \end{cases}$$

As $X_{min}^{(2)}$ belongs to $\mathcal{M}_2([a, b]; \mu_1)$, we know that the lower bound on $F_X(t)$ can only be zero before μ_1 and the upper bound can only be 1 after μ_1. The support of Z must contain t and either a or b. If $a < t < \mu_1$ then a cannot be a support point because a and t both smaller than μ_1 is impossible so that the support must be of the form $\{t, b\}$. Another argument is as follows: for such a t, we have $\mathrm{P}[Z < t] = 0$ so that a cannot belong to the support of Z. We then get

$$Z = \begin{cases} t & \text{with probability } \frac{b-\mu_1}{b-t}, \\ b & \text{with probability } \frac{\mu_1-t}{b-t}. \end{cases}$$

Notice that Z tends to $X_{max}^{(2)}$ as $t \to a$. If $\mu_1 < t < b$ then b cannot be a support point because b and t both larger than μ_1 is impossible so that the support must be of the form $\{a, t\}$. Another argument is as follows: for such a t, we have $\mathrm{P}[Z \leqslant t] = 1$ so that b cannot belong to the support of Z. We then get

$$Z = \begin{cases} a & \text{with probability } \frac{t-\mu_1}{t-a}, \\ t & \text{with probability } \frac{\mu_1-a}{t-a}. \end{cases}$$

Notice that Z tends to $X_{max}^{(2)}$ as $t \to b$ and also observe that the discussion on bounds for $F_X(t)$ depend on the localisation of t with respect to

$$\mathrm{supp}(X_{min}^{(2)}) \cup \mathrm{supp}(X_{max}^{(2)}) = \{a, \mu_1, b\}.$$

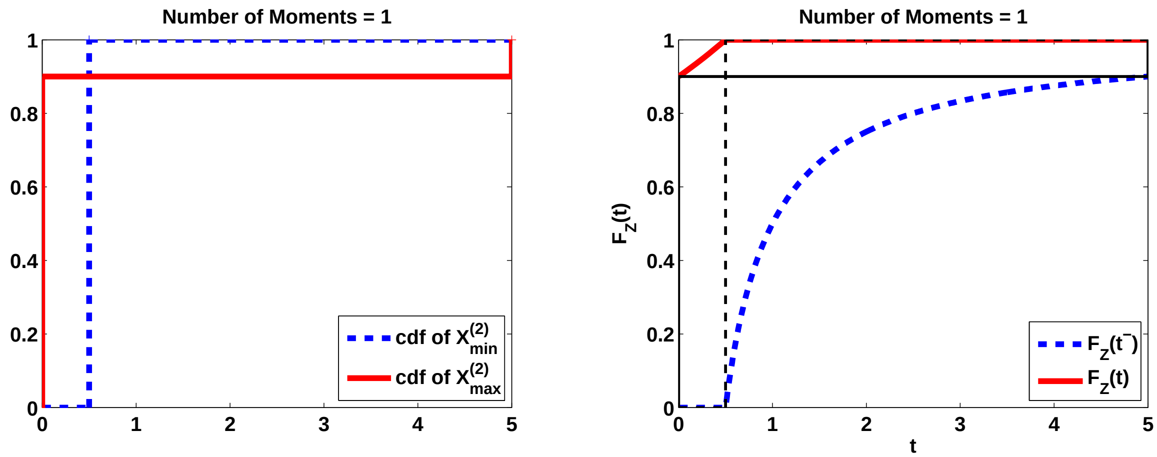
The final result we get is as follows:

- if $a < t < \mu_1$:

$$0 = \mathrm{P}[Z < t] \leqslant F_X(t) \leqslant \mathrm{P}[Z \leqslant t] = \mathrm{P}[Z = t] = \frac{b - \mu_1}{b - t}.$$

- if $\mu_1 < t < b$:

$$\frac{t - \mu_1}{t - a} = \mathrm{P}[Z = a] = \mathrm{P}[Z < t] \leqslant F_X(t) \leqslant \mathrm{P}[Z \leqslant t] = 1.$$



Panel A

Panel B

Figure 3.1: Case $s = 2$, $m = 1$: The left panel displays the extreme cdfs of $X^{(2)}_{min}$ and $X^{(2)}_{max}$ and the right panel draws $F_Z(t)$ and $F_Z(t^-) = \mathrm{P}[Z < t]$ on top of these extreme cdfs with $\mu_1 = \frac{1}{2}$, $a = 0$ and $b = 5$. Note that we omitted the indice in the notation of the random variable Z but for each t, the random variable Z depends on t. The curves represented in Panel B and denoted by abuse of notation $F_Z(t)$ and $F_Z(t^-)$ are $F_{Z_t}(t)$ and $F_{Z_t}(t^-)$. The lower and upper curves in Panel B should be interpreted as bounds on the $F_Z(t)$ when two moments are fixed.

Even number of moments fixed: Let us now assume that s is odd, i.e. $s = 2m + 1$ for some integer m. In this case, we consider a discrete variable Z in $\mathcal{M}_s([a, b]; \mu_1, \mu_2, \ldots, \mu_{s-1})$ with support that is either of the form $\{z_1, \ldots, z_{m+2}\}$ with $a = z_1 < \ldots < z_{m+2} = b$, or $\{z_1, \ldots, z_{m+1}\}$ with $a < z_1 < \ldots < z_{m+1} < b$. Indeed, if the support is $\{z_1, \ldots, z_{m+2}\}$ with $z_1 = a$ and $z_{m+2} = b$, it readily shows that one has a total of $2m + 1 = s$ crossings with any other element X in $\mathcal{M}_s([a, b]; \mu_1, \mu_2, \ldots, \mu_{s-1})$ ($X \neq X^{(s)}_{\min}$, $X \neq X^{(s)}_{\max}$):

- one crossing over each flat part above (z_j, z_{j+1}), $j = 1, \ldots, m + 1$ (i.e. a total of $m + 1$ crossings);
- one crossing at each atom $z_2, \ldots, z_{m+1}$ (i.e. a total of m crossings).

If the support is $\{z_1, \ldots, z_{m+1}\}$ with $a < z_1 < \ldots < z_{m+1} < b$, we also have a total of $2m + 1 = s$ crossings:

- one crossing over each flat part above (z_j, z_{j+1}), $j = 1, \ldots, m$ (i.e. a total of m crossings);
- one crossing at each atom $z_1, \ldots, z_{m+1}$ (i.e. a total of $m + 1$ crossings).

As before, depending on the location of t with respect to the support points of the s-convex extrema $X^{(s)}_{\min}$ and $X^{(s)}_{\max}$, there is only one possible Z, including either a and b as support points or none of them. To explain this further let us remark that the support points of $X^{(s)}_{\min}$ are $a = x_1 < x_2 < \ldots < x_{m+1} < b$ and the support points of $X^{(s)}_{\max}$ are $a < y_1 < y_2 < \ldots < y_m < y_{m+1} = b$. The x_j and y_j satisfy

$$a = x_1 < y_1 < x_2 < y_2 < \ldots < y_m < x_{m+1} < y_{m+1} = b. \tag{3.2}$$

At each atom z_j of Z, F_Z must cross both $F_{X^{(s)}_{\min}}$ and $F_{X^{(s)}_{\max}}$ so that two cases are possible:

- if $t \in (x_j, y_j)$ for some $j = 1, \ldots, m+1$ then each atom of Z must belong to an interval (x_ℓ, y_ℓ), $\ell = 1, \ldots, m+1$ and no probability mass is placed on the extremities a and b. The support of Z thus consists of

  $$a < z_1 < z_2 < \ldots < z_m < z_{m+1} < b$$

  with each $z_k \in (x_k, y_k)$ and $z_j = t$.

