Probability, Uncertainty and Quantitative Risk
Vol. 10, No. 1, 2025, 59–66
https://www.aimsciences.org/PUQR
DOI:10.3934/puqr.2025004

An algorithm for the calculation of upper variance under multiple probabilities and its application to quadratic programming

Xinpeng Li[1,2,*], Miao Yu[1], Shiyi Zheng[3]

[1]*Research Center for Mathematics and Interdisciplinary Sciences, Shandong University, Qingdao* 266237, *China*
[2]*Frontiers Science Center for Nonlinear Expectations, Ministry of Education, Shandong University, Qingdao* 266237, *China*
[3]*School of Mathematics, Shandong University, Jinan* 250100, *China*

Email: *lixinpeng@sdu.edu.cn*, *yu-miao@mail.sdu.edu.cn*, *zhshiyi_1@163.com*

Abstract The concept of upper variance under multiple probabilities is defined through a corresponding minimax optimization problem. This study proposes a simple algorithm to solve this optimization problem exactly. Additionally, we provide a probabilistic representation for a class of quadratic programming problems, demonstrating the practical application of our approach.

Keywords Multiple probabilities, Quadratic programming, Sublinear expectation, Upper variance

2020 Mathematics Subject Classification 60E05, 90C20

1. Introduction

The concepts of upper and lower variances generalize the classical notion of variance to situations involving multiple probabilities or problems with uncertainty. For example, consider a random variable X representing daily stock return. Without loss of generality, we assume $X \sim N(0.1, 0.4)$ and $X \sim N(-0.1, 0.4)$ in bull and bear markets, respectively. The mean and variance of these normal distributions can be estimated from the stock's historical data. Although we may attempt to estimate the "risk" (variance) of the stock after one month, we do not know whether the stock market will be in a bull or bear state in the future. In this case, the upper and lower variances are useful for measuring risk. Specifically, let $\mathcal{P} = \{P_1, \cdots, P_K\}$ be a set of probability measures on measurable space $(\Omega, \mathcal{F})$. For each random variable X with $\max_{1\leqslant i\leqslant K} E_{P_i}[X^2] < \infty$, the upper and lower variances of X under $\mathcal{P}$ are respectively defined as follows:

$$\overline{V}(X) = \min_{\mu\in\mathbb{R}} \max_{1\leqslant i\leqslant K} E_{P_i}[(X-\mu)^2], \quad \underline{V}(X) = \min_{\mu\in\mathbb{R}} \min_{1\leqslant i\leqslant K} E_{P_i}[(X-\mu)^2].$$

Received 12 January 2024; Accepted 1 December 2024; Early access 16 January 2025
*Corresponding author

Walley [8] indicated that calculating the lower variance is relatively straightforward, whereas calculating the upper variance is more difficult (see Appendix G in [8]). In this study, we introduce a simple algorithm for calculating the upper variance $\overline{V}(X)$.

Upper and lower variances are widely used, especially in mathematical finance. Li et al. [1] used these variances to develop a general G-Var model with mean-uncertainty, which generalized the G-Var model with zero-mean proposed by Peng et al. [6] and Pei et al. [4]. In this study, we introduce a new application of the upper variance. Because the upper variance $\overline{V}(X)$ can be regarded as the supremum of the variance over the convex hull of $\mathcal{P}$ (see Theorem 2.3), it can be calculated using a quadratic programming problem on the unit simplex in $\mathbb{R}_+^K$ (see Proposition 4.1). The proposed algorithm also facilitates solving a class of quadratic programming problems, where the related optimal value corresponds to the upper variance under multiple probabilities.

The remainder of this paper is organized as follows. Section 2 recalls the concepts and properties of upper and lower variances. Section 3 presents the algorithm for calculating the upper variance. Section 4 explores the application of the upper variance in quadratic programming.

2. Preliminaries

We begin by recalling the preliminaries of the sublinear expectation theory introduced by Peng [5], which serves as a powerful tool for addressing problems involving uncertainty, such as problems under multiple probability measures.