- if $t \in (y_j, x_{j+1})$ for some $j = 1, \ldots, m$ then each atom of Z must belong to an interval $(y_\ell, x_{\ell+1})$, $\ell = 1, \ldots, m$ and a and b must both belong to the support of Z. The support of Z thus consists in

  $$a = z_1 < z_2 < \ldots < z_{m+1} < z_{m+2} = b$$

  with each $z_{k+1} \in (y_k, x_{k+1})$ for $k = 1, \ldots, m$ and $z_{j+1} = t$.

Thus, we see that the location of t with respect to the partition of $[a, b]$ created by the union of the supports of $X^{(s)}_{\min}$ and $X^{(s)}_{\max}$ satisfying (3.2) determines whether both a and b belong to the support of Z or none of them. The support of Z is then obtained by solving equation (2.2) with either $\alpha_1 = t$ and $i = 1$ or $\alpha_1 = a$, $\alpha_2 = t$, $\alpha_3 = b$ and $i = 3$.

We illustrate this further with a detailed example.

Example 3.2 (Case $s = 3$)**.** In $\mathcal{M}_3([a, b]; \mu_1, \mu_2)$ we find that,

$$X^{(3)}_{min} = \begin{cases} a & \text{with probability } \frac{\mu_2-\mu_1^2}{(a-\mu_1)^2+\mu_2-\mu_1^2}, \\ \mu_1 + \frac{\mu_2-\mu_1^2}{\mu_1-a} & \text{with probability } \frac{(a-\mu_1)^2}{(a-\mu_1)^2+\mu_2-\mu_1^2}, \end{cases} \tag{3.3}$$

and

$$X^{(3)}_{max} = \begin{cases} \mu_1 - \frac{\mu_2-\mu_1^2}{b-\mu_1} & \text{with probability } \frac{(b-\mu_1)^2}{(b-\mu_1)^2+\mu_2-\mu_1^2}, \\ b & \text{with probability } \frac{\mu_2-\mu_1^2}{(b-\mu_1)^2+\mu_2-\mu_1^2}. \end{cases} \tag{3.4}$$

As both $X^{(3)}_{min}$ and $X^{(3)}_{max}$ belong to $\mathcal{M}_3([a, b]; \mu_1, \mu_2)$, the lower bound on $F_X(t)$ can only be zero before $\mu_1 - \frac{\mu_2-\mu_1^2}{b-\mu_1}$ because $F_{X^{(3)}_{max}}$ vanishes there and that the upper bound can only be 1 after $\mu_1 + \frac{\mu_2-\mu_1^2}{\mu_1-a}$ because $F_{X^{(3)}_{min}} = 1$ over that interval. The support of Z must be either $\{a, t, b\}$ or $\{t, t'\}$ where t and $t' \in (a, b)$. If $a < t < \mu_1 - \frac{\mu_2-\mu_1^2}{b-\mu_1}$ then a cannot be a support point because $F_Z(t-) = 0$ there. Therefore, the support must be of the form $\{t, t'\}$. From Lemma 1 in Jansen et al. (1986), we know that

$$\begin{array}{l} t' = \mu_1 + \frac{\mu_2-\mu_1^2}{\mu_1-t} \in (a', b) := \left(\mu_1 + \frac{\mu_2-\mu_1^2}{\mu_1-a}, b\right) \\ \mathrm{P}[Z = t] = \frac{t'-\mu_1}{t'-t}, \quad \mathrm{P}[Z = t'] = \frac{\mu_1-t}{t'-t}. \end{array}$$

Hence,

$$Z = \begin{cases} t & \text{with probability } \frac{t'-\mu_1}{t'-t}, \\ t' & \text{with probability } \frac{\mu_1-t}{t'-t}. \end{cases}$$

If $\mu_1 - \frac{\mu_2-\mu_1^2}{b-\mu_1} < t < \mu_1 + \frac{\mu_2-\mu_1^2}{\mu_1-a}$, i.e. $b' < t < a'$ then

$$\begin{array}{l} \mathrm{P}[Z < t] \geqslant F_{X_{min}^{(3)}}(t) > 0 \\ \mathrm{P}[Z > t] = 1 - F_Z(t) > 1 - F_{X_{max}^{(3)}}(t) > 0 \end{array}$$

so that there must be at least two other support points, one to the left of t and another one to the right of t. Thus, the support must be of the form $\{a, t, b\}$. The corresponding probability masses can be found from

$$\begin{aligned} \mathrm{P}[Z = a] &= \frac{E[(Z-t)(Z-b)]}{(a-t)(a-b)} = \frac{\mu_2 - (t+b)\mu_1 + bt}{(a-t)(a-b)}, \\ \mathrm{P}[Z = t] &= \frac{E[(Z-a)(Z-b)]}{(t-a)(t-b)} = \frac{\mu_2 - (a+b)\mu_1 + ab}{(t-a)(t-b)}, \\ \mathrm{P}[Z = b] &= \frac{E[(Z-a)(Z-t)]}{(b-a)(b-t)} = \frac{\mu_2 - (a+t)\mu_1 + at}{(b-a)(b-t)}. \end{aligned} \tag{3.5}$$

If $\mu_1 + \frac{\mu_2-\mu_1^2}{\mu_1-a} < t < b$, i.e. $a' < t < b$ then b cannot be a support point because $F_Z(t) = 1$ here. Hence, the support must consist of two interior points of $[a, b]$ and

$$Z = \begin{cases} t' & \text{with probability } \frac{t-\mu_1}{t-t'}, \\ t & \text{with probability } \frac{\mu_1-t'}{t-t'}. \end{cases}$$

We observe again that the discussion on t is with respect to

$$\mathrm{supp}(X_{min}^{(3)}) \cup \mathrm{supp}(X_{max}^{(3)}) = \left\{a, \mu_1 - \frac{\mu_2-\mu_1^2}{b-\mu_1}, \mu_1 + \frac{\mu_2-\mu_1^2}{\mu_1-a}, b\right\}.$$

The final result we get is as follows:

- if $a < t < \mu_1 - \frac{\mu_2-\mu_1^2}{b-\mu_1}$:

$$0 = \mathrm{P}[Z < t] \leqslant F_X(t) \leqslant \mathrm{P}[Z \leqslant t] = \frac{t'-\mu_1}{t'-t} = \frac{\mu_2-\mu_1^2}{(\mu_1-t)^2 + \mu_2 - \mu_1^2}.$$

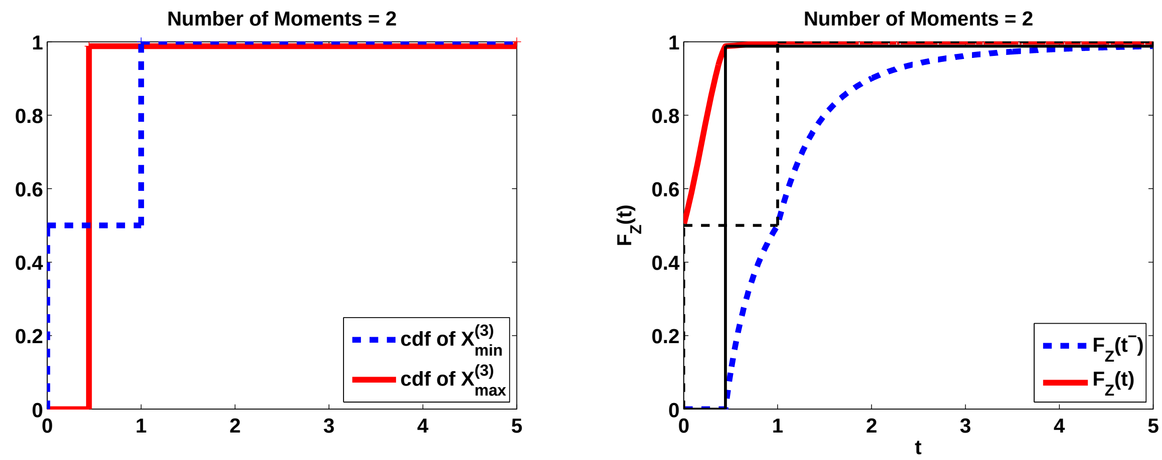
- if $\mu_1 - \frac{\mu_2-\mu_1^2}{b-\mu_1} < t < \mu_1 + \frac{\mu_2-\mu_1^2}{\mu_1-a}$:

$$\mathrm{P}[Z=a] = \mathrm{P}[Z<t] \leqslant F_X(t) \leqslant \mathrm{P}[Z \leqslant t] = \mathrm{P}[Z=b]$$

  where $\mathrm{P}[Z=a]$ and $\mathrm{P}[Z=b]$ are given by (3.5).

- if $\mu_1 + \frac{\mu_2-\mu_1^2}{\mu_1-a} < t < b$:

$$\frac{(x-\mu_1)^2}{(t-\mu_1)^2+\mu_2-\mu_1^2} = \frac{t-\mu_1}{t-t'} = \mathrm{P}[Z=t'] = \mathrm{P}[Z<t] \leqslant F_X(t) \leqslant \mathrm{P}[Z \leqslant t] = 1.$$



Panel A

Panel B

Figure 3.2: Case $s=3$: the left panel represents the cdfs of $X^{(3)}_{min}$ and $X^{(3)}_{max}$ and the right panel displays $F_Z(t)$ and $F_Z(t^-) = \mathrm{P}[Z<t]$ with $\mu_1 = \frac{1}{2} = \mu_2$, $a=0$ and $b=5$. Note that we omitted the indice in the notation of the random variable Z but for each t, the random variable Z depends on t. The curves represented in Panel B and denoted by abuse of notation $F_Z(t)$ and $F_Z(t^-)$ are $F_{Z_t}(t)$ and $F_{Z_t}(t^-)$. The lower and upper curves in Panel B should be interpreted as bounds on $F_X(t)$ when two moments are fixed.

4 Moment bounds on VaR

4.1 Derivation

Our aim is now to derive bounds on the Value-at-Risk (VaR) of X at level p,

$$VaR_p[X] := F_X^{-1}(p) = \inf\{x \in \mathbb{R} | F_X(x) \geqslant p\}.$$

The approach to derive VaR bounds is closely related to the one we described for finding bounds on its distribution. The only adjustment we need to make is that we now impose

$F_Z(z_j) = p$ for some atom z_j of Z. Any distribution function in the moment space must then cross the plateau at probability level p and hence $[F_Z^{-1}(p); F_Z^{-1+}(p)]$ provides the range of admissible values for $VaR_p[X]$ where F_Z^{-1+} is defined as

$$F_Z^{-1+}(p) = \sup\{x \in \mathbb{R} | F_Z(x) \leqslant p\}.$$

More precisely, the position of the level p in the union of the ranges of the distribution functions of the s-convex extrema $X_{\min}^{(s)}$ and $X_{\max}^{(s)}$ indicates the atom at which F_Z must reach level p (to ensure that we indeed have s crossings). We explain this further as follows.

Let us first assume that an even number of moments are known. Thus, we assume that s is odd, i.e. $s = 2m+1$. Then, the support points of $X_{\min}^{(s)}$ are $a = x_1 < x_2 < \ldots < x_{m+1} < b$ and the support points of $X_{\max}^{(s)}$ are $a < y_1 < y_2 < \ldots < y_m < y_{m+1} = b$. The x_j and y_j satisfy (3.2). We distinguish the two following cases:

- if $p \in \left(F_{X_{\min}^{(s)}}(x_j), F_{X_{\max}^{(s)}}(y_j)\right)$, i.e. $p \in \left(\sum_{k=1}^{j} p_k, \sum_{k=1}^{j} q_k\right)$ then F_Z must plateau at level p from before $F^{-1}_{X_{\max}^{(s)}}(p) = y_j$ to after $F^{-1}_{X_{\min}^{(s)}}(p) = x_{j+1}$. This means that Z has no atom in (y_j, x_{j+1}) so that each of the $m+1$ support points of Z belongs to an interval (x_ℓ, y_ℓ), $\ell = 1, \ldots, m+1$ and a and b do not belong to the support of Z. Level p is reached by F_Z at the atom located in (x_j, y_j), that is, $F_Z(z_j) = p$ and

  $$z_j \leqslant VaR_p[X] \leqslant z_{j+1} \text{ for any } X \in \mathcal{M}_s([a,b]; \mu_1, \mu_2, \ldots, \mu_{s-1}).$$

  This also provides the following bounds on $VaR_p[X]$ in terms of support points of the s-convex extrema:

  $$x_j \leqslant VaR_p[X] \leqslant y_{j+1} \text{ for any } X \in \mathcal{M}_s([a,b]; \mu_1, \mu_2, \ldots, \mu_{s-1}).$$

- if $p \in \left(F_{X_{\max}^{(s)}}(y_j), F_{X_{\min}^{(s)}}(x_{j+1})\right)$, i.e. $p \in \left(\sum_{k=1}^{j} q_k, \sum_{k=1}^{j+1} p_k\right)$ then F_Z must plateau at level p from before $F^{-1}_{X_{\min}^{(s)}}(p) = x_{j+1}$ to after $F^{-1}_{X_{\max}^{(s)}}(p) = y_{j+1}$. This means that Z has no atom in (x_{j+1}, y_{j+1}) and the support points of Z must be a, b and (unique) elements in each interval $(y_\ell, x_{\ell+1})$, $\ell = 1, \ldots, m$. Level p is reached by F_Z at the atom located in (y_j, x_{j+1}), that is, $F_Z(z_{j+1}) = p$ and

  $$z_{j+1} \leqslant VaR_p[X] \leqslant z_{j+2} \text{ for any } X \in \mathcal{M}_s([a,b]; \mu_1, \mu_2, \ldots, \mu_{s-1}).$$

  This also provides the following bounds on $VaR_p[X]$ in terms of support points of the s-convex extrema:

  $$y_j \leqslant VaR_p[X] \leqslant x_{j+2} \text{ for any } X \in \mathcal{M}_s([a,b]; \mu_1, \mu_2, \ldots, \mu_{s-1}).$$

Next we consider the situation in which an odd number of moments are known. Hence, $s = 2m$. Then, the support points of $X_{\min}^{(s)}$ are $a < x_1 < x_2 < \ldots < x_m < b$ and the support points of $X_{\max}^{(s)}$ are $a = y_1 < y_2 < \ldots < y_m < y_{m+1} = b$. The x_j and y_j satisfy (3.1). The following two cases can be identified:

- if $p \in \left(F_{X_{\max}^{(s)}}(y_j), F_{X_{\min}^{(s)}}(x_j)\right)$, i.e. $p \in \left(\sum_{k=1}^{j} q_k, \sum_{k=1}^{j} p_k\right)$ then then F_Z must plateau at level p from before $F^{-1}_{X_{\min}^{(s)}}(p) = x_j$ to after $F^{-1}_{X_{\max}^{(s)}}(p) = y_{j+1}$. Hence, Z has no atom in (x_j, y_{j+1}) so that the support points of Z must include b, and (unique) elements in each interval (y_ℓ, x_ℓ), $\ell = 1, \ldots, m$. Level p is reached by F_Z at the atom located in (y_j, x_j), that is, $F_Z(z_j) = p$ and

$$z_j \leqslant VaR_p[X] \leqslant z_{j+1} \text{ for any } X \in \mathcal{M}_s([a, b]; \mu_1, \mu_2, \ldots, \mu_{s-1}).$$