Let $(\Omega, \mathcal{F})$ be a measurable space and $\mathcal{P}$ be a set of probability measures on $(\Omega, \mathcal{F})$ that characterizes Knightian uncertainty. We define the corresponding upper expectation $\mathbb{E}^{\mathcal{P}}$ as follows:

$$\mathbb{E}^{\mathcal{P}}[X] = \sup_{P \in \mathcal{P}} E_P[X].$$

Clearly, $\mathbb{E}^{\mathcal{P}}$ is a sublinear expectation that satifies the following properties:

(i) Monotonicity: $\mathbb{E}^{\mathcal{P}}[X] \leqslant \mathbb{E}^{\mathcal{P}}[Y]$, if $X \leqslant Y$;
(ii) Constant preserving : $\mathbb{E}^{\mathcal{P}}[c] = c, \forall c \in \mathbb{R}$;
(iii) Sub-additivity: $\mathbb{E}^{\mathcal{P}}[X+Y] \leqslant \mathbb{E}^{\mathcal{P}}[X] + \mathbb{E}^{\mathcal{P}}[Y]$;
(iv) Positive homogeneity: $\mathbb{E}^{\mathcal{P}}[\lambda X] = \lambda \mathbb{E}^{\mathcal{P}}[X], \forall \lambda \geqslant 0$.

We call $(\Omega, \mathcal{F}, \mathbb{E}^{\mathcal{P}})$ as the sublinear expectation space.

In this study, we denote $\mathrm{co}(\mathcal{P})$ as the convex hull of $\mathcal{P}$. It is evident that:

$$\mathbb{E}^{\mathcal{P}}[X] = \sup_{P \in \mathcal{P}} E_P[X] = \sup_{P \in \mathrm{co}(\mathcal{P})} E_P[X].$$

The upper expectation $\mathbb{E}^{\mathcal{P}}[X]$ and the lower expectation $-\mathbb{E}^{\mathcal{P}}[-X]$ of X are denoted by $\overline{\mu}_X$ and $\underline{\mu}_X$, respectively. These are referred to as the upper and lower means of X, respectively. The interval $[\underline{\mu}_X, \overline{\mu}_X]$ characterizes the mean-uncertainty of X, denoted by M_X.

Next, we recall the concept of upper and lower variances, first introduced by Walley [8] for bounded random variables under coherent prevision and later generalized by Li et al. [1].

Let $(\Omega, \mathcal{F}, P)$ be a probability space and X be a random variable with a finite second moment. The variance of X, denoted by $V_P(X)$, is defined as follows:

$$V_P(X) = E_P[(X - E_P[X])^2].$$

We note that

$$E_P[(X-\mu)^2] = V_P(X) + (E_P[X] - \mu)^2,$$

then we have

$$V_P(X) = \min_{\mu\in\mathbb{R}} E_P[(X-\mu)^2].$$

Using $\mathbb{E}^{\mathcal{P}}$ instead of E_P, we obtain the following definition of upper and lower variances.

Definition 2.1 *For a random variable X on the sublinear expectation space $(\Omega, \mathcal{F}, \mathbb{E}^{\mathcal{P}})$ with $\mathbb{E}^{\mathcal{P}}[X^2] < \infty$, the upper variance of X is defined as*

$$\overline{V}(X) := \min_{\mu\in M_X} \{\mathbb{E}^{\mathcal{P}}[(X-\mu)^2]\},$$

and the lower variance of X is

$$\underline{V}(X) := \min_{\mu\in M_X} \{-\mathbb{E}^{\mathcal{P}}[-(X-\mu)^2]\}.$$

Remark 2.2 *Because $\mathbb{E}^{\mathcal{P}}[(X-\mu)^2]$ is a strictly convex function of μ, there exists a unique $\mu^* \in M_X$ such that*

$$\overline{V}(X) = \mathbb{E}^{\mathcal{P}}[(X-\mu^*)^2].$$

We may then prove that

$$\overline{V}(X) = \min_{\mu\in\mathbb{R}}\{\mathbb{E}^{\mathcal{P}}[(X-\mu)^2]\}, \quad \underline{V}(X) = \min_{\mu\in\mathbb{R}}\{-\mathbb{E}^{\mathcal{P}}[-(X-\mu)^2]\}.$$

For more details, refer to studies by Walley [8] and Li et al. [1].

Using the minimax theorem, we obtain the following variance envelop theorem.