This also provides the following bounds on $VaR_p[X]$ in terms of support points of the s-convex extrema:

$$y_j \leqslant VaR_p[X] \leqslant x_{j+1} \text{ for any } X \in \mathcal{M}_s([a, b]; \mu_1, \mu_2, \ldots, \mu_{s-1}).$$

- if $p \in \left(F_{X_{\min}^{(s)}}(x_j), F_{X_{\max}^{(s)}}(y_{j+1})\right)$, i.e. $p \in \left(\sum_{k=1}^{j} p_k, \sum_{k=1}^{j+1} q_k\right)$ then F_Z must plateau at level p from before $F^{-1}_{X_{\max}^{(s)}}(p) = y_{j+1}$ to after $F^{-1}_{X_{\min}^{(s)}}(p) = x_{j+1}$. This means that Z has no atom in (y_{j+1}, x_{j+1}) and the support points of Z must be a and (unique) elements in each interval $(x_\ell, y_{\ell+1})$, $\ell = 1, \ldots, m$. Level p is reached by F_Z at the atom located in (x_j, y_{j+1}), that is, $F_Z(z_j) = p$ and

$$z_j \leqslant VaR_p[X] \leqslant z_{j+1} \text{ for any } X \in \mathcal{M}_s([a, b]; \mu_1, \mu_2, \ldots, \mu_{s-1}).$$

This also provides the following bounds on $VaR_p[X]$ in terms of support points of the s-convex extrema:

$$x_j \leqslant VaR_p[X] \leqslant y_{j+2} \text{ for any } X \in \mathcal{M}_s([a, b]; \mu_1, \mu_2, \ldots, \mu_{s-1}).$$

We illustrate the construction with an example.

Example 4.1 (Case $s = 3$)**.** The discussion is with respect to the probability levels

$$\text{range}(F_{X^{(3)}_{min}}) \cup \text{range}(F_{X^{(3)}_{max}}) = \left\{0, \frac{\mu_2 - \mu_1^2}{(a - \mu_1)^2 + \mu_2 - \mu_1^2}, \frac{(b - \mu_1)^2}{(b - \mu_1)^2 + \mu_2 - \mu_1^2}, 1\right\}.$$

- For $p < \frac{\mu_2 - \mu_1^2}{(a-\mu_1)^2 + \mu_2 - \mu_1^2}$ we select Z with support $\{a, t, b\}$ and we impose $\mathrm{P}[Z = a] = p$ so that

$$p = \frac{\mu_2 - (t + b)\mu_1 + bt}{(a - t)(a - b)} \Leftrightarrow t = \frac{pa(a - b) + b\mu_1 - \mu_2}{b - \mu_1 + p(a - b)}.$$

The bounds are then given by

$$VaR_p[X^{(3)}_{min}] = a = VaR_p[Z] \leqslant VaR_p[X] \leqslant VaR^+_p[Z] = t = \frac{pa(a - b) + b\mu_1 - \mu_2}{b - \mu_1 + p(a - b)}.$$

- For $\frac{\mu_2-\mu_1^2}{(a-\mu_1)^2+\mu_2-\mu_1^2} < p < \frac{(b-\mu_1)^2}{(b-\mu_1)^2+\mu_2-\mu_1^2}$ we select Z with support $\{t,t'\}$ in (a,b) such that $\mathrm{P}[Z=t]=p$ so that

$$p = \frac{t'-\mu_1}{t'-t} = \frac{\frac{\mu_2-\mu_1^2}{\mu_1-t}}{\mu_1-t+\frac{\mu_2-\mu_1^2}{\mu_1-t}} \Leftrightarrow (\mu_1-t)^2 = \frac{1-p}{p}(\mu_2-\mu_1^2) \Leftrightarrow t = \mu_1-\sqrt{\frac{1-p}{p}(\mu_2-\mu_1^2)}$$

as $t<\mu_1<t'$ must hold. The other support point of Z is then given by

$$t' = \mu_1 + \frac{\mu_2-\mu_1^2}{\sqrt{\frac{1-p}{p}(\mu_2-\mu_1^2)}} = \mu_1 + \sqrt{\frac{p}{1-p}(\mu_2-\mu_1^2)}.$$

The bounds on $VaR_p[X]$ are then given by

$$\mu_1 - \sqrt{\frac{1-p}{p}(\mu_2-\mu_1^2)} = t = VaR_p[Z] \leqslant VaR_p[X]$$
$$\leqslant VaR_p^+[Z] = t' = \mu_1 + \sqrt{\frac{p}{1-p}(\mu_2-\mu_1^2)}.$$

- For $p > \frac{(b-\mu_1)^2}{(b-\mu_1)^2+\mu_2-\mu_1^2}$ we consider Z with support $\{a,t,b\}$ with $\mathrm{P}[Z\leqslant t]=p \Leftrightarrow \mathrm{P}[Z=b]=1-p$ so that

$$1-p = \frac{\mu_2-(a+t)\mu_1+at}{(b-a)(b-t)} \Leftrightarrow t = \frac{(1-p)b(b-a)+a\mu_1-\mu_2}{a-\mu_1+(1-p)(b-a)}.$$

Then,

$$\frac{(1-p)b(b-a)+a\mu_1-\mu_2}{a-\mu_1+(1-p)(b-a)} = t = VaR_p[Z] \leqslant VaR_p[X] \leqslant VaR_p^+[Z] = b.$$

VaR bounds for an exponential like claim: Let us next compute VaR bounds for a variable that behaves like an exponential claim. Specifically, we consider random variables X on a bounded interval $[0,50]$ such that their first moments μ_i $(i=1,2,...,10)$ match those of an exponential distribution with parameter $\lambda=10$. Thus, $\mu_i=\frac{i!}{10^i}$ with $i=1,2,...,10$. Note that for the given exponential distribution, the probability to exceed 50 is equal to $1-e^{-500}$, i.e. we obtain that this event has zero probability at machine precision. In other words, the restriction to the bounded interval $[0,50]$ is not a real constraint. In Table 4.1 we provide the bounds on VaR at the different levels equal to 90%, 95% and 99%. When only two moments are given, then we can directly build on Example 4.1 to obtain closed-form expressions for the VaR bounds. Hence, taking $a=0, b=50$, $\mu_1=\frac{1}{10}$, $\mu_2=\frac{2}{100}$ we find in this case that

- If $p<\frac{1}{2}$,

$$0 < VaR_p(X) < \frac{498}{4990-5000p}.$$

- If $\frac{1}{2} \leqslant p \leqslant \frac{499^2}{499^2+1}$,

$$\frac{1}{10} - \frac{1}{10}\sqrt{\frac{1-p}{p}} \leqslant VaR_p(X) \leqslant \frac{1}{10} + \frac{1}{10}\sqrt{\frac{p}{1-p}}.$$

- If $p > \frac{499^2}{499^2+1}$,

$$\frac{100(1-p)50^2 - 2}{(1-p)5000 - 10} \leqslant VaR_p(X) \leqslant 50$$

# of moments	$VaR_{90\%}$	$VaR_{95\%}$	$VaR_{99\%}$
1	(0.005 ; 1.00)	(0.005 ; 2.00)	(0.005 ; 10.0)
2	(0.07 ; 0.40)	(0.08 ; 0.54)	(0.09 ; 1.09)
3	(0.09 ; 0.38)	(0.125 ; 0.46)	(0.16 ; 0.72)
4	(0.095 ; 0.37)	(0.135 ; 0.45)	(0.23 ; 0.64)
5	(0.10 ; 0.36)	(0.14 ; 0.44)	(0.24 ; 0.63)
6	(0.11 ; 0.35)	(0.16 ; 0.44)	(0.24 ; 0.62)
7	(0.12 ; 0.35)	(0.17 ; 0.43)	(0.27 ; 0.61)
8	(0.13 ; 0.33)	(0.17 ; 0.43)	(0.28 ; 0.60)
9	(0.13 ; 0.33)	(0.18 ; 0.42)	(0.29 ; 0.60)
10	(0.13 ; 0.33)	(0.19 ; 0.41)	(0.31 ; 0.59)

Table 4.1: We compute the minimum and maximum possible values of VaR for assuming that only the first series of moments μ_i of the loss are known ($a = 0$, $b = 50$). The VaR bounds are reported as (;) in which the first value is the lower bound and the second value is the upper bound. The VaR numbers that are obtained when assuming the loss is exponentially distributed (parameter $\lambda = 10$ and moments $\mu_i = \frac{i!}{\lambda^i}$) are given by $VaR_{90\%} = 0.2302$, $VaR_{95\%} = 0.2996$ and $VaR_{99\%} = 0.4605$.

We observe from Table 4.1 that there is still a lot of model risk on VaR assessments even when ten moments are known (without any statistical uncertainty). This observation holds for VaRs obtained at different probability levels as shown by comparing the width of the bounds for confidence levels at 90%, 95% and 99%.

4.2 Incorporating statistical uncertainty

In practice, the moments of the variable X at hand are not always precisely known ex-ante but are usually to be estimated from available data. For example, when X reflects the number of defaults of a credit loan portfolio, its $k-$th moment depends on the (joint) probabilities that k different loans default. While for low values of k ($k = 1, 2$) one may perhaps assume that these probabilities (and thus the corresponding moments) are known such assumption is clearly not realistic for larger values of k.

In this subsection, in order to capture the statistical uncertainty on the moments, we propose a robust approach in which we fix the first moment, but allow higher moments to take a value that is smaller than some maximum value μ_k ($k = 2, 3, ..., s-1$). It is clear that doing so is prudent in the sense that the bounds on statistical quantities will become wider as compared to the situation that we considered before (known moments). Hence, let us denote by $\mathfrak{U}_s([a,b]; \mu_1, \mu_2, \ldots, \mu_{s-1})$ the class of all the random variables X valued in the interval $[a,b]$ ($a, b \in \mathbb{R}$) with $E[X] = \mu_1$ and $E[X^k] \leqslant \mu_k$ ($k = 2, 3, \ldots, s-1$). In what follows, we restrict to non-negative random variables, i.e. $0 \leqslant a < b$. As $x \mapsto x^k$ is convex on $(0, \infty)$ it follows that $\mu_k \geqslant (\mu_1)^k$ ($k = 1, ..., s-1$) must hold. In particular, the variable $X = \mu_1$ a.s. belongs to $\mathfrak{U}_s([a,b]; \mu_1, \mu_2, \ldots, \mu_{s-1})$.

Consider now discrete random variables Z_α ($\alpha \geqslant 0$, $0 < p < 1$) of the form,

$$Z_\alpha = \begin{cases} \mu_1 - \alpha\frac{1-p}{p} & \text{with probability } p, \\ \mu_1 + \alpha & \text{with probability 1-}p. \end{cases} \tag{4.1}$$

Remark that $E[Z_\alpha] = \mu_1$ and Z_α takes values on $[a,b]$ whenever

$$0 \leqslant \alpha \leqslant \alpha^* := \min\left(b - \mu_1, (\mu_1 - a)\frac{p}{1-p}\right).$$

In particular, Z_0 belongs to $\mathfrak{U}_s([a,b]; \mu_1, \mu_2, \ldots, \mu_{s-1})$. Furthermore, for all $k = 1, ..., s-1$, one has that $\alpha \to E[Z_\alpha^k]$ is increasing on $[0, \alpha^*]$ and there exists a maximum value β such $Z_\beta \in \mathfrak{U}_s([a,b]; \mu_1, \mu_2, \ldots, \mu_{s-1})$. We claim now that

$$\mu_1 - \beta\frac{1-p}{p} \leqslant VaR_p[X] \leqslant \mu_1 + \beta \text{ for any } X \in \mathfrak{U}_s([a,b]; \mu_1, \mu_2, \ldots, \mu_{s-1}).$$

To this end, note that there exists $\ell \in \{2, 3, ..., s-1\}$ such that for $E[Z_\beta^\ell] = \mu_\ell$ (otherwise by continuity of $\alpha \to E[Z_\alpha^k]$ one can find $\beta^* > \beta$, $Z_\beta^\ell \in \mathfrak{U}_s([a,b]; \mu_1, \mu_2, \ldots, \mu_{s-1})$, which is a contradiction). Take any $X \in \mathfrak{U}_s([a,b]; \mu_1, \mu_2, \ldots, \mu_{s-1})$ and assume that $VaR_p[X] > VaR_p^+[Z_\beta] = \mu_1 + \beta$. From the construction of Z_β it is immediately clear that $S^-(F_{Z_\beta} - F_X) \leqslant 1$ and that the last sign of $F_{Z_\beta} - F_X$ is a "+". If there is no crossing then $E[Z_\beta^\ell] = \mu_\ell < E[X^\ell]$ which contradicts with $X \in \mathfrak{U}_s([a,b]; \mu_1, \mu_2, \ldots, \mu_{s-1})$. If there is one crossing then $Z_\beta \preceq_{2-cx} X$ must hold. Since $x \mapsto x^\ell$ is convex, it must hold that $E[X^\ell] = E[Z_\beta^\ell] = \mu_\ell$. Invoking the convexity of $x \mapsto x^\ell$ again, this can only be true if Z_β and X have the same distribution, which is a contradiction with the assumption $VaR_p[X] > VaR_p^+[Z_\beta]$. Hence, $VaR_p[X] \leqslant \mu_1 + \beta$ must hold. The other inequality, namely $\mu_1 - \beta\frac{1-p}{p} \leqslant VaR_p[X]$, can be proven similarly.

5 Applications

Stochastic s-convex extrema are known to provide reasonably accurate bounds on several risk quantities, including ruin probabilities (Denuit and Lefevre (1997)) or zero-exponential utility premiums (Denuit (1999)). They have also been successfully applied to bound gambler's ruin probabilities (Hurlimann (2005)) or the possible extinction in a branching process (Sawaya and Klaere (1997)). These problems have in common that they involve bounds on expectations of higher order differentiable functions. By contrast, in this paper we essentially aim at finding bounds for distribution functions (i.e., expectations of step functions) and VaRs.

5.1 Model uncertainty of credit risk models

Financial institutions employ credit portfolio models to monitor the risk they run on their different lending activities. Such monitoring is of utmost importance, as credit risk is typically the most important risk driver of a bank, and failure of one or more banks can destabilize an entire economy. Hence, let us consider a portfolio sum $S = v_1 I_1 + v_2 I_2 + \cdots + v_n I_n$ in which the I_i $(i = 1, 2, \ldots, n)$ are the indicator variables taking value one in the case of a default (with probability q_i) and zero otherwise, and in which the v_i $(i = 1, 2, \ldots, n)$ are the net exposures upon default. In order to assess the portfolio risk, one essentially needs to know the probabilities that two or more loans default together, but these quantities are hard to compute due to lack of (sufficiently rich) default statistics. In other words, credit risk portfolio models are inherently subject to model uncertainty. In this section, we shed light on the magnitude of this uncertainty. For convenience, but without impact on the conclusions of the analysis, we focus on large homogeneous portfolios. Hence, let us denote $q_i = q$ and let $v_i = 1$ $(i = 1, 2, \ldots, n)$. We are interested in the properties of the variable $\frac{S}{n}$ for $n \to \infty$.