Theorem 2.3 *Let X be a random variable on the sublinear expectation space $(\Omega, \mathcal{F}, \mathbb{E}^{\mathcal{P}})$ with $\mathbb{E}^{\mathcal{P}}[X^2] < \infty$. Then*

(i) $\overline{V}(X) = \sup_{P\in \mathrm{co}(\mathcal{P})} V_P(X)$.

(ii) $\underline{V}(X) = \inf_{P\in \mathrm{co}(\mathcal{P})} V_P(X) = \inf_{P\in\mathcal{P}} V_P(X)$.

Proof We note that M_X is convex and compact, and $\mathrm{co}(\mathcal{P})$ is convex, and furthermore, $E_P[(X-\mu)^2]$ is an affine function of P and a convex function of μ, by minimax theorem in Sion [7], we have

$$\begin{aligned}\overline{V}(X) &= \min_{\mu\in M_X}\sup_{P\in\mathcal{P}} E_P[(X-\mu)^2] = \min_{\mu\in M_X}\sup_{P\in\mathrm{co}(\mathcal{P})} E_P[(X-\mu)^2] \\ &= \sup_{P\in\mathrm{co}(\mathcal{P})}\min_{\mu\in M_X} E_P[(X-\mu)^2] = \sup_{P\in\mathrm{co}(\mathcal{P})} V_P(X).\end{aligned}$$

It is also evident that

$$\begin{aligned}\underline{V}(X) &= \min_{\mu\in M_X}\inf_{P\in\mathcal{P}} E_P[(X-\mu)^2] = \min_{\mu\in M_X}\inf_{P\in\mathrm{co}(\mathcal{P})} E_P[(X-\mu)^2] \\ &= \inf_{P\in\mathrm{co}(\mathcal{P})}\min_{\mu\in M_X} E_P[(X-\mu)^2] = \inf_{P\in\mathrm{co}(\mathcal{P})} V_P(X),\end{aligned}$$

and

$$\underline{V}(X) = \min_{\mu\in M_X}\inf_{P\in\mathcal{P}} E_P[(X-\mu)^2] = \inf_{P\in\mathcal{P}}\min_{\mu\in M_X} E_P[(X-\mu)^2] = \inf_{P\in\mathcal{P}} V_P(X).$$

□

Remark 2.4 *Unlike the envelop theorems in Walley [8] and Li et al. [1], we do not require $\mathcal{P}$ to be weakly compact. However, the convexity of $\mathcal{P}$ is necessary, as illustrated in Example 2.5.*

Example 2.5 *Let X be normally distributed with $X \sim N(0.1, 0.4)$ under P_1 and $X \sim N(-0.1, 0.4)$ under P_2. Taking $\mathcal{P} = \{P_1, P_2\}$, we obtain*

$$\overline{V}(X) = V_{P^*}(X) = 0.41 > 0.4 = \max\{V_{P_1}(X), V_{P_2}(X)\},$$

where $P^* = \frac{1}{2}(P_1 + P_2)$ *and*

$$\underline{V}(X) = 0.4 = V_{P_1}(X) = V_{P_2}(X).$$

In this study, we focus on the case where $\mathcal{P}$ consists of a finite number of probability measures, i.e., $\mathcal{P} = \{P_1, \cdots, P_K\}$. In this case, the convex hull $\mathrm{co}(\mathcal{P})$ is given by

$$\mathrm{co}(\mathcal{P}) = \{P_{\boldsymbol{\lambda}} : P_{\boldsymbol{\lambda}} = \lambda_1 P_1 + \cdots + \lambda_K P_K, \quad \forall\, \boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_K)^T \in \Delta^K\},$$

where $\Delta^K = \{\boldsymbol{\lambda} \in \mathbb{R}^K : \lambda_1 + \cdots + \lambda_K = 1, \lambda_i \geqslant 0, 1 \leqslant i \leqslant K\}$.

As indicated by Walley [8], calculating the lower variance $\underline{V}(X)$ is usually easy using Theorem 2.3 because $\underline{V}(X) = \min_{1\leqslant i\leqslant K} V_{P_i}(X)$. For the upper variance $\overline{V}(X)$, there exists $\boldsymbol{\lambda}^* \in \Delta^K$ such that $\overline{V}(X) = V_{P_{\boldsymbol{\lambda}^*}}(X)$, but $P_{\boldsymbol{\lambda}^*}$ is not generally an extreme point (see Example 2.5). Therefore, the calculation of the upper variance $\overline{V}(X)$ is more challenging.