KMV model: The KMV model is the standard credit risk model used in the financial industry, which is also prominently present in the Solvency II and Basel III regulatory frameworks where it is used to when determining the capital requirements that insurers respectively banks need to hold as a safeguard against insolvency; see also McKinsey (2009), the Basel Committee on Banking Supervision (2010) and the Committee of Insurance and Occupational Pension Supervisors (2008). The basic idea of the KMV model is that default of a loan occurs when the assets of the underlying debtor are insufficient to meet the liabilities Assuming that the asset returns of the obligors are multivariate normally distributed (with correlation parameter ρ), one obtains for the asymptotic distribution of the portfolio loss $\frac{S}{n}$,

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{S}{n} \leqslant s\right) = \Phi(\frac{(1-\rho)^{\frac{1}{2}}\Phi^{-1}(s) - \Phi^{-1}(q)}{\sqrt{\rho}}), \tag{5.1}$$

where Φ is the distribution of the standard normal random variable N. Hence, in the KMV framework one obtains that, asymptotically, $\frac{S}{n} \sim \Phi(\mu + \sigma N)$ where $\mu = \frac{\Phi^{-1}(q)}{\sqrt{(1-\rho)}}$ and $\sigma = \frac{\sqrt{\rho}}{\sqrt{(1-\rho)}}$, i.e., one obtains a probit norm distribution for the portfolio loss with parameters that depend on q and ρ only (see also Vasicek (2001) or McNeil et al. (2005)).

Model uncertainty: The KMV framework makes the crucial assumption that (log) asset returns are multivariate normally distributed allowing to have full knowledge on the distribution. However, it is clear that lack of data makes it hard, if not impossible, to backtest this modeled distribution. In fact, it is more realistic to assume that only the first series of moments of the portfolio loss are known, and, thus, that there are several models (one of which is the KMV model) that are statistically indistinguishable in the sense that their loss distributions are consistent with the given moments. Therefore, it is of interest to find the maximum and the minimum possible VaR that can be attained by a model given the information on a limited number of moments.

Using Table 8.6 page 365 from McNeil et al. (2005), we fix $\mu = -1.71$ and $\sigma = 0.264$ (corresponding to a default probability $p = 0.049$ and an asset correlation $\rho = 0.0652$). We use Monte-Carlo simulation to obtain the higher order moments and we find $\mu_1 = 0.04913$, $\mu_2 = 0.003149$, $\mu_3 = 0.0002529$, $\mu_4 = 0.00002466$ and $\mu_5 = 0.000002840$.

# of moments	$VaR_{70\%}$	$VaR_{90\%}$	$VaR_{95\%}$	$VaR_{99.5\%}$
1	(0.0000 ; 0.1638)	(0.0000 ; 0.4913)	(0.0000 ; 0.9826)	(0.0444 ; 1.0000)
2	(0.0314 ; 0.0905)	(0.0401 ; 0.1305)	(0.0429 ; 0.1673)	(0.0472 ; 0.4316)
3	(0.0315 ; 0.0903)	(0.0457 ; 0.1206)	(0.0508 ; 0.1424)	(0.0588 ; 0.2597)
4	(0.0318 ; 0.0890)	(0.0459 ; 0.1205)	(0.0603 ; 0.1362)	(0.0831 ; 0.1995)
5	(0.0347 ; 0.0836)	(0.0469 ; 0.1200)	(0.0610 ; 0.1358)	(0.0932 ; 0.1897)

Table 5.1: We compute the minimum and maximum possible values of VaR assuming that only the first n moments of the portfolio loss are known ($n = 1, 2, 3, 4$ or 5). The VaR bounds are reported as (;) in which the first value is the lower bound and the second value is the upper bound. The VaR numbers that are obtained when using the KMV model are $\text{VaR}_{70\%} = 0.0580$, $\text{VaR}_{90\%} = 0.0851$, $\text{VaR}_{95\%} = 0.1010$ and $\text{VaR}_{99.5\%} = 0.1515$.

We observe from Table 5.1 that VaR assessments of credit risk portfolio models are subject to a huge amount of model risk when only the first moment is known (arising from knowledge of the marginal distributions), but reduces significantly when the second moment as source of dependence information is available. Hence, these results comply well with the findings of Bernard et al. (2013) who already indicated that adding dependence information can sharpen significantly the VaR bounds. Having information on higher order moments reduces the model risk further, but the speed of sharpening the VaR bounds is rather slow. The model risk reduces significantly when lowering the probability level used for assessing VaR. For example, reducing the probability level from 99.5% to 95% has a huge impact on the maximum possible and minimum possible values of VaR. Nevertheless, model risk remains a concern in all instances.

5.2 Model uncertainty of the collective risk model

The collective risk model is typically used in insurance to describe (non-life) insurance portfolios. It is also used to model credit risk portfolios (called the "actuarial" approach); see

CreditRisk$^+$ (Gordy (2000) or Vandendorpe et al (2008)). So, let N be an $\mathbb{N}$-valued random variable denoting the number of claims that arise from a portfolio of policies in a given time period and let $X_1, X_2, ...$ be a sequence of independent and identically distributed individual risks with known distributions F. Consider now the aggregate risk

$$S = \sum_{i=1}^{N} X_i$$

We want to compare the risk numbers obtained by using a fully specified model for S with its bounds that are derived assuming that only some moments of S are known.

Denote by $m_r = E[X_i^r]$. Assume that N and X_i are independent and assume that N has a Poisson distribution with parameter λ. Then we can compute the moments of S recursively using the following expression

$$\mu_r = E[S^r] = \lambda \sum_{k=0}^{r-1} \binom{r-1}{k} \mu_k m_{r-k}, \quad \mu_0 = 1.$$

For example, when N is Poisson distributed with mean λ and X_i are Exponentially distributed with parameter θ, then the moments are given by $m_r = \frac{r!}{\theta^r}$. In particular the average number of claims per period is $\frac{1}{\lambda}$ and the average claim value is $\frac{1}{\theta}$.

In the case of Exponential distribution, Klugman et al. (2012), Section 9.3, Example 9.8 page 229 on Exponential Severities, propose to use this formula to compute the distribution of the aggregate claims

$$\mathrm{P}[S \leqslant x] = \sum_{n=0}^{\infty} p_n \mathrm{P}[S \leqslant x] = p_0 + \sum_{n=1}^{\infty} p_n \Gamma(n; x\theta)$$

where $\Gamma(n; x) = 1 - \sum_{j=0}^{n-1} \frac{x^j e^{-x}}{j!}$ and $p_n = \mathrm{P}[N = n] = \frac{\lambda^n e^{-\lambda}}{n!}$. We find after inverting the cdf that when $\lambda = 1$ and $\theta = 10$, the VaR at 99% is 0.618.

We report upper and lower bounds on $VaR_{0.99}$ values when a given number of moments is known.

Moments	VaRmin	VaRmax	Moments	VaRmin	VaRmax
1	.05	10	5	.32	.85
2	.1	1.55	10	.41	.78
3	.24	.93	55	.53	.69
4	.31	.85	100	.56	.67

Table 5.2: We compute the minimum and maximum possible values of VaR at the 0.99 probability level for a collective risk model when n of its moments are known ($a = 0$ and $b = 30$). We assume that N is Poisson(1) distributed and the claim size is exponentially distributed with parameter $\theta = 10$.

From Table 5.2 we see that using a collective model for computing (VaR based) capital requirements of an insurance portfolio can be a slippery slope when there is no strong motivation for the particular model (here, a compound Poisson model with exponential like claims) other than mathematical convenience. If one only knows two moments of the (unknown) loss distribution then there are many models possible that are consistent with this information and the range of possible VaR numbers is wide. However, in line with the previous results on the assessment of model risk for a credit risk portfolio, we find that knowledge of the portfolio variance makes it possible to sharpen significantly the VaR bounds that arise when only the portfolio mean is given. In other words, adding dependence information reduces model error. Note that also in this case the marginal added value of having information on the higher order moments is decreasing with the order and model risk cannot be wiped out even when up to hundred moments are known.

6 Conclusion

In this paper, we study risk measurement of portfolios. If the marginal distributions of the risky components are known but their interdependence is not, then it is possible to identify models that give rise to the maximum and minimum possible values for VaR (see Puccetti and Rüschendorf (2012a) and Embrechts et al. (2013)). It appears that in many practical situations of interest these bounds are too wide to be very useful and dependence information should be useful to reducing them. In this context, using information on the higher order moments of the portfolio sum is a natural device to incorporate dependence information, but, unfortunately, doing so makes the problem hard to deal with. In this paper, we bypass the technical complexity by ignoring the knowledge of the marginal distributions and using the portfolio mean as source of information instead. Doing so leads to little loss of information while making it possible to derive bounds on distributions of portfolios and their VaRs. Indeed, we are able to present a new and simple technique to obtain upper and lower bounds on the distribution, the stop-loss premium and the Value-at-Risk of a given portfolio. We also propose a approach that allows dealing with additional statistical uncertainty on the imposed moment constraints.

Our results allow to assess the robustness of risk models and hence to address a growing concern of regulators. Assume for instance that an insurance company uses a certain model for setting capital requirements. The model is calibrated using statistical estimates of (higher order) moments. We are then able to derive the best and worst model that is consistent with this information, in the sense that these models provide the highest and lowest feasible value for the capital requirement. If the bounds diverge too much then this is a clear signal that the outcomes of the model should be considered with care (for that particular purpose), or, that the risk measure may not be suitable. We illustrate this feature by assessing the model uncertainty (i.e., the difference between upper and lower bounds) of a collective risk model. We observe that even when many moments are known, there may still be a lot of model uncertainty on the VaR of a portfolio especially at higher probability levels.

Supervisory authorities may find the results also useful for comparing risk models across institutions. For example, banks and insurers employ models to assess the riskiness of credit default portfolios. In this context the $k-$th moment ($k = 1, 2, ...$) of the portfolio is driven by the probabilities that k different obligors default together. It is then clear that it is very difficult to obtain reliable estimates for the third and perhaps even the second moment of the portfolio. Hence, one may expect that financial institutions will find very different risk numbers for portfolios with *similar* characteristics. This observation is no longer acceptable for the Basel Committee (2013) who insist on comparability to enhance a level playing field. The results of our paper allow to identify such issues of incomparability of risk models.

A Appendix: Requirements for $\mathcal{M}_{2m}([a,b];\mu_1,\mu_2,\dots,\mu_{2m-1})$ resp. $\mathcal{M}_{2m+1}([a,b];\mu_1,\mu_2,\dots,\mu_{2m})$ to be admissible as introduced in Section 2.1.

We need to introduce some moment matrices. To this end, let $\mu_0 \equiv 1$, and m a (strictly) positive integer. For a given sequence $\mu_0,\mu_1,\dots,\mu_{2m-1}$ of real numbers, define the $m\times m$ matrices

$$\boldsymbol{P}_{m-1} = \begin{pmatrix} \mu_0 & \mu_1 & \mu_2 & \cdots & \mu_{m-1} \\ \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_m \\ \mu_2 & \mu_3 & \mu_4 & \cdots & \mu_{m+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{m-1} & \mu_m & \mu_{m+1} & \cdots & \mu_{2m-2} \end{pmatrix}, \qquad \boldsymbol{Q}_{m-1} = \begin{pmatrix} \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_m \\ \mu_2 & \mu_3 & \mu_4 & \cdots & \mu_{m+1} \\ \mu_3 & \mu_4 & \mu_5 & \cdots & \mu_{m+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_m & \mu_{m+1} & \mu_{m+2} & \cdots & \mu_{2m-1} \end{pmatrix}.$$

For a given sequence $\mu_0,\mu_1,\dots,\mu_{2m}$ we also define in addition,

$$\boldsymbol{R}_{m-1} = \begin{pmatrix} \mu_2 & \mu_3 & \mu_4 & \cdots & \mu_{m+1} \\ \mu_3 & \mu_4 & \mu_5 & \cdots & \mu_{m+2} \\ \mu_4 & \mu_5 & \mu_6 & \cdots & \mu_{m+3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{m+1} & \mu_{m+2} & \mu_{m+3} & \cdots & \mu_{2m} \end{pmatrix}.$$

Further, for any square matrix $\boldsymbol{A}$, let $\|\boldsymbol{A}\|$ denote the determinant of $\boldsymbol{A}$. For $m = 1,2,\dots,$ consider the conditions

$$\|\boldsymbol{Q}_{m-1} - a\boldsymbol{P}_{m-1}\| > 0 \text{ and } \|b\boldsymbol{P}_{m-1} - \boldsymbol{Q}_{m-1}\| > 0, \tag{A.1}$$

and

$$\|\boldsymbol{P}_{m-1}\| > 0 \text{ and } \|-ab\boldsymbol{P}_{m-2} + (a+b)\boldsymbol{Q}_{m-2} - \boldsymbol{R}_{m-2}\| > 0. \tag{A.2}$$

The conditions (A.1) and (A.2) guarantee that the elements of the associated moment space have a minimum number of support points or, equivalently, that the moment space contains more than one element (and thus many elements). More precisely, let us denote for any random variable X, by supp(X) the support of X, i.e. the set of all possible values for X. Furthermore, for any set A, let $\#(A)$ denote the number of points in A. We can now state the following result (see also Proposition 3.1 in Denuit et al. (2000)):

(i) Let $m \geqslant 1$ be an integer and consider the moment space $\mathcal{M}_{2m}([a,b];\mu_1,\mu_2,\dots,\mu_{2m-1})$. Then $\mathcal{M}_{2m}([a,b];\mu_1,\mu_2,\dots,\mu_{2m-1})$ contains at least two elements $\Longleftrightarrow \#(supp(X)\setminus\{a\}) \geqslant m$ and $\#(supp(X)\setminus\{b\}) \geqslant m$ for all $X \in \mathcal{M}_{2m}([a,b];\mu_1,\mu_2,\dots,\mu_{2m-1}) \Leftrightarrow$ (A.1) holds for $k \leqslant m-1$ and (A.2) holds for $k \leqslant m-1$.

(ii) Let $m \geqslant 1$ be an integer and consider the moment space $\mathcal{M}_{2m+1}([a,b];\mu_1,\mu_2,\dots,\mu_{2m})$. Then $\mathcal{M}_{2m+1}([a,b];\mu_1,\mu_2,\dots,\mu_{2m})$ contains at least two elements $\Leftrightarrow \#(supp(X)) \geqslant m+1$ and $\#(supp(X)\setminus\{a,b\}) \geqslant m$ for all $X \in \mathcal{M}_{2m+1}([a,b];\mu_1,\mu_2,\dots,\mu_{2m}) \Leftrightarrow$ (A.1) holds for $k \leqslant m-1$ and (A.2) holds for $k \leqslant m$.