3. Calculation of the upper variance

In this section, we propose a simple algorithm to calculate the upper variance $\overline{V}(X)$ under the set of probability measures $\mathcal{P} = \{P_1, \cdots, P_K\}$.

Our goal is to solve the following minimax problem:

$$\overline{V}(X) = \min_{\mu\in M_X} \max_{1\leqslant i\leqslant K} E_{P_i}[(X-\mu)^2]. \tag{1}$$

We rewrite (1) as

$$\overline{V}(X) = \min_{\mu\in M_X} \max_{1\leqslant i\leqslant K} (\mu^2 - 2\mu_i\mu + \kappa_i), \tag{2}$$

where $\mu_i = E_{P_i}[X]$ and $\kappa_i = E_{P_i}[X^2]$, $1 \leqslant i \leqslant K$.

Without loss of generality, we assume $\mu_1 \leqslant \mu_2 \leqslant \cdots \leqslant \mu_K$.

To prove our main theorem, we introduce the following two lemmas. The first lemma characterizes the position of the optimal μ^* for (1). The second lemma calculates the upper variance for two probability measures.

Lemma 3.1. *The unique optimal* $\mu^* \in M_X$ *in (1) must satisfy one of the following two conditions:*

(i) μ^* *is the minimum of some parabola* $f_i(\mu) := \mu^2 - 2\mu_i\mu + \kappa_i$.

(ii) μ^* *is the intersection of two parabolas* f_i *and* f_j, *i.e.,* $f_i(\mu^*) = f_j(\mu^*)$.

Proof If μ^* is not the intersection of two parabolas, without loss of generality, we assume that

$$f_1(\mu^*) < f_2(\mu^*) < \cdots < f_K(\mu^*).$$

There exists a neighborhood O of μ^* such that

$$f_1(x) < f_2(x) < \cdots < f_K(x), \quad \forall x \in O.$$

Thus, we have

$$\min_{x\in O} \max_{1\leqslant i\leqslant K} f_i(x) = \min_{x\in O} f_K(x),$$

and μ^* is the minimum of f_K. □

Lemma 3.2 *The upper variance of* X *under* $\mathcal{P} = \{P_1, P_2\}$ *can be calculated as*

$$\overline{V}(X) = \max\{\kappa_1 - \mu_1^2, \kappa_2 - \mu_2^2, h(\mu_{12})\},$$

where

$$\mu_{12} = \begin{cases} \left(\mu_1 \vee \dfrac{\kappa_2 - \kappa_1}{2(\mu_2 - \mu_1)}\right) \wedge \mu_2, & \mu_1 < \mu_2, \\ \mu_1, & \mu_1 = \mu_2, \end{cases}$$

and $h(x) = x^2 - 2\mu_1 x + \kappa_1$.

Proof We complete the proof by considering two cases.

First, if $\mu_1 = \mu_2$, we have

$$g(\mu) := \max\{f_1(\mu), f_2(\mu)\} = \begin{cases} f_1(\mu), & \kappa_1 \geqslant \kappa_2, \\ f_2(\mu), & \kappa_1 < \kappa_2. \end{cases}$$

It is easy to check that

$$\overline{V}(X) = \min_{\mu \in M_X} g(\mu) = \begin{cases} \kappa_1 - \mu_1^2, & \kappa_1 \geqslant \kappa_2, \\ \kappa_2 - \mu_2^2, & \kappa_1 < \kappa_2. \end{cases}$$

Furthermore, since $h(\mu_{12}) = \kappa_1 - \mu_1^2$, we have $\overline{V}(X) = \max\{\kappa_1 - \mu_1^2, \kappa_2 - \mu_2^2, h(\mu_{12})\}$.