To understand why these statements follow from the conditions (A.1) and (A.2) indeed, one may recall from Proposition A.1 in Denuit et al. (2000) that the representations

$$\begin{aligned}\big\|\boldsymbol{Q}_{m-1}-a\boldsymbol{P}_{m-1}\big\| &= \frac{1}{m!}E\Big[\prod_{i=0}^{m-1}(X_i-a)\prod_{0\leqslant j<k\leqslant m-1}(X_j-X_k)^2\Big]\\ \big\|b\boldsymbol{P}_{m-1}-\boldsymbol{Q}_{m-1}\big\| &= \frac{1}{m!}E\Big[\prod_{i=0}^{m-1}(b-X_i)\prod_{0\leqslant k<l\leqslant m-1}(X_j-X_k)^2\Big]\end{aligned}$$

hold for any $m\geqslant 1$, where $X_0,X_1,\ldots,X_{m-1}$ are independent and identically distributed random variables belonging to $\mathcal{M}_{2m}([a,b];\mu_1,\mu_2,\ldots,\mu_{2m-1})$. Also, the representations

$$\begin{aligned}\big\|\boldsymbol{P}_m\big\| &= \frac{1}{(m+1)!}E\Big[\prod_{0\leqslant j<k\leqslant m}(X_j-X_k)^2\Big]\\ \big\|-ab\boldsymbol{P}_{m-1}+(a+b)\boldsymbol{Q}_{m-1}-\boldsymbol{R}_{m-1}\big\| &= \frac{1}{m!}E\Big[\prod_{i=0}^{m-1}(X_i-a)(b-X_i)\prod_{0\leqslant j<k\leqslant m-1}(X_j-X_k)^2\Big]\end{aligned}$$

hold for any $m\geqslant 1$, where $X_0,X_1,\ldots,X_{m-1}$ are independent and identically distributed random variables belonging to $\mathcal{M}_{2m+1}([a,b];\mu_1,\mu_2,\ldots,\mu_{2m})$.

Now, imposing $\big\|\boldsymbol{P}_m\big\|>0$ excludes all the discrete random variables with less than $m+1$ support points as we must have $X_j=X_k$ for at least one pair (j,k) in this case. Similarly, imposing $\big\|-ab\boldsymbol{P}_{m-1}+(a+b)\boldsymbol{Q}_{m-1}-\boldsymbol{R}_{m-1}\big\|>0$ excludes discrete random variables with atoms a, b and less than m other support points in (a,b), for the same reason.

□

References

Basel Committee on Banking Supervision (2010). Basel III: A global regulatory framework for more resilient banks and banking systems.

Bernard, C., Rüschendorf, L., Vanduffel S. (2013). Value-at-Risk bounds with variance constraints. *Working paper available at SSRN: http://ssrn.com/abstract=2342068.*

Bloomberg http://www.bloomberg.com/news/2013-06-23/-window-dressing-undermines-bank-risk-weight-trust-bis.html

Basel Committee on Banking Supervision (2013). The regulatory framework: balancing risk sensitivity, simplicity and comparability. Consultative Document July 2013.

Chernih, A., Henrard, L., Vanduffel, S. (2010). Reconciling asset correlations. *The Journal of Risk Model Validation* 4, 47–64.

Committee of Insurance and Occupational Pension Supervisors (CEIOPS) (2008). Quantitative Impact Study 4. Available at www.ceiops.org.

Cheung, K.C., Vanduffel, S. (2013). Bounds for sums of random variables when the marginals and the variance of the sum are known *The Scandinavian Actuarial Journal* 2013, 103-118.

Denuit, M. (1999). The exponential premium calculation principle revisited. *ASTIN Bulletin* 29, 215-226.

Denuit, M., Genest, C., Marceau, É. (1999). Stochastic bounds on sums of dependent risks. *Insurance: Mathematics and Economics* 25, 85–104.

Denuit, M., Lefevre, Cl. (1997). Some new classes of stochastic order relations among arithmetic random variables, with applications in actuarial sciences. *Insurance: Mathematics and Economics* 20, 197-214.

Denuit, M., De Vijlder, F.E., Lefevre, Cl. (1999). Extremal generators and extremal distributions for the continuous s-convex stochastic orderings. *Insurance: Mathematics and Economics* 24, 201-217.

Denuit, M., Lefevre, Cl., Shaked, M. (1998). The s-convex orders among real random variables, with applications. *Mathematical Inequalities and Their Applications* 1, 585-613.

Ekern, S. (1980). Increasing n-th degree risk. *Economics Letters*, 6, 329-333.

Denuit, M., Lefevre, Cl. Shaked, M. (2000). S-convex approximations. *Advances in Applied Probability* 32, 994-1010.

De Schepper, A., Heijnen, B. (2010). How to estimate the Value-at-Risk under incomplete information. *Journal of Computational and Applied Mathematics* 233, 2213-2226.

De Vijlder, F.E. (1996). Advanced Risk Theory. A Self-Contained Introduction. Editions de l'Université Libre de Bruxelles - Swiss Association of Actuaries. Bruxelles.

Embrechts, P., Puccetti, G. (2006). Bounds for functions of dependent risks. *Finance and Stochastics* 10, 341–352.

Embrechts, P., Puccetti, G., (2009). Bounds for the sum of dependent risks having overlapping marginals. *Journal of Multivariate Analysis* 101, 177–190.

Embrechts, P., Puccetti, G., Rüschendorf, L., (2013). Model uncertainty and VaR aggregation, *Journal of Banking and Finance* 37, 2750–2764.

Gordy, M. (2000). A comparative anatomy of credit risk models. *Journal of Banking and Finance* 24, 119–149.

Huber, P.J., Ronchetti, E.M. (2009). Robust Statistics. Wiley Series in Probability and Statistics.

Hürlimann, W. (1998). Extremal Moment Methods and Stochastic Orders. Application in Actuarial Science. Monograph manuscript (available from the author).

Hürlimann, W. (2002). Analytical bounds for two Value-at-Risk functionals. *Astin Bulletin* 32, 235-265

Hürlimann, W. (2005). Improved analytical bounds for gambler's ruin probabilities. *Methodology and Computing in Applied Probability* 7, 79–95.

Jansen, K., Haezendonck, J., Goovaerts, M.J. (1986). Upper bounds on stop-loss premiums in case of known moments up to the fourth order. *Insurance: Mathematics and Economics* 5, 315-334.

Kaas, R., Goovaerts, M.J. (1986a). Best bounds for positive distributions with fixed moments. *Insurance: Mathematics and Economics*, 5, 87-92.

Kaas, R., Goovaerts, M.J. (1986b). Extremal values of stop-loss premiums under moment constraints. *Insurance: Mathematics and Economics* 5, 279-283.

Karlin, S. (1968). Total Positivity. Stanford University Press. Stanford. California.

Klugman, Stuart A., Harry H. Panjer, and Gordon E. Willmot (2012). Loss Models: From Data to Decisions. 3rd edition. John Wiley & Sons.

McKinsey (2009). Best Practices for estimating Ecocomic Capital. *Mc Kinsey's Working Papers on Risk*, 11, April 2009.

McNeil, A. J., Frey, R., Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques, Tools.* Princeton, NJ: Princeton University Press.

Puccetti, G., Rüschendorf, L., (2012a). Computation of sharp bounds on the distribution of a function of dependent risks. *Journal of Computational and Applied Mathematics* 236, 1833–1840.

Puccetti, G., Rüschendorf L. (2012b). Asymptotic equivalence of conservative VaR- and ES-based capital charges. *Journal of Risk*, in press.

Rüschendorf, L. (1982). Random variables with maximum sums. *Advances in Applied Probability* 14, 623–632.

Sawaya, S., Klaere, S. (2013). Extinction in a branching process: Why some of the fittest strategies cannot guarantee survival. *Working paper*, available on http://arxiv.org/pdf/1209.2074.pdf.

Rüschendorf, L. (1991) Bounds for distributions with multivariate marginals. In: *Stochastic Orders and Decisions*, K. Mosler, M. Scarsini (eds.). IMS Lecture Notes **19**, 285–310.

Vandendorpe, A., Ho, N., Vanduffel, S., Van Dooren, P. (2008). On the parameterization of the CreditRisk+ model for estimating credit portfolio risk. *Insurance: Mathematics and Economics* 42, 736–745.

Vasicek, O. (2002). Loan portfolio value, *Risk Magazine.* December 2002, 160–162.