Second, if $\mu_1 < \mu_2$, the solution of the equation $f_1(\mu) = f_2(\mu)$ is

$$\hat{\mu} = \frac{\kappa_2 - \kappa_1}{2(\mu_2 - \mu_1)}.$$

We consider the following function

$$g(\mu) := \max\{f_1(\mu), f_2(\mu)\} = \begin{cases} f_2(\mu), & \mu < \hat{\mu}, \\ f_1(\mu), & \mu \geqslant \hat{\mu}. \end{cases}$$

We discuss the location of $\hat{\mu}$.

(1) If $\hat{\mu} \in [\mu_1, \mu_2]$, it is easy to verify that

$$\overline{V}(X) = \min_{\mu \in M_X} g(\mu) = f_2(\hat{\mu}) = f_1(\hat{\mu}) = h(\mu_{12}) = \max\{\kappa_1 - \mu_1^2, \kappa_2 - \mu_2^2, h(\mu_{12})\}.$$

(2) If $\hat{\mu} < \mu_1$, then $\kappa_1 - \mu_1^2 > \kappa_2 - \mu_2^2$ and $h(\mu_{12}) = f_1(\mu_1)$ since $\mu_{12} = \mu_1$, therefore,

$$\overline{V}(X) = \min_{\mu \in M_X} g(\mu) = \min_{\mu \in M_X} f_1(\mu) = f_1(\mu_1) = \max\{\kappa_1 - \mu_1^2, \kappa_2 - \mu_2^2, h(\mu_{12})\}.$$

(3) If $\hat{\mu} > \mu_2$, then we have $\kappa_2 - \mu_2^2 > \kappa_1 - \mu_1^2$ and

$$\kappa_2 - \mu_2^2 > \mu_2^2 - 2\mu_1\mu_2 + \kappa_1 = h(\mu_{12}),$$

and therefore,

$$\overline{V}(X) = \min_{\mu \in M_X} g(\mu) = \min_{\mu \in M_X} f_2(\mu) = f_2(\mu_2) = \max\{\kappa_1 - \mu_1^2, \kappa_2 - \mu_2^2, h(\mu_{12})\}.$$

□

Theorem 3.3 *The upper variance of* X *under* $\mathcal{P} = \{P_1, \cdots, P_K\}$ *can be calculated as*

$$\overline{V}(X) = \max\left\{ \max_{1 \leqslant i \leqslant K} (\kappa_i - \mu_i^2), \max_{1 \leqslant i < j \leqslant K} h_{ij}(\mu_{ij}) \right\},$$

where

$$\mu_{ij} = \begin{cases} \left(\mu_i \vee \dfrac{\kappa_j - \kappa_i}{2(\mu_j - \mu_i)}\right) \wedge \mu_j, & \mu_i < \mu_j, \\ \mu_i, & \mu_i = \mu_j, \end{cases} \quad 1 \leqslant i < j \leqslant K$$

and $h_{ij}(x) = x^2 - 2\mu_i x + \kappa_i$.

Proof For $1 \leqslant i < j \leqslant K$, let $\overline{V}_{ij}(X)$ denote the upper variance under the two probability measures P_i and P_j. Then, using Lemma 3.2, we obtain

$$\overline{V}_{ij}(X) = \max\{\kappa_i - \mu_i^2, \kappa_j - \mu_j^2, h_{ij}(\mu_{ij})\}.$$

It is obvious that $\overline{V}(X) \geqslant \overline{V}_{ij}(X)$ holds for any $1 \leqslant i < j \leqslant K$, and we obtain

$$\begin{aligned}\overline{V}(X) &\geqslant \max_{1\leqslant i<j\leqslant K}\left\{\max\{\kappa_i-\mu_i^2,\kappa_j-\mu_j^2,h_{ij}(\mu_{ij})\}\right\}\\ &=\max\left\{\max_{1\leqslant i\leqslant K}(\kappa_i-\mu_i^2),\max_{1\leqslant i<j\leqslant K}h_{ij}(\mu_{ij})\right\}. \end{aligned} \tag{3}$$

Let $\mu^*\in M_X$ be the optimal point in (1). According to Lemma 3.1, for Case (i), we note that $\kappa_i-\mu_i^2=\min_{\mu\in\mathbb{R}} f_i(\mu)$, $1\leqslant i\leqslant K$. Therefore, if Case (i) holds, we have

$$\overline{V}(X)\leqslant \max_{1\leqslant i\leqslant K}(\kappa_i-\mu_i^2).$$

For Case (ii), we note that each parabola f_i is determined by the corresponding underlying probability measure P_i. In this case, $\overline{V}(X)$ can be calculated under two probability measures P_i and P_j, $1\leqslant i<j\leqslant K$. Thus, we have

$$\overline{V}(X)\leqslant \max_{1\leqslant i<j\leqslant K}\overline{V}_{ij}(X).$$

Combining these two cases, it can be seen that

$$\overline{V}(X)\leqslant \max\left\{\max_{1\leqslant i\leqslant K}(\kappa_i-\mu_i^2),\max_{1\leqslant i<j\leqslant K}h_{ij}(\mu_{ij})\right\}. \tag{4}$$

Therefore, from (3) and (4), it follows that

$$\overline{V}(X)= \max\left\{\max_{1\leqslant i\leqslant K}(\kappa_i-\mu_i^2),\max_{1\leqslant i<j\leqslant K}h_{ij}(\mu_{ij})\right\}.$$

□

Corollary 3.4 *The upper variance of X under $\mathcal{P}=\{P_1,\cdots,P_K\}$ can be calculated as:*

$$\overline{V}(X)=\max_{1\leqslant i<j\leqslant K}\left\{\overline{V}_{ij}(X)\right\},$$

where $\overline{V}_{ij}(X)$ is the upper variance under P_i and P_j, $1\leqslant i<j\leqslant K$.

Next, we present our algorithm. Owing to Lemma 3.2, Theorem 3.3, and Corollary 3.4, our algorithm does not require the assumption $\mu_1\leqslant\mu_2\leqslant\cdots\leqslant\mu_K$.

Given a random variable X under $\mathcal{P}=\{P_1,\cdots,P_K\}$, we calculate $\mu_i=E_{P_i}[X]$ and $\kappa_i=E_{P_i}[X^2]$, $1\leqslant i\leqslant K$.

Algorithm:

Step (1): For each pair of probabilities P_i and P_j, $1\leqslant i<j\leqslant K$, we calculate μ_{ij} as

$$\mu_{ij}=\begin{cases}\left(\underline{\mu}_{ij}\vee\dfrac{\kappa_j-\kappa_i}{2(\mu_j-\mu_i)}\right)\wedge\overline{\mu}_{ij}, & \mu_i\neq\mu_j,\\ \mu_i, & \mu_i=\mu_j,\end{cases}$$

$$\underline{\mu}_{ij}=\mu_i\wedge\mu_j,\quad \overline{\mu}_{ij}=\mu_i\vee\mu_j.$$

Step (2): Output

$$\overline{V}(X)= \max\left\{\max_{1\leqslant i\leqslant K}(\kappa_i-\mu_i^2),\max_{1\leqslant i<j\leqslant K}h_{ij}(\mu_{ij})\right\},$$

where $h_{ij}(x)=x^2-2\mu_i x+\kappa_i$.

In addition, the lower variance is calculated as

$$\underline{V}(X)=\min_{1\leqslant i\leqslant K}(\kappa_i-\mu_i^2).$$

In practice, μ_i and κ_i can be easily estimated from the data. For example, let X be the daily return of one stock. In the real market, we can obtain the daily return data $\{x_i\}_{i\in I}$ (resp. $\{x_j\}_{j\in J}$) from the bull (resp. bear) market, where I (resp. J) denotes the periods of the bull (resp. bear) market. Then, we can estimate the sample means and variances as

$$\hat{\mu}_1 = \frac{\sum_{i\in I} x_i}{|I|}, \quad \hat{\mu}_2 = \frac{\sum_{j\in J} x_j}{|J|},$$

and

$$\hat{\sigma}_1^2 = \frac{\sum_{i\in I}(x_i - \hat{\mu}_1)^2}{|I|-1}, \quad \hat{\sigma}_2^2 = \frac{\sum_{j\in J}(x_j - \hat{\mu}_2)^2}{|J|-1}.$$

We then take $\mu_i = \hat{\mu}_i$ and $\kappa_i = \hat{\sigma}_i^2 + \mu_i^2$ $(i = 1, 2)$ to calculate the upper and lower variances.

4. Application: Quadratic programming

By Theorem 2.3, (1) is equivalent to the following convex quadratic programming problem.

Proposition 4.1 *Let $\overline{V}(X)$ be the upper variance of X under $\{P_1, \cdots, P_K\}$, then*

$$\overline{V}(X) = \max_{\boldsymbol{\lambda}\in\Delta^K} (\boldsymbol{\lambda}^T\boldsymbol{\kappa} - (\boldsymbol{\lambda}^T\boldsymbol{\mu})^2), \tag{5}$$

where $\boldsymbol{\kappa} = (E_{P_1}[X^2], \cdots, E_{P_K}[X^2])^T$, and $\boldsymbol{\mu} = (E_{P_1}[X], \cdots, E_{P_K}[X])^T$.

Proof By Theorem 2.3, we obtain

$$\overline{V}(X) = \max_{\boldsymbol{\lambda}\in\Delta^K} V_{P_{\boldsymbol{\lambda}}}(X).$$

It is easily observable that

$$V_{P_{\boldsymbol{\lambda}}}(X) = E_{P_{\boldsymbol{\lambda}}}[X^2] - (E_{P_{\boldsymbol{\lambda}}}[X])^2 = \boldsymbol{\lambda}^T\boldsymbol{\kappa} - (\boldsymbol{\lambda}^T\boldsymbol{\mu})^2.$$

□

As shown in (5), this is a convex quadratic programming problem. Numerous numerical algorithms are available to solve such problems, such as the polynomial-time interior-point algorithm (Nesterov and Nemirovskii [3]) and accelerated gradient method (Nesterov [2]). However, most existing algorithms provide only approximate solutions. In this section, we propose a simple method to solve such this problem exactly using upper variance under multiple probability distributions.

We now state the following theorem:

Theorem 4.2 *Given $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_K)^T \in \mathbb{R}^K$ and $\boldsymbol{\kappa} = (\kappa_1, \cdots, \kappa_K)^T \in \mathbb{R}^K$, the solution to the following maximization problem:*

$$V = \max_{\boldsymbol{\lambda}\in\Delta^K} (\boldsymbol{\lambda}^T\boldsymbol{\kappa} - (\boldsymbol{\lambda}^T\boldsymbol{\mu})^2)$$

is given by

$$V = \max\left\{\max_{1\leqslant i\leqslant K}(\kappa_i - \mu_i^2), \max_{1\leqslant i<j\leqslant K} h_{ij}(\mu_{ij})\right\},$$

where

$$\mu_{ij} = \begin{cases} \left(\underline{\mu}_{ij} \vee \dfrac{\kappa_j - \kappa_i}{2(\mu_j - \mu_i)}\right) \wedge \overline{\mu}_{ij}, & \mu_i \neq \mu_j, \\ \mu_i, & \mu_i = \mu_j, \end{cases}$$

$$\underline{\mu}_{ij} = \mu_i \wedge \mu_j, \quad \overline{\mu}_{ij} = \mu_i \vee \mu_j,$$

and $h_{ij}(x) = x^2 - 2\mu_i x + \kappa_i$.

If there exists i_0 such that $V=\kappa_{i_0}-\mu_{i_0}^2$, then the optimal $\boldsymbol{\lambda}^*$ is given by $\lambda_{i_0}^*=1$ and $\lambda_j^*=0$, $j\neq i_0$. Otherwise, there exists $1\leqslant i_0<j_0\leqslant K$ such that $V=h_{i_0j_0}(\mu_{i_0j_0})$. The optimal $\boldsymbol{\lambda}^*$ is then given by $\lambda_{i_0}^*=\frac{\mu_{j_0}}{\mu_{j_0}-\mu_{i_0}}+\frac{\kappa_{i_0}-\kappa_{j_0}}{2(\mu_{i_0}-\mu_{j_0})^2}$, $\lambda_{j_0}^*=1-\lambda_{i_0}^*$, and $\lambda_j^*=0$, $j\neq i_0,j_0$.

Proof We provide only a sketch of the proof of this theorem. Let $C=\min_{1\leqslant i\leqslant K}\{\kappa_i-\mu_i^2\}$.

(1) $C>0$. In this case, (2) is equivalent to (1). We can calculate V by Theorem 3.3 and Corollary 3.4. Now we consider the optimal $\boldsymbol{\lambda}^*$.

If there exists i_0 such that $V=\kappa_{i_0}-\mu_{i_0}^2$, then it is clear that $\lambda_{i_0}^*=1$ and $\lambda_j^*=0$ for $j\neq i_0$.

If there exists $1\leqslant i_0<j_0\leqslant K$ such that $V=h_{i_0j_0}(\mu_{i_0j_0})$, we consider the following maximize problem:

$$\max_{\lambda_{i_0}+\lambda_{j_0}=1}\left\{\lambda_{i_0}\kappa_{i_0}+\lambda_{j_0}\kappa_{j_0}-(\lambda_{i_0}\mu_{i_0}+\lambda_{j_0}\mu_{j_0})^2\right\}. \tag{6}$$

The optimal solution of (6) is

$$\lambda_{i_0}=\frac{\mu_{j_0}}{\mu_{j_0}-\mu_{i_0}}+\frac{\kappa_{i_0}-\kappa_{j_0}}{2(\mu_{i_0}-\mu_{j_0})^2},$$

and

$$V=\frac{\mu_{j_0}\kappa_{i_0}-\mu_{i_0}\kappa_{j_0}}{\mu_{j_0}-\mu_{i_0}}+\frac{(\kappa_{i_0}-\kappa_{j_0})^2}{4(\mu_{i_0}-\mu_{j_0})^2}=h_{i_0j_0}(\mu_{i_0j_0}).$$

Similar to the proof of Case (ii) in Theorem 3.3, we know in this case that $0\leqslant\lambda_{i_0}\leqslant 1$.

(2) $C\leqslant 0$. In this case, we define $\bar{\boldsymbol{\kappa}}=\boldsymbol{\kappa}-C+1$. Now $\bar{\kappa_i}-\mu_i^2>0$, $1\leqslant i\leqslant K$, and noting that $\boldsymbol{\lambda}^T\bar{\boldsymbol{\kappa}}=\boldsymbol{\lambda}^T\boldsymbol{\kappa}-C+1$, the optimal $\boldsymbol{\lambda}^*$ is the same as in the case where $C>0$. □

Remark 4.3 *If $\kappa_i-\mu_i^2>0$, $1\leqslant i\leqslant K$, then V can be regarded as the upper variance of a random variable X under the set of probability measures $\{P_1,\cdots,P_K\}$ with $E_{P_i}[X]=\mu_i$ and $E_{P_i}[X^2]=\kappa_i$, $1\leqslant i\leqslant K$. Specifically, we can assume that X is normally distributed under each P_i with $X\sim N(\mu_i,\kappa_i-\mu_i^2)$, $1\leqslant i\leqslant K$.*

Acknowledgements

This work is supported by the NSF of China (Grant Nos. 12326603 and 11601281), NSF of Shandong Province (Grant No. ZR2021MA018), National Key R&D Program of China (Grant No.2018YFA0703900), the Qilu Young Scholars Program of Shandong University, and the Fundamental Research Funds for the Central Universities.

References

[1] Li, S., Li, X. and Yuan, G. X., Upper and lower variances under model uncertainty and their applications in finance, International Journal of Financial Engineering, 2022, 9(1): 2250007.

[2] Nesterov, Y., Introductory Lectures on Convex Optimization: A Basic Course, Springer, New York, 2003.

[3] Nesterov, Y. and Nemirovskii, A., Interior-Point Polynomial Algorithms in Convex Programming, Society for Industrial and Applied Mathematics, 1987

[4] Pei, Z., Wang, X., Xu, Y. and Yue, X., A worst-case risk measure by G-VaR, Acta Mathematicae Applicatae Sinica, English Series, 2021, 37(2): 421–440.

[5] Peng, S., Nonlinear Expectations and Stochastic Calculus under Uncertainty, Springer, Berlin, Heidelberg, 2019.

[6] Peng, S., Yang, S. and Yao, J., Improving value-at-risk prediction under model uncertainty, Journal of Financial Econometrics, 2023, 21: 228–259.

[7] Sion, M., On general minimax theorems, Pacific Journal of Mathematics, 1958, 8(1): 171–176.

[8] Walley, P., Statistical Reasoning with Imprecise Probabilities, Chapman and Hall, 1991